名古屋大学　大学院理学研究科　物理学教室　最終講義
再収録　November 18, 2022

計算機シミュレーションによる研究を振り返って（その４）

岡本 祐幸（おかもと ゆうこう）

名古屋大学
情報基盤センター 大規模計算支援環境研究部門
国際本部 グローバル・エンゲージメントセンター
e-mail: okamoto{a}cc.nagoya-u.ac.jp
URL: https://yuko-okamoto.github.io/homepage/index.shtml

研究内容の話に戻ります。

研究内容については、今年（2022年）３月１４日（月）に日本生物物理学会のサブグループ「生体分子シミュレーション・モデリング」の第２回研究会で、「拡張アンサンブル法による生体分子シミュレーションによる研究を振り返って」というタイトルで講演させて頂き、その動画が既に以下のYoutubeサイト

招待講演 岡本祐幸先生（2022/3/14）
https://www.youtube.com/watch?v=qmPxJG_gT10

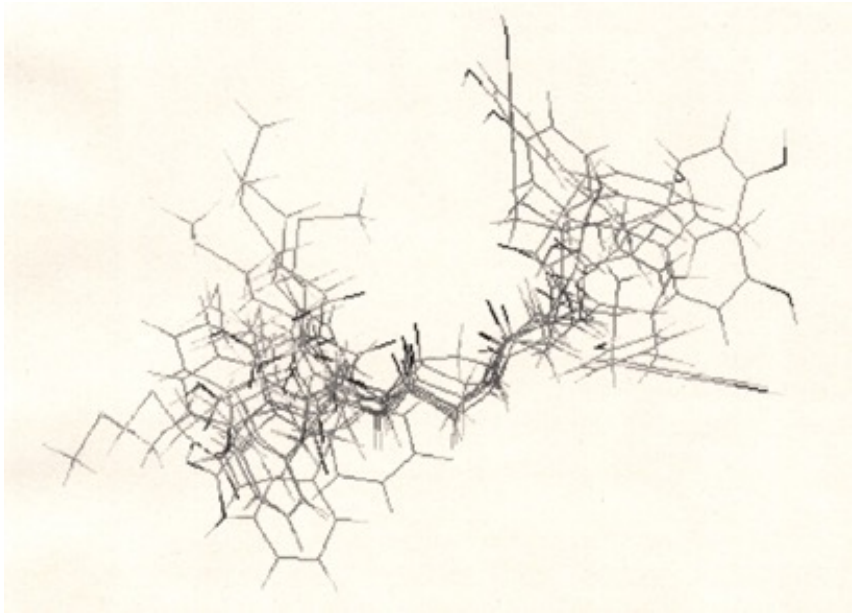
にアップロードされています。

今回、もし時間切れとなり、研究内容を十分お話しできませんでしたら、このYoutube動画をご覧頂けると幸いです。

Met-enkephalin in aqueous solution (RISM) (Level 3)

In gas phase, the peptide is compact due to
intra-chain hydorgen bonds (β-hairpin structures).
In aqueous solution, however, the peptide is extended:
In agreement with NMR experiments.

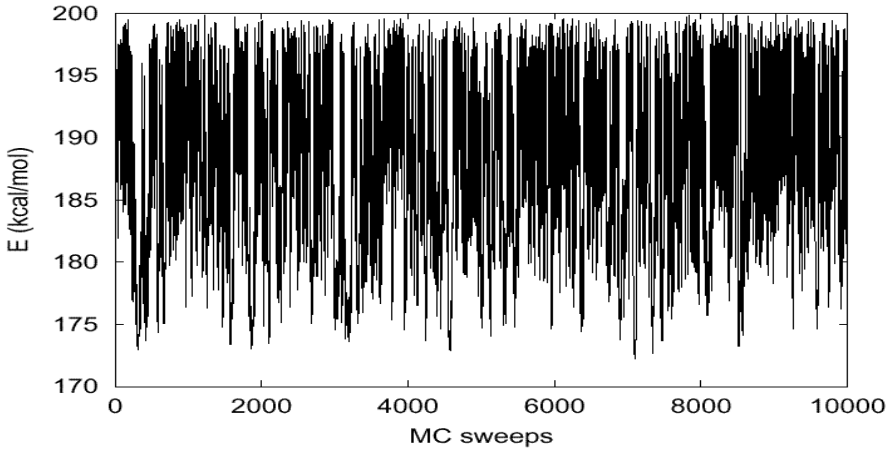
8 low-energy conformations from
Simulated Annealing (SA) runs
are superposed.

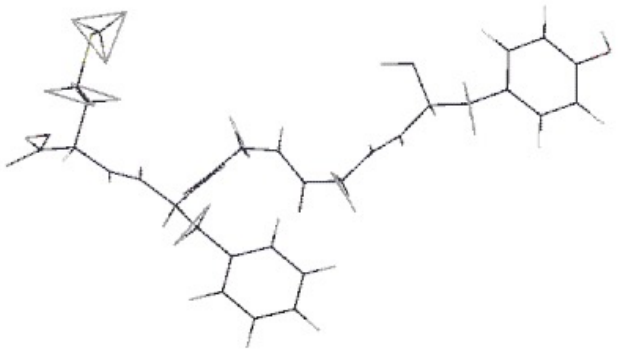


M. Kinoshita, Y. O. & F. Hirata, *J. Am. Chem. Soc.* **120**, 1855 (1998).

Met-enkephalin in aqueous solution (RISM) (Level 3)

- Multicanonical Simulation



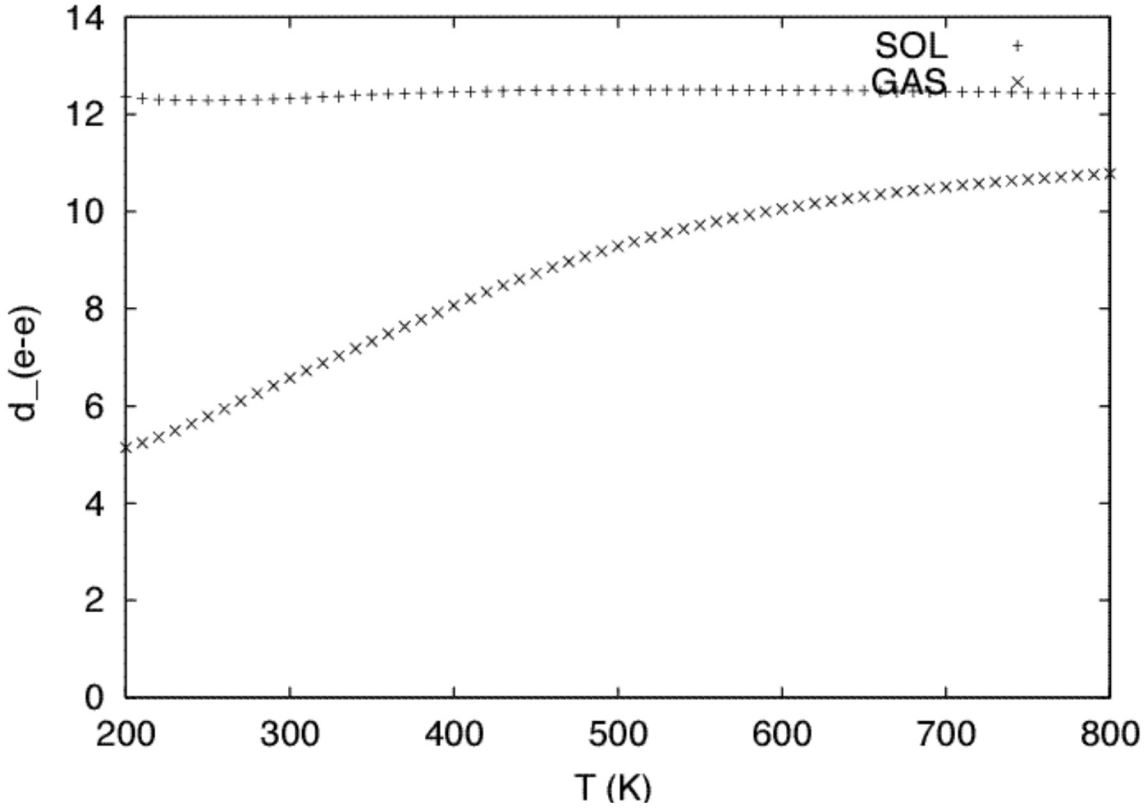


Simulation and movie by A. Mitsutake

In aqueous solution, the peptide is extended:
In agreement with NMR experiments.

A. Mitsutake, M. Kinoshita, Y. O. & F. Hirata,
Chem. Phys. Lett. **329**, 295 (2000).
See also: A. Mitsutake, M. Kinoshita, Y. O. & F. Hirata,
J. Phys. Chem. B **108**, 19002 (2004).

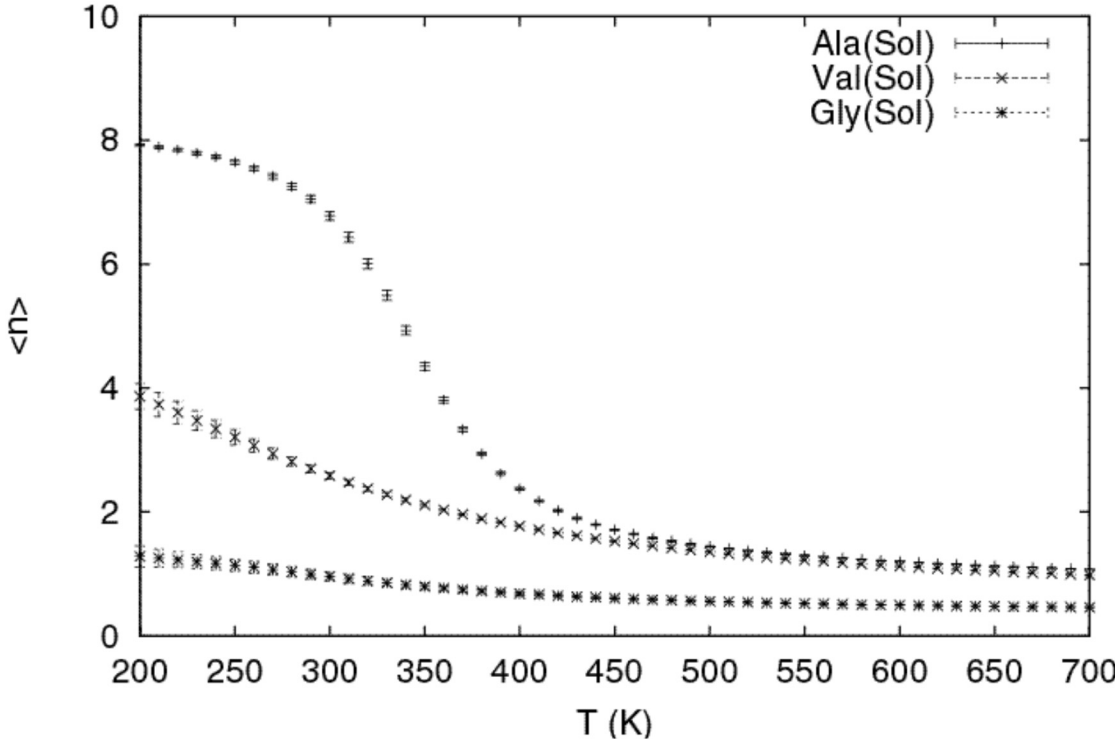
Average End-to-End Distance of Met-enkephain in Gas Phase and in Aquesou Solution (RISM): ECEPP/2



A. Mitsutake, M. Kinoshita, Y. O. & F. Hirata, *Chem. Phys. Lett.* **329**, 295 (2000).
See also: A. Mitsutake, M. Kinoshita, Y. O. & F. Hirata, *J. Phys. Chem. B* **108**, 19002 (2004).

Helix Propensities of Homopolymers: $(Ala)_{10}$, $(Val)_{10}$ & $(Gly)_{10}$

Average Helicity (Level 2)



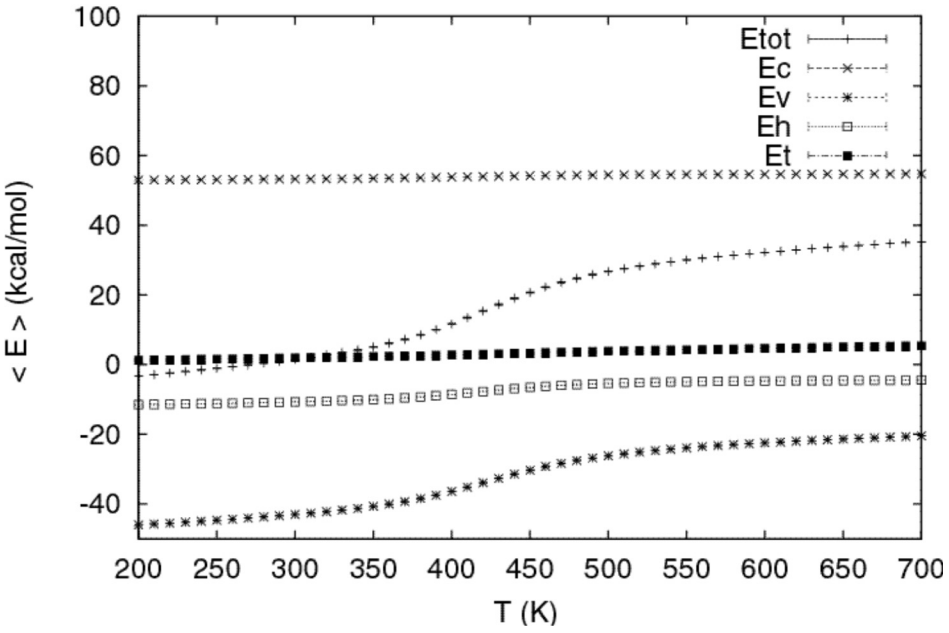
Level 0: Y.O. & U. Hansmann, *J. Phys. Chem.* **99**, 11276 (1995).

Level 2: A. Mitsutake & Y.O., *Chem. Phys. Lett.* **309**, 95 (1999).

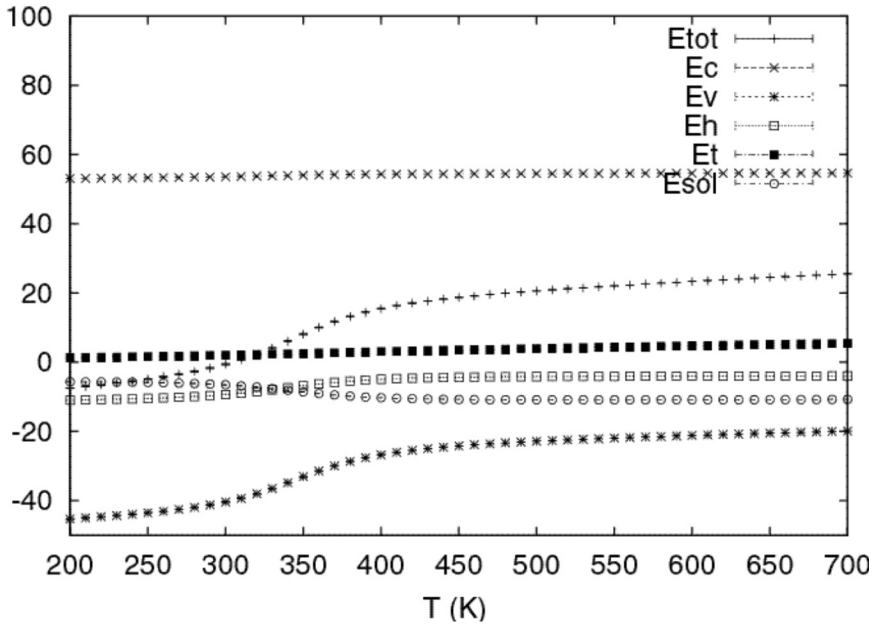
A. Mitsutake & Y.O., *J. Chem. Phys.* **112**, 10638 (2000).

Average Energy of (Ala)10

Level 0 (Gas Phase)

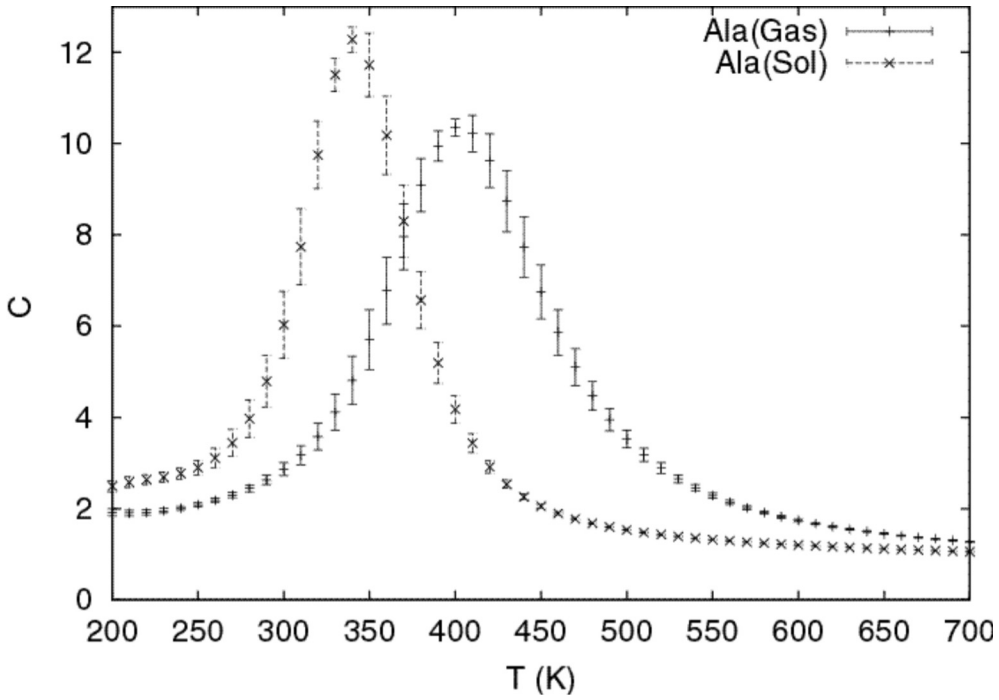


Level 2



A. Mitsutake & Y.O., *Chem. Phys. Lett.* **309**, 95 (1999).

Specific Heat of (Ala)10



Helix-Coil Transition Temperatures:

T = 420 K (Level 0)

T = 340 K (Level 2)

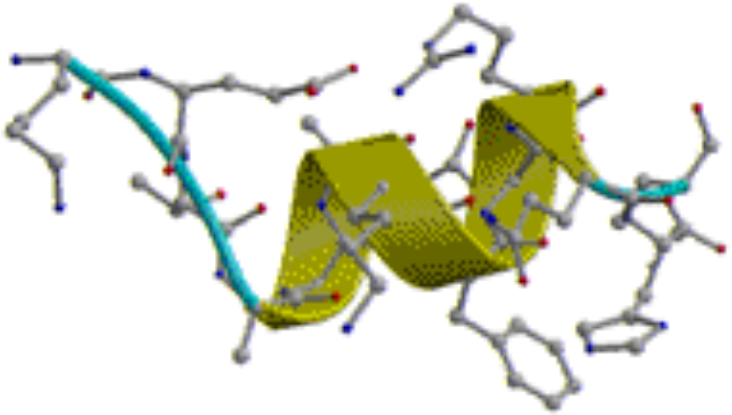
A. Mitsutake & Y.O., *Chem. Phys. Lett.* **309**, 95 (1999).

Importance of Solvent Effects

- Level 1 Distance-Dependent Dielectric Function $\varepsilon(r)$
  (Crude but computationally cheap)
- Level 2 Term Proportional to Solvent Accessible Surface Area
- Level 3 Explicit Solvent Molecules Included
  (Accurate but time-consuming)
  * Statistical Mechanical Theory of Liquids, etc.
    Reference Interaction Site Model (RISM)
    Scaled-Particle Theory + Poisson-Boltzmann
  * Direct Inclusion of Water Molecules

We compared the lowest-energy conformations obtained from simulations with the native conformations (more accurate solvation theory gives better agreement).

C-Peptide (N = 13): X-Ray

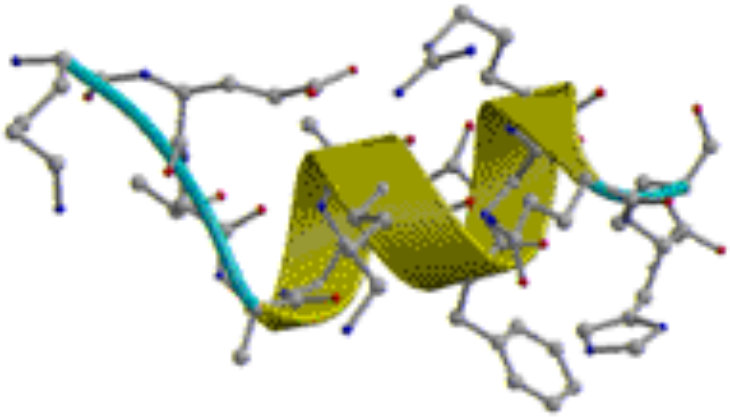


C-Peptide of Ribonuclease A

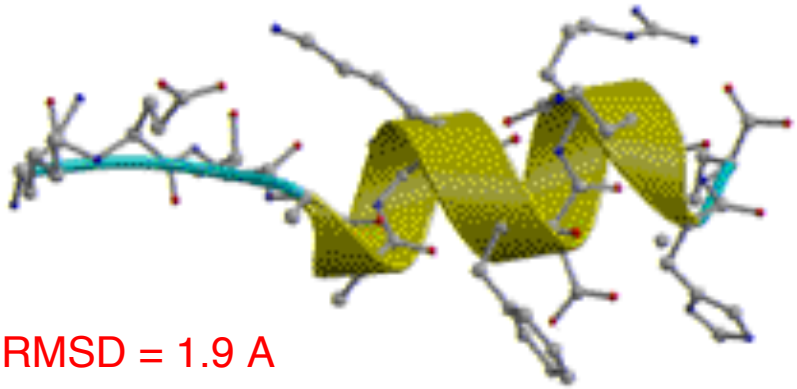
Amino-Acid Sequence:

Lys-Glu-Thr-Ala-Ala-Ala-Lys-Phe-Glu-Arg-Gln-His-Met

C-Peptide (N = 13): X-Ray



Level 0 (Gas Phase)

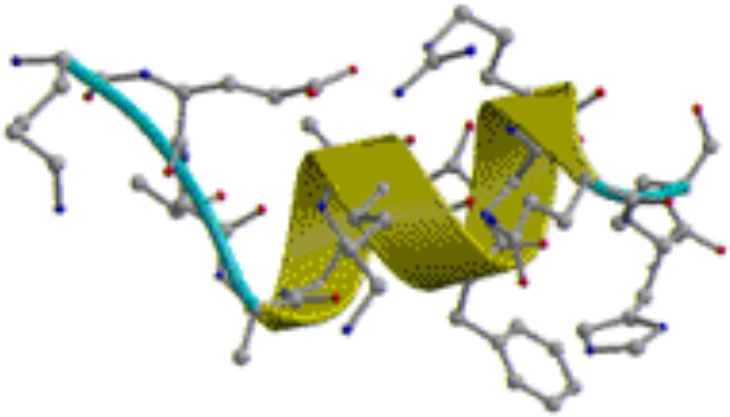


RMSD = 1.9 A

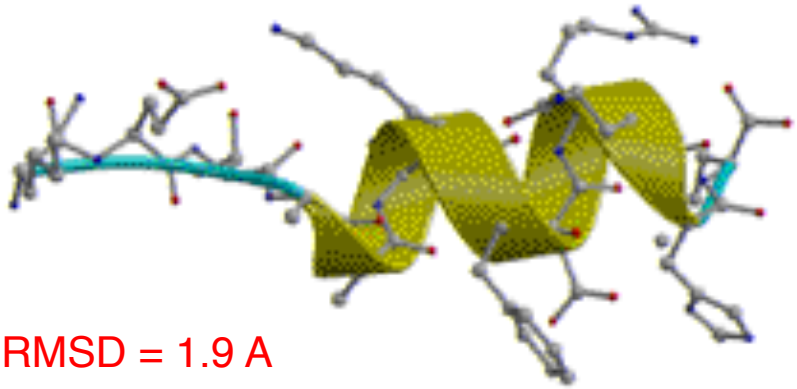
H. Kawai, Y.O., M. Fukugita, T. Nakazawa & T. Kikuchi, *Chem. Lett.* **1991**, 213.

Y.O., M. Fukugita, T. Nakazawa & H. Kawai, *Protein Eng.* **4**, 639 (1991).

C-Peptide (N = 13): X-Ray

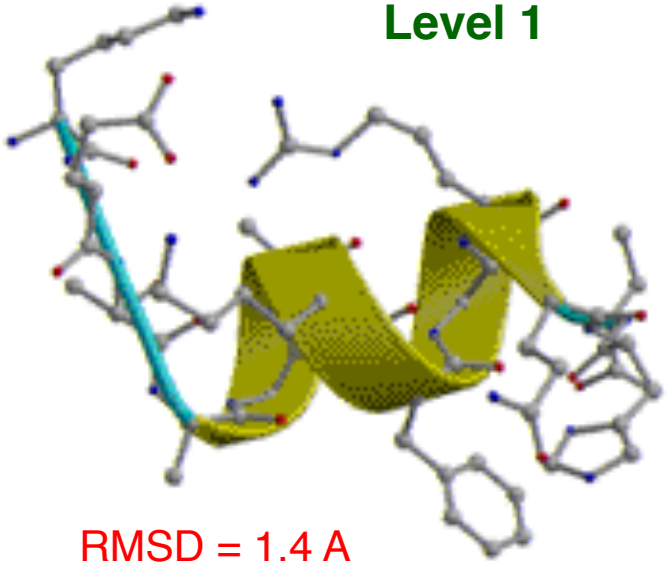


Level 0 (Gas Phase)



RMSD = 1.9 A

Level 1

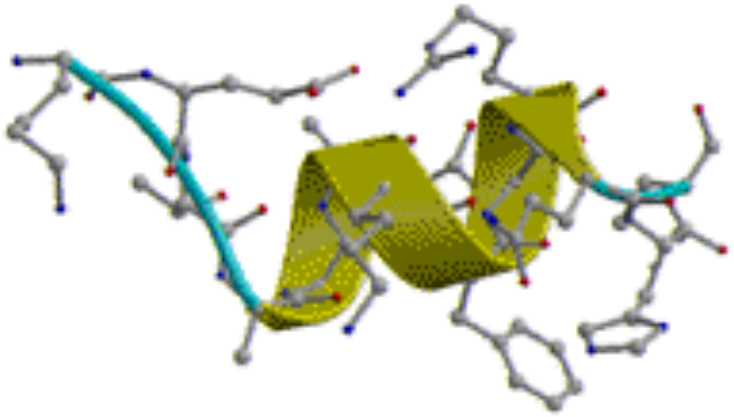


RMSD = 1.4 A

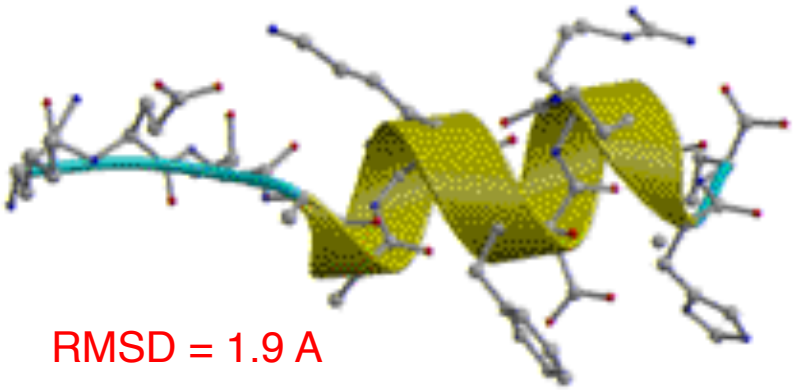
U. Hansmann & Y.O., *J. Phys. Chem. B* **102**, 653 (1998).

U. Hansmann & Y.O., *J. Phys. Chem. B* **103**, 1595 (1999).

C-Peptide (N = 13): X-Ray

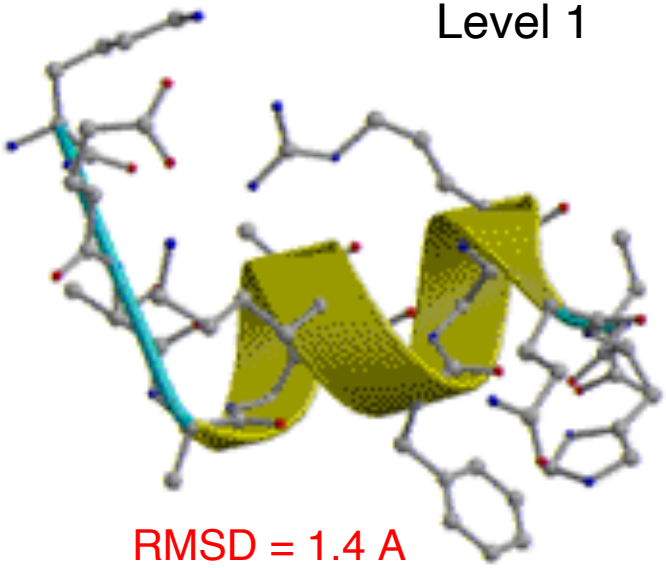


Level 0 (Gas Phase)



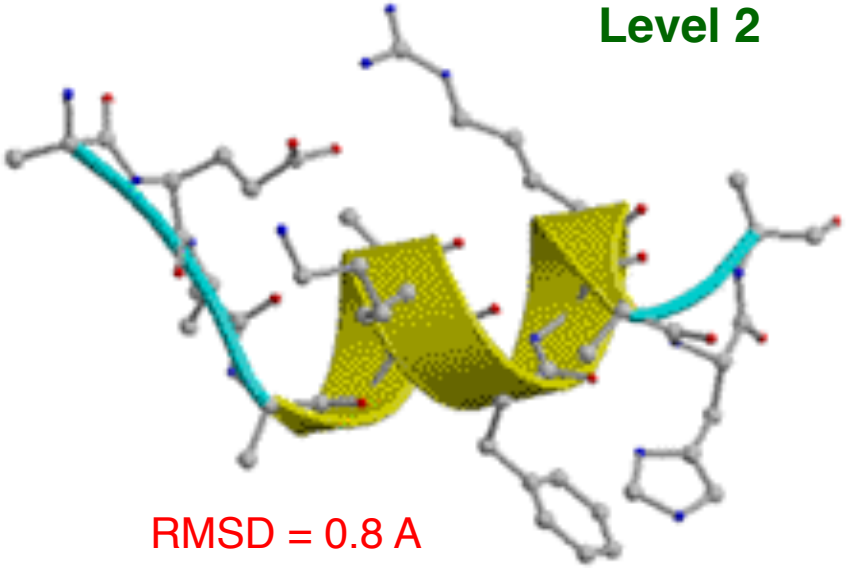
RMSD = 1.9 A

Level 1



RMSD = 1.4 A

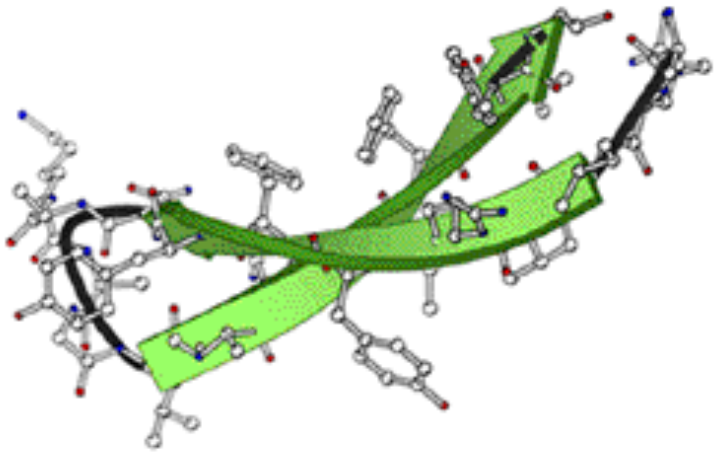
Level 2



RMSD = 0.8 A

M. Masuya & Y.O., unpublished.

BPTI(16-36) (N = 21): X-Ray



Fragment of Bovine Pancreatic Trypsin Inhibitor

BPTI(16-36)

Amino-Acid Sequence:

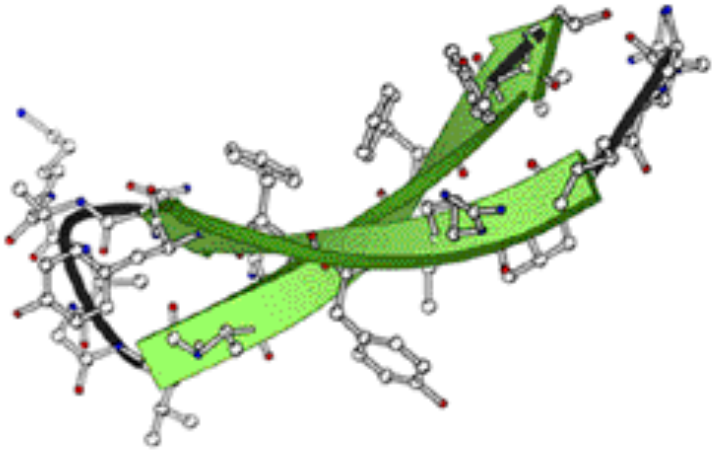
Ala-Arg-Ile-Ile-Arg-Tyr-Phe-Tyr-Asn-Ala-

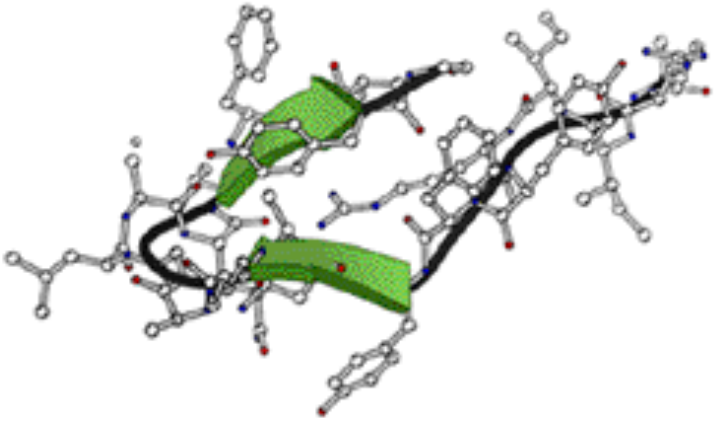
Lys-Ala-Gly-Leu-Cys-Gln-Thr-Phe-Val-Tyr-

Gly

BPTI(16-36) (N = 21): X-Ray

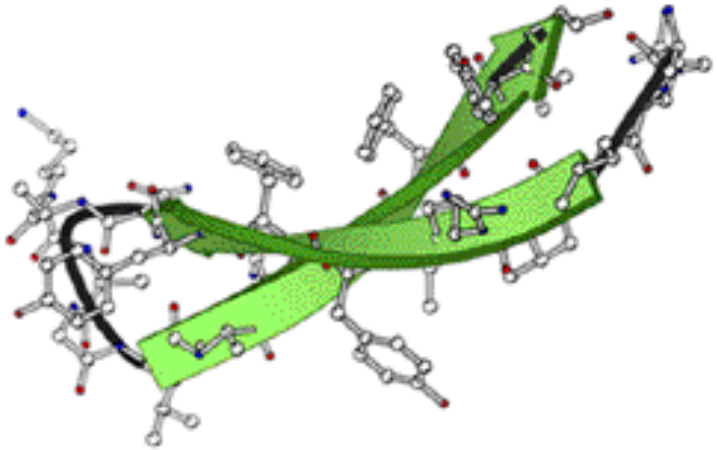
Level 0 (Gas Phase)



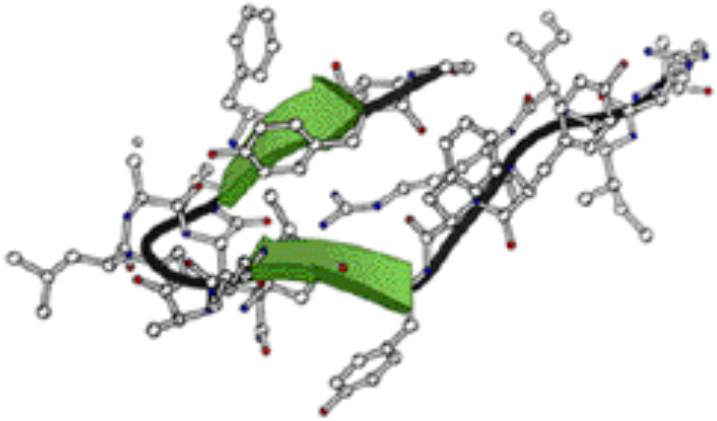


T. Nakazawa, H. Kawai, Y.O. & M. Fukugita, *Protein Eng*. **5**, 495 (1992).

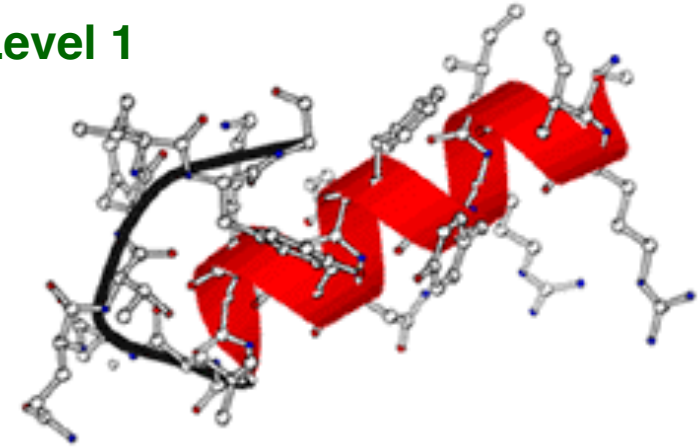
BPTI(16-36) (N = 21): X-Ray



Level 0 (Gas Phase)

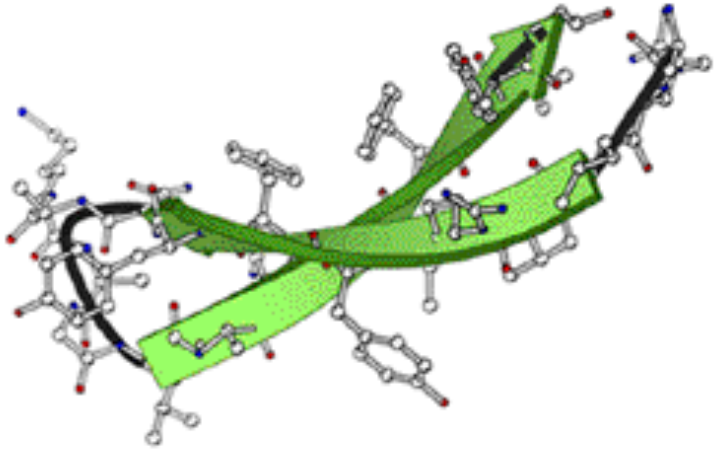


Level 1

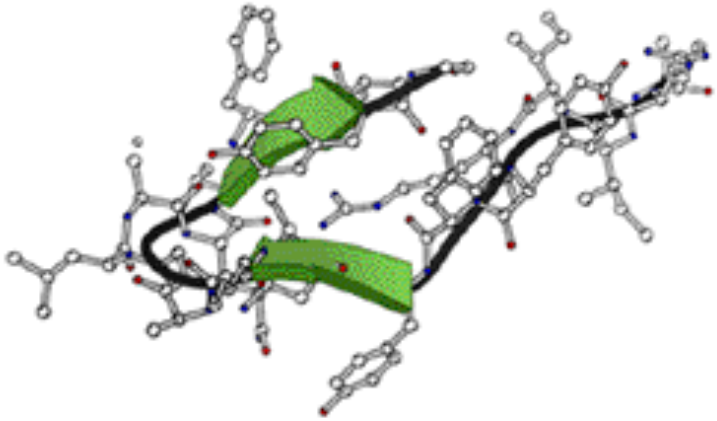


T. Nakazawa & Y.O., *J. Peptide Res.* **54**, 230 (1999).

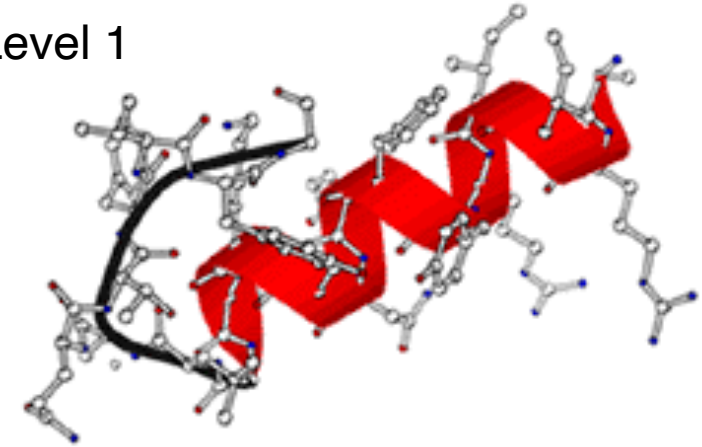
BPTI(16-36) (N = 21): X-Ray



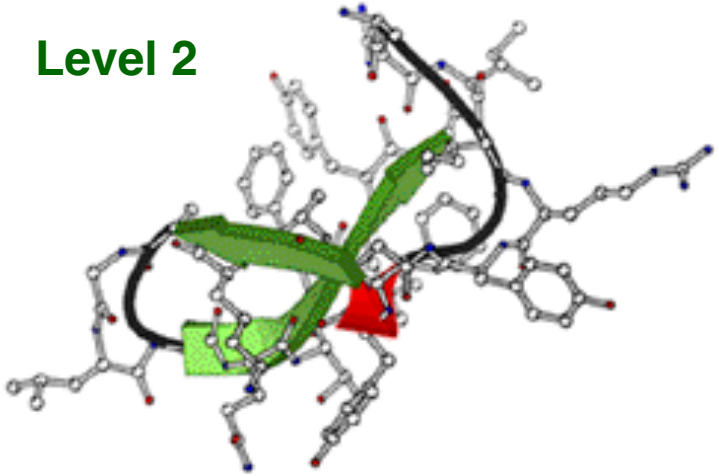
Level 0 (Gas Phase)



Level 1



Level 2



Y.O., M. Masuya, M. Nabeshima & T. Nakazawa, *Chem. Phys. Lett.* **299**, 17 (1999).

C-Peptide of Ribonuclease A (N=13) in Water (Level 3)

Amino-Acid Sequence:

ACE-Ala-Glu-Thr-Ala-Ala-Ala-Lys-Phe-Leu-Arg-Ala-His-Ala-NME

AMBER 99 Force Field (peptide)
TIP3P (water)

13 residue peptide

1387 Water Molecules

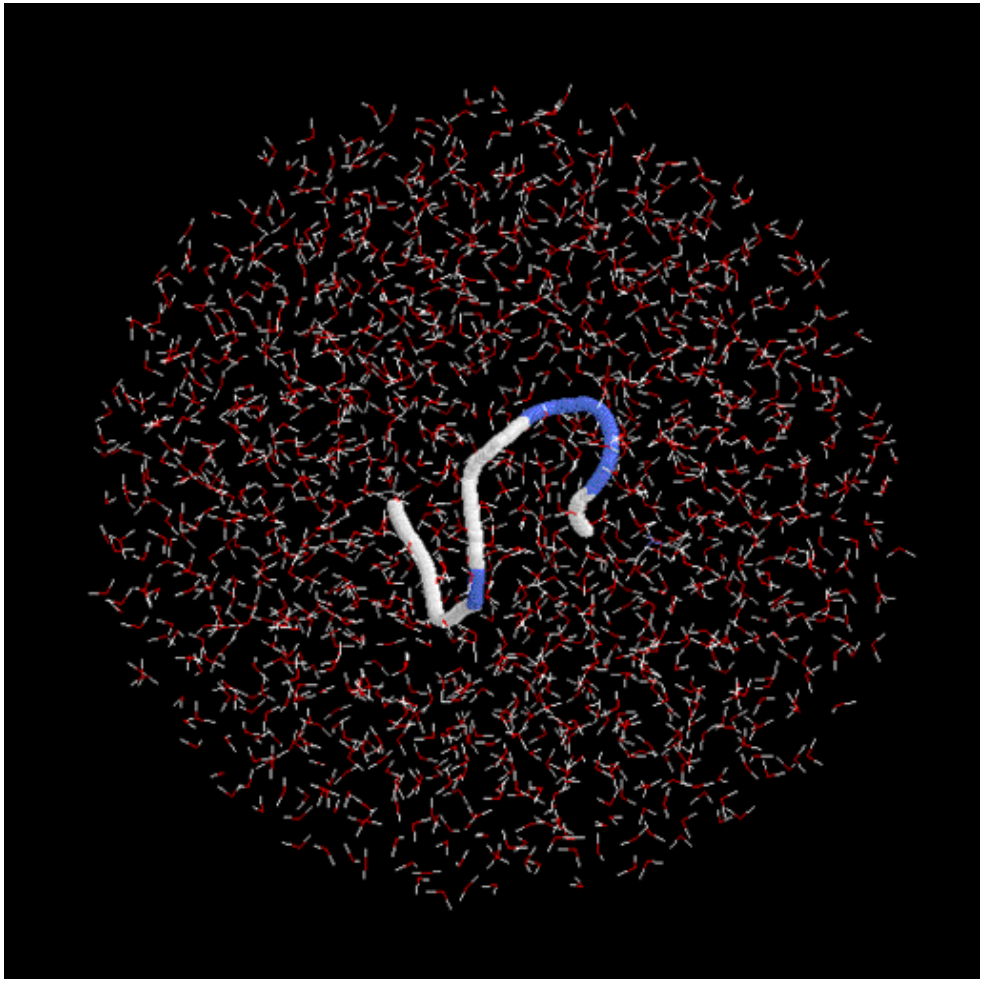
(R = 22 Å)

Total 4365 atoms

NMR Studies (Osterhout 1988)

50-60% helix in solution
First three residues are distorted
Formation of a salt bridge between Glu2 and Arg10

Initial Configuration

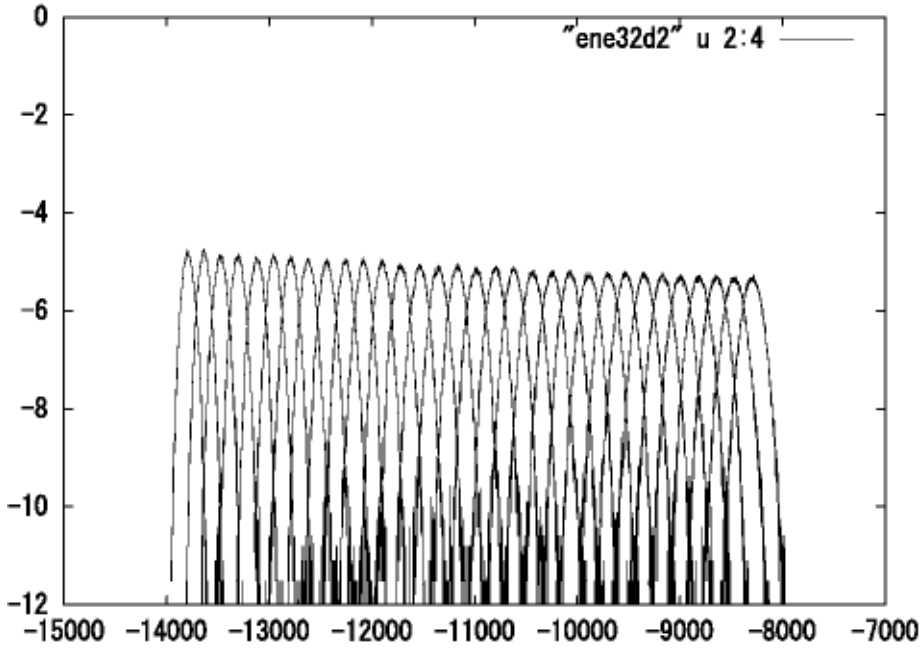


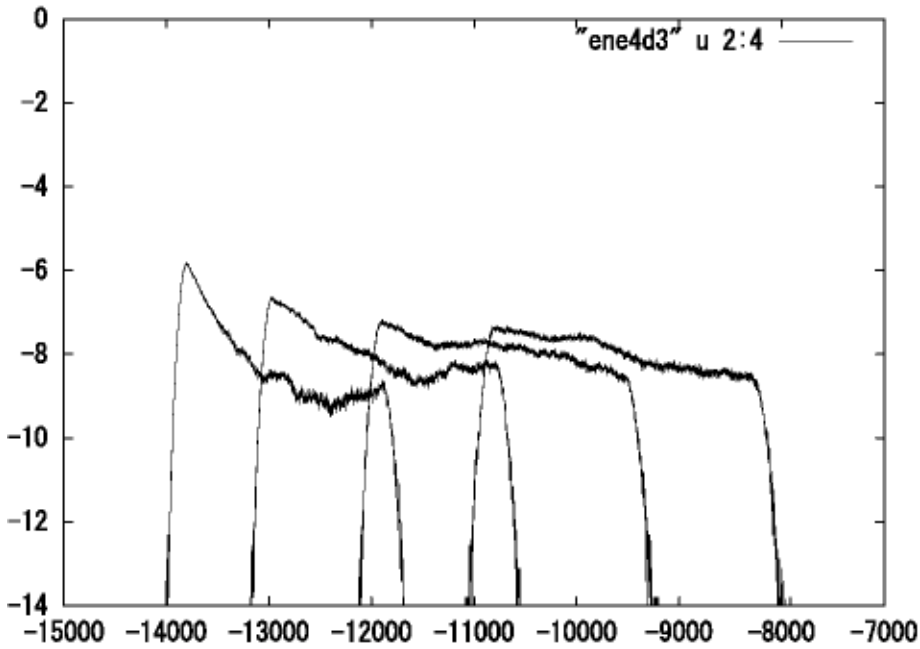
C-Peptide of Ribonuclease A (N=13) in Water

Y. Sugita & Y. O., *Biophys. J.* 88, 3180 (2005).

1. REM (32 replicas)
   100 ps X 32

2. MUCAREM (4 replicas)
   1 ns X 4





# of tunneling events = 0
(-12850, -8250)

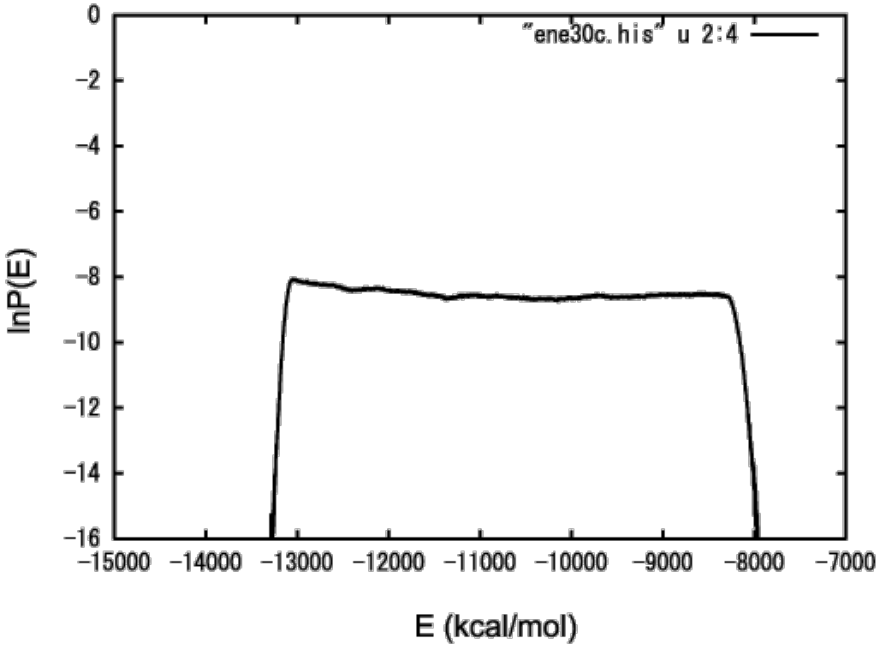
# of tunneling events = 5
(-12850, -8250)

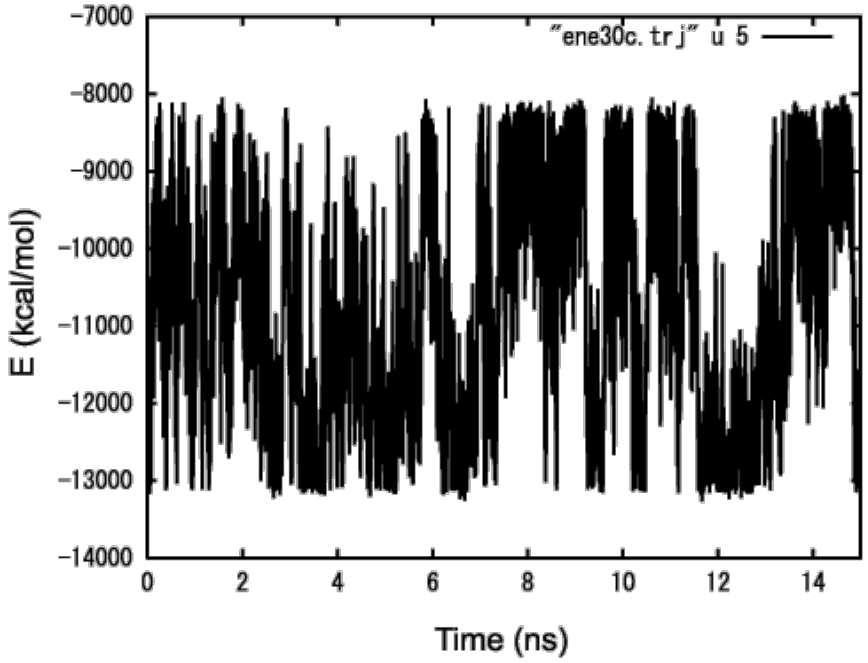
C-Peptide of Ribonuclease A (N=13) in Water

Y. Sugita & Y. O., *Biophys. J.* 88, 3180 (2005).

3. REMUCA (1 replica)
   15 ns

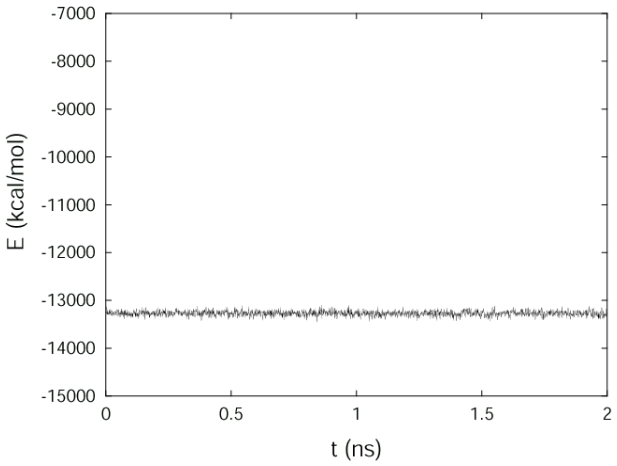
Time series of potential energy

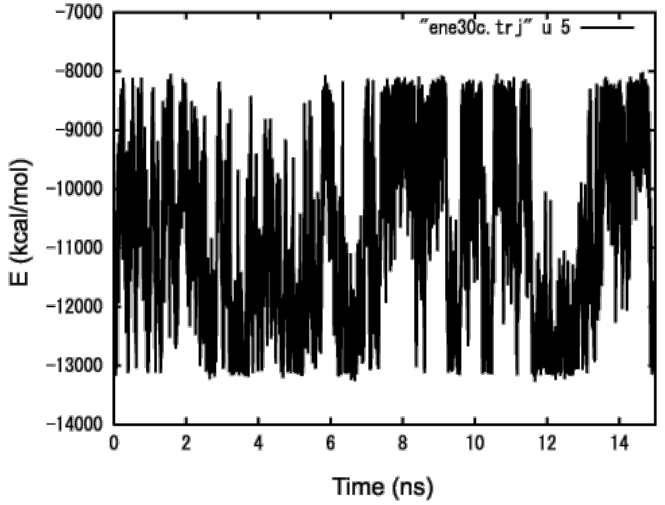




# of tunneling events = 46
(-12850, -8250)

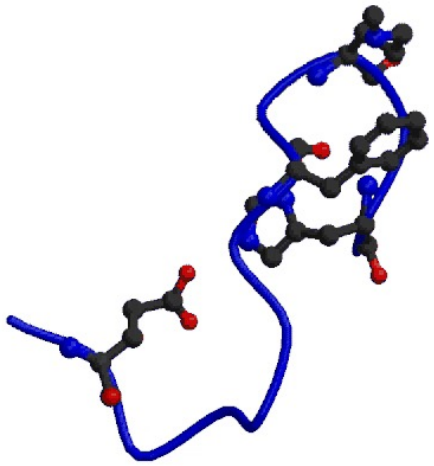
C-Peptide of Ribonuclease A (N=13) in Water

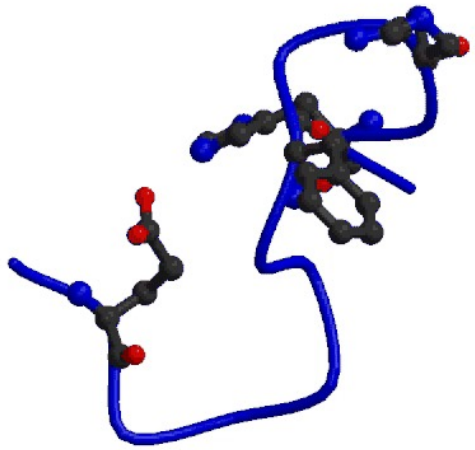




Canonical (T = 275 K)

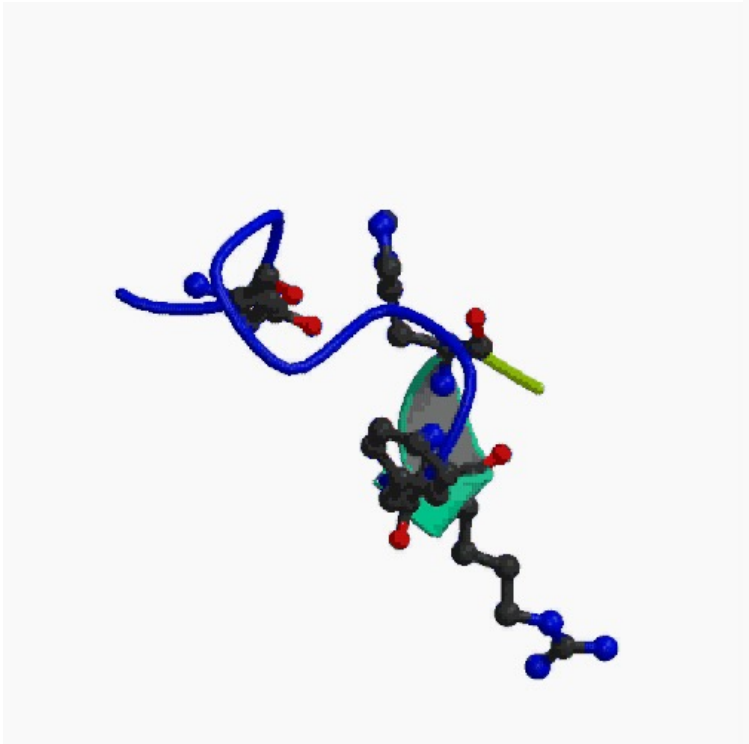
REMUCA

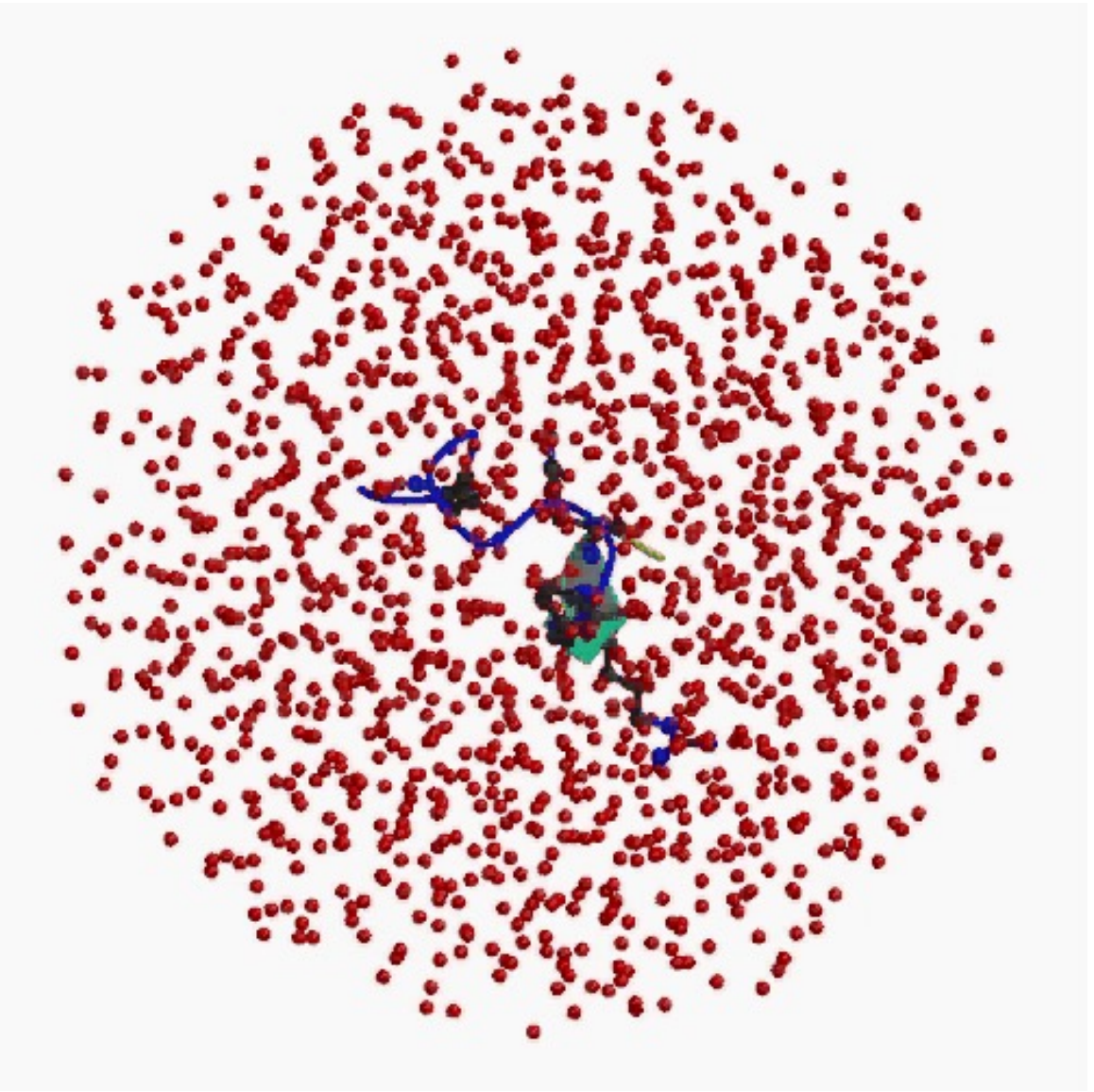




Simulation and movie by Y. Sugita

C-Peptide of Ribonuclease A (N=13) in Water
REMUCA Simulation with AMBER (parm99)





Simulation and movie by Y. Sug

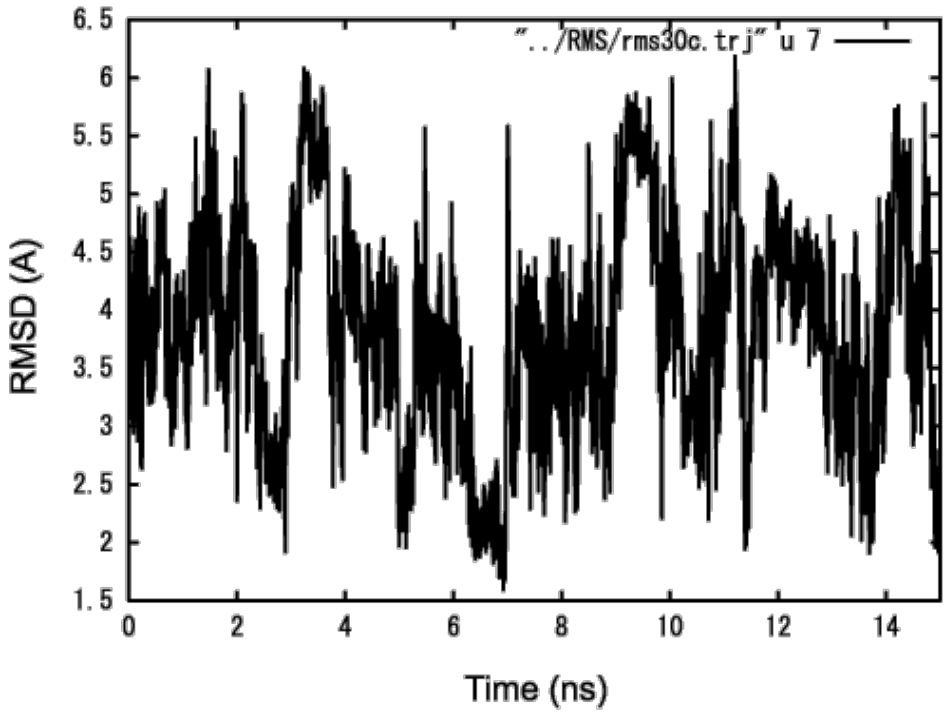
C-Peptide of Ribonuclease A (N=13) in Water

Y. Sugita & Y. O., *Biophys. J.* 88, 3180 (2005).

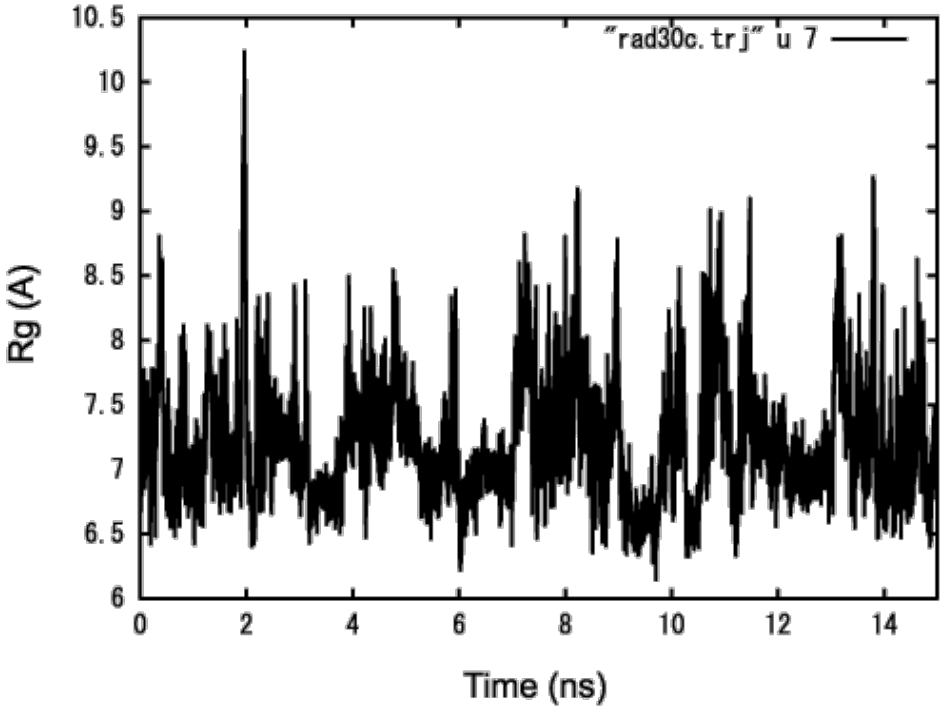
RMSD from “native structure”

Rg: radius of gyration

Time Series of RMSD



Time Series of Rg



ヘルムホルツ自由エネルギー

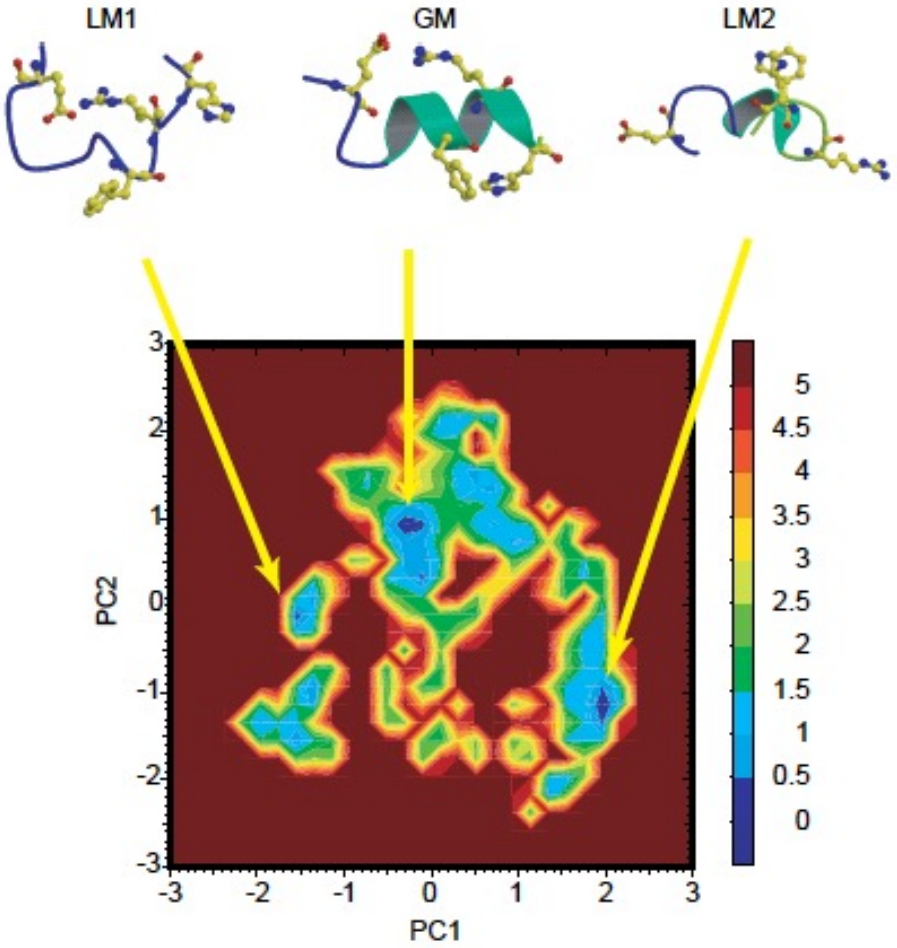
$$F(T) = - \; k_B T \log Z = < E >_T - \; TS(T)$$

$$P_{\mathrm{eq}}(x) = \frac{1}{Z} W_B(x;T) = \frac{1}{Z} \mathrm{e}^{- \; \mathrm{b} E(x)}$$

$$\text{where } Z = \int dx \, W_B(x;T) \text{ and } \mathrm{b} = \frac{1}{k_B T}$$

Free Energy Landscape (C-peptide)

T = 300 K



At T = 300 K, 3 characteristic structural groups exist, among which the group of native-like structure with Glu2-Arg10 salt bridge is the global minimum in free energy.

Y. Sugita & Y. O., *Biophys. J.* 88, 3180 (2005).

G-Peptide (N=16) of Protein G in Water

Amino-Acid Sequence:

Gly-Glu-Trp-Thr-Tyr-Asp-Asp-Ala-Thr-Lys-Thr-Phe-Thr-Val-Thr-Glu

16 residue peptide

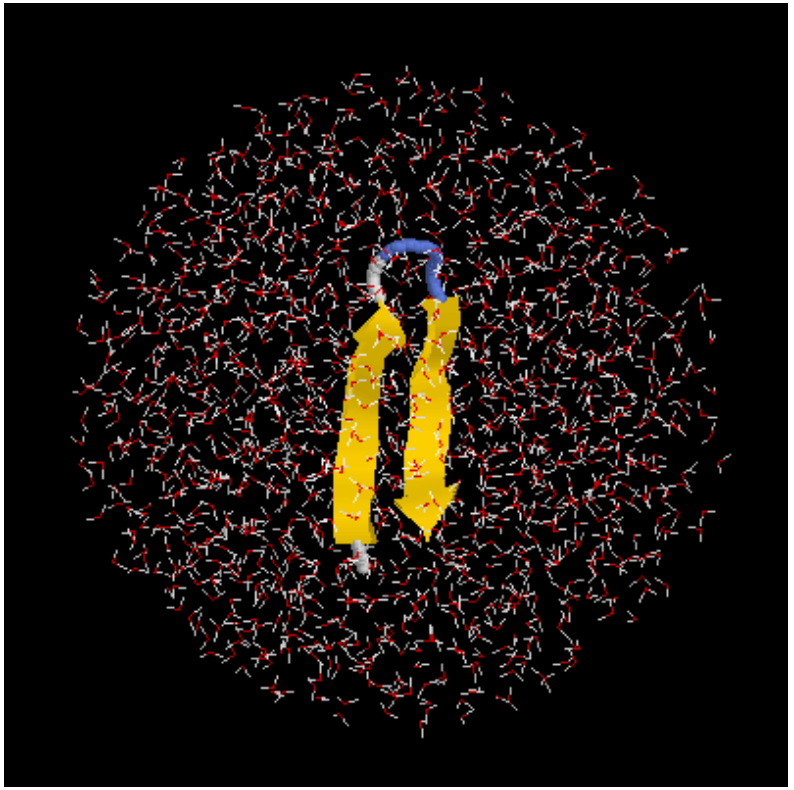
2328 TIP3P water molecules

(R = 26 Å)

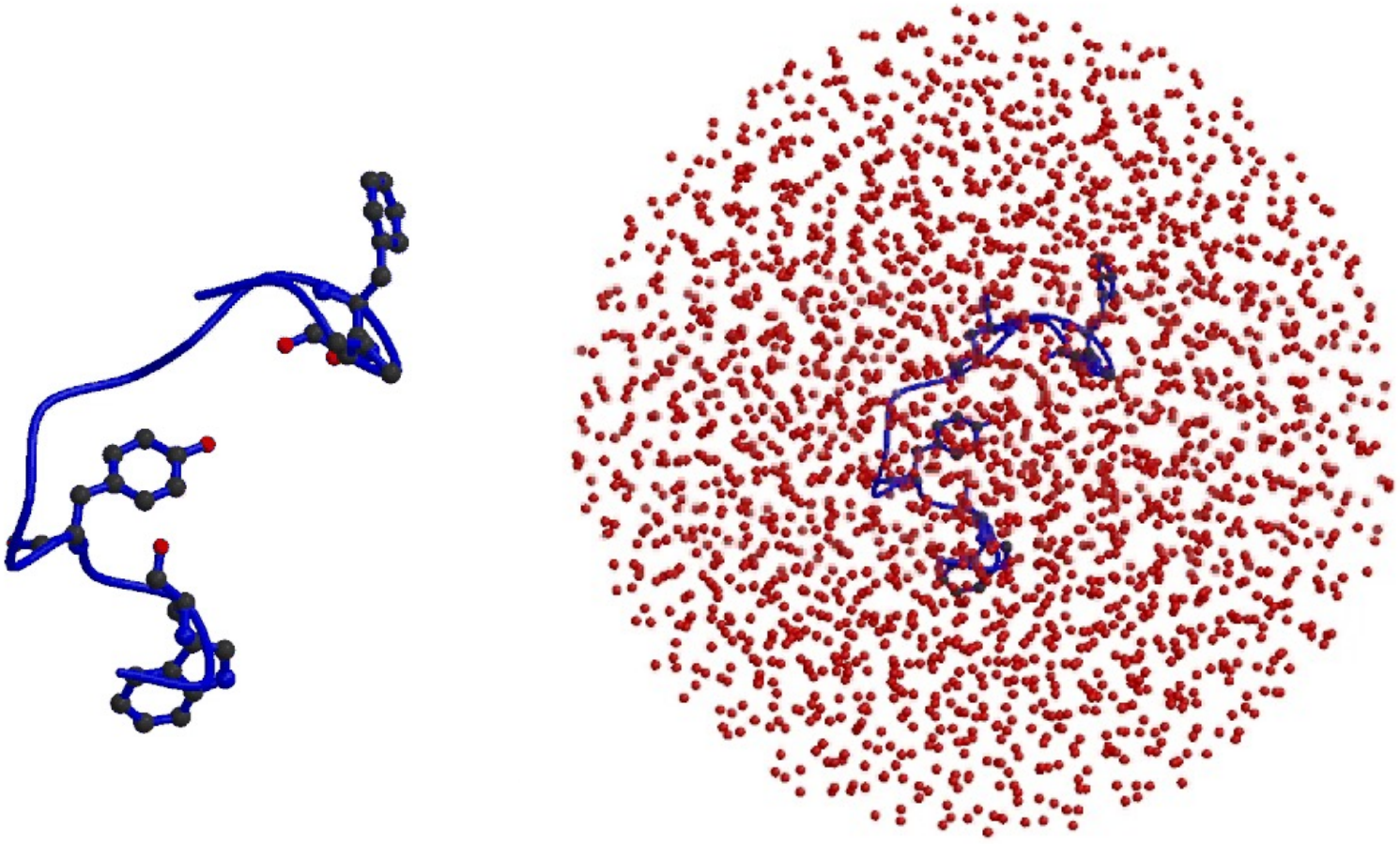
Total: 7231 atoms

- G-peptide is a C-terminal fragment (41-56 residues) of protein G B1 domain.
- This β-hairpin structure is stable in aqueous solution (F. Blanco, G. Rivas & L. Serrano, *Nature Struct. Biol.* **1**, 584 (1994); N. Kobayashi, S. Honda, H. Yoshii, H. Uedaira & E. Munekata, *FEBS Lett.* **366**, 99 (1995)).

"Native Structure" of G-Peptide

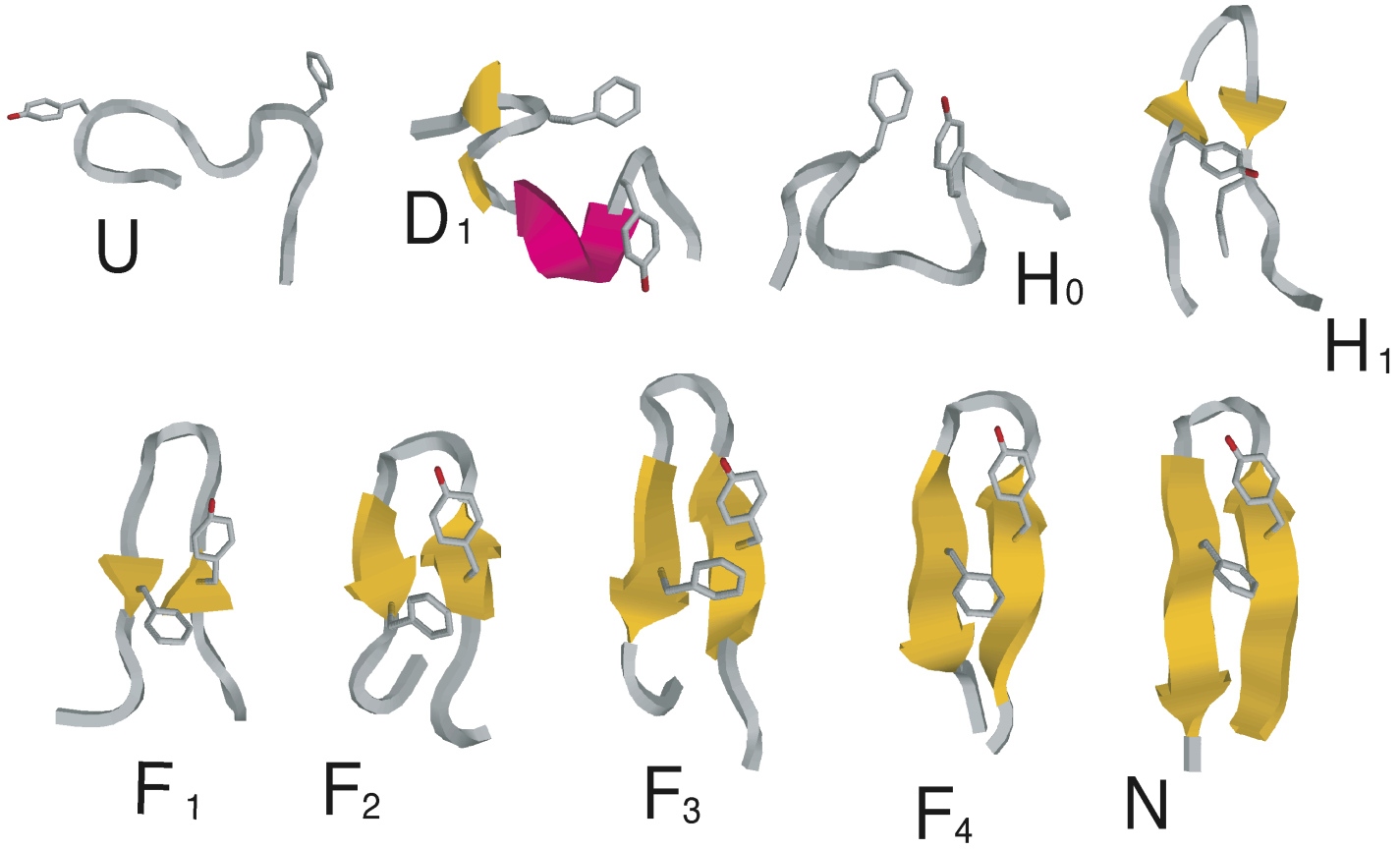


G-Peptide of Protein G (N=16) in Water
MUCAREM Simulation with CHARMM (ver. 22)



Simulation and movie by T. Yoda

Snap-shot Structures of G-peptide observed in the Simulation



T. Yoda, Y. Sugita & Y.O., *Proteins* **66**, 846 (2007).

Comparisons of force field parameters for all-atom protein models with explicit solvent

T. Yoda, Y. Sugita & Y.O., *Chem. Phys. Lett.* **386**, 460-467 (2004);

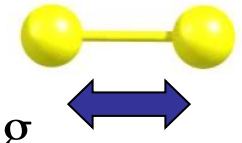
T. Yoda, Y. Sugita & Y.O., *Chem. Phys.* **307**, 269-283 (2004).

Force-field parameters for all-atom models

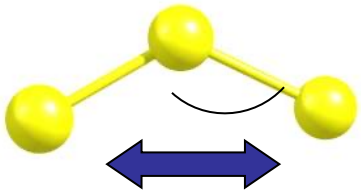
Commonly used conformational potential energy

$$E_{conf} = \sum_{bonds} K_r (r - r_{eq})^2$$
$$+ \sum_{angles} K_\theta (\theta - \theta_{eq})^2$$
$$+ \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]$$
$$+ \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + 332 \frac{q_i q_j}{\varepsilon R_{ij}} \right]$$

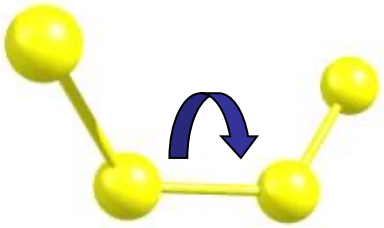
Bond-stretching



Bond-bending



Dihedral angle

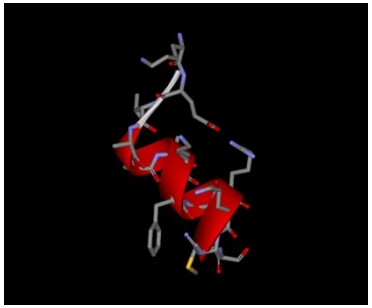


Non-bonding interactions
(Lennard-Jones and electrostatic)

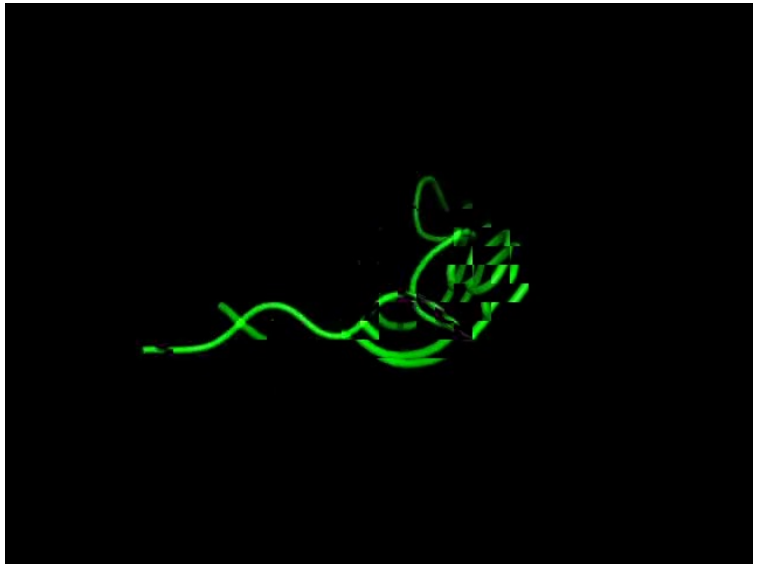
These energy terms include some force-field parameters (blue color)

The existing force fields have different force-field parameters

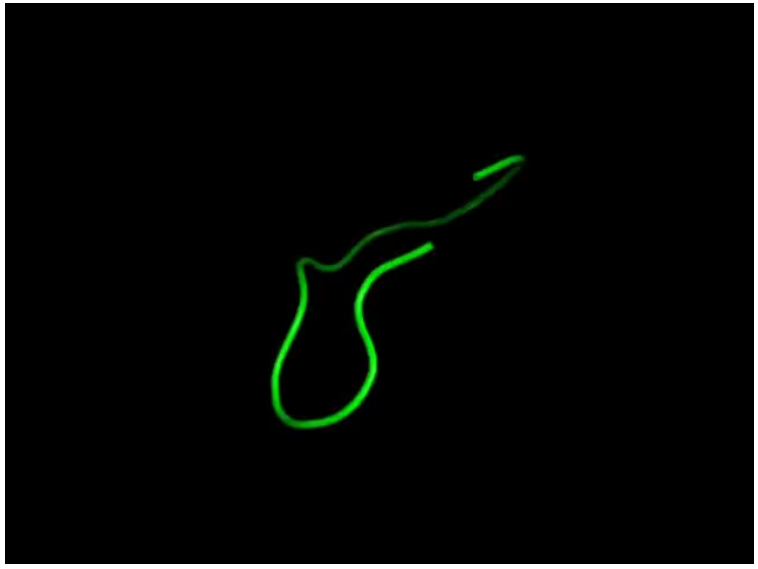
Typical example of folding simulations using different force fields



C-peptide
(13 residues)
Lys-Glu-Thr-Ala-Ala-Ala-Lys-Phe-Glu-Arg-Gln-His-Met



AMBER ff94

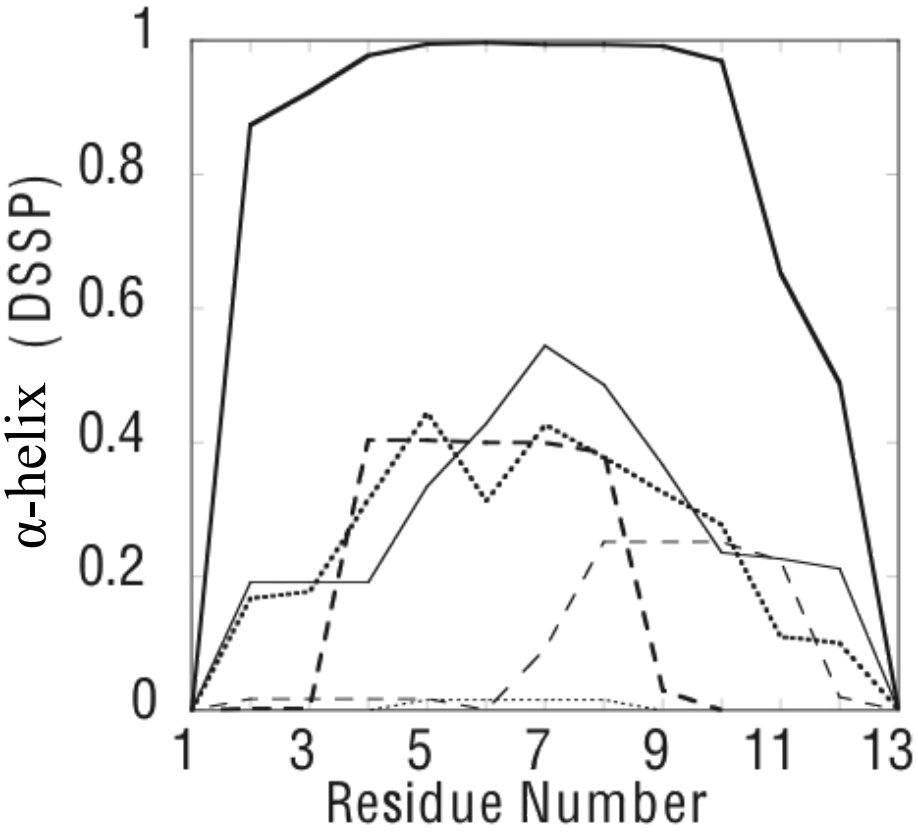


AMBER ff96

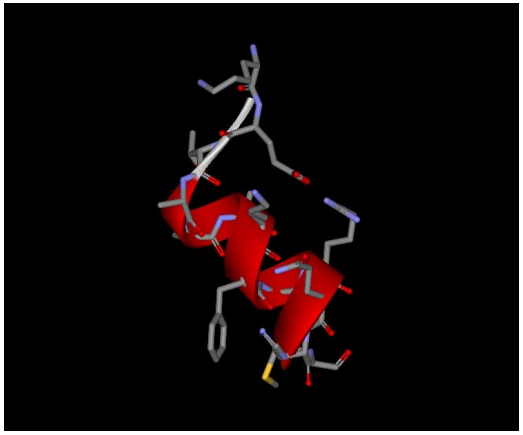
Simulation and movie by Y. Sakae

Method : Simulated annealing, Simulation time : 1.0 nsec, Temperature : 700～200 K, Solvent model : GB/SA

Force-field dependency of secondary-structure properties



Helicity of C-peptide



C-peptide
(13 residues)

AMBER parm94
AMBER parm96
AMBER parm99
CHARMM22
OPLS-AA/L
GROMOS96

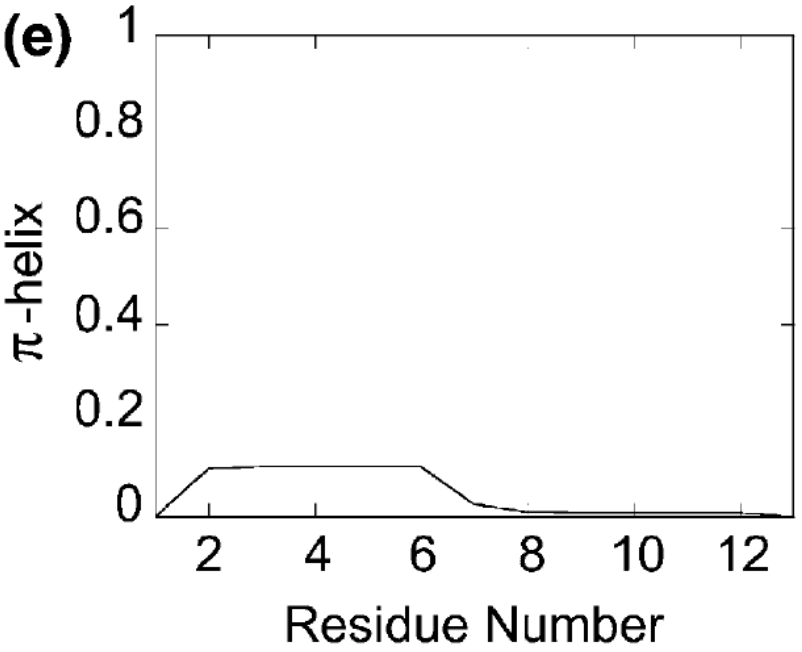
T. Yoda, Y. Sugita & Y.O.,

Chem. Phys. Lett. **386**, 460 (2004).

7. 分子科学研究所 (1995-2005)

2001年8月、米国カリフォルニア州San Diegoの隣のLa JollaのScripps研究所Charlie Brooksを杉田有治さんと訪問しました。

I gave a seminar talk.
In one of the slides I had the unpublished figure on the right. Namely, π-helix formation was found in C-peptide by MUCAREM simulations with CHARMM22.



T. Yoda, Y. Sugita & Y.O., *Chem. Phys. Lett.* **386**, 460 (2004);
T. Yoda, Y. Sugita & Y.O., *Chem. Phys.* **307**, 269 (2004).

I think this influenced the development of CMAP of CHARMM force field:
M. Feig, A.D. MacKerell Jr. & C.L. Brooks III, *J. Phys. Chem. B* **107**, 2831 (2003).
A.D. MacKerell Jr., M. Feig & C.L. Brooks III, *J. Am. Chem. Soc.* **126**, 698 (2004)

Force field refinement for all-atom protein models

Y. Sakae & Y.O., *Chem. Phys. Lett*. **382**, 626-636 (2003);

Y. Sakae & Y.O., *J. Phys. Soc. Jpn.* **75**, 054802 (9 pages) (2006);

Y. Sakae & Y.O., *Mol. Sim*. **36**, 138-158 (2010);

Y. Sakae & Y.O., *J. Chem. Phys*. **138**, 064103 (2013), etc.

Review:

Y. Sakae & Y.O., in *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes – from Bioinformatics to Molecular Quantum Mechanics*, A. Liwo (ed.) (Springer-Verlag, Berlin Heidelberg, 2019) pp. 203-256.

Conformational Energy

$$E_{conf} = E_{BL} + E_{BA} + E_{torsion} + E_{nonbond}$$

Backbone-torsion energy term

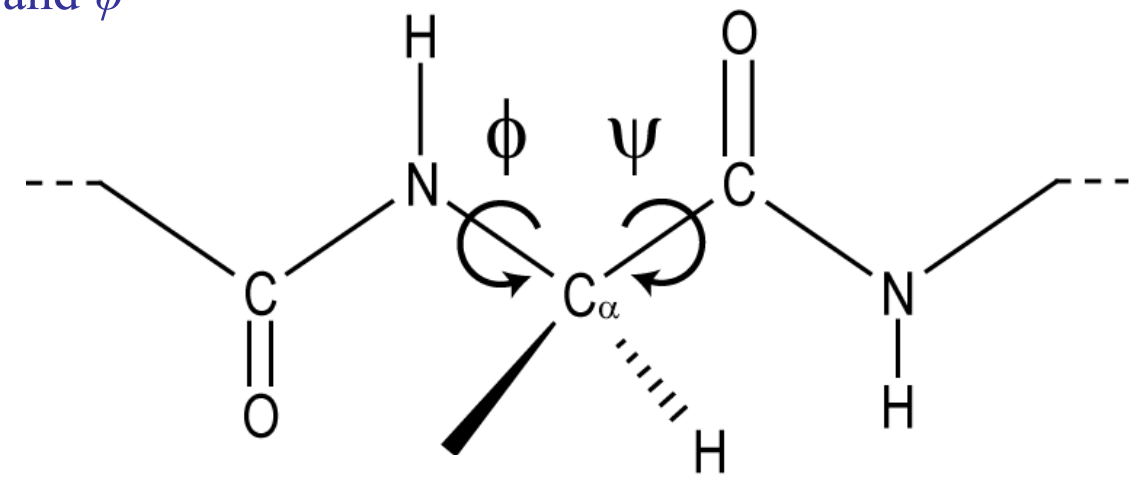
$$E_{torsion} = \sum_{\Phi}\sum_{n} \frac{V_n}{2}[1+\cos(n\Phi - \gamma_n)]$$

$$= E(\varphi,\psi) + E_{rest}$$

backbone dihedral angles φ and ψ

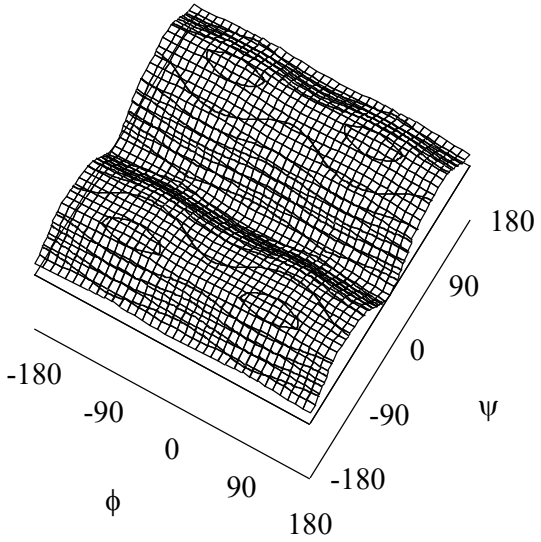
rest of the torsion terms

Φ : all dihedral angles
n : number of waves
γ_n : phase
V_n : Fourier coefficient

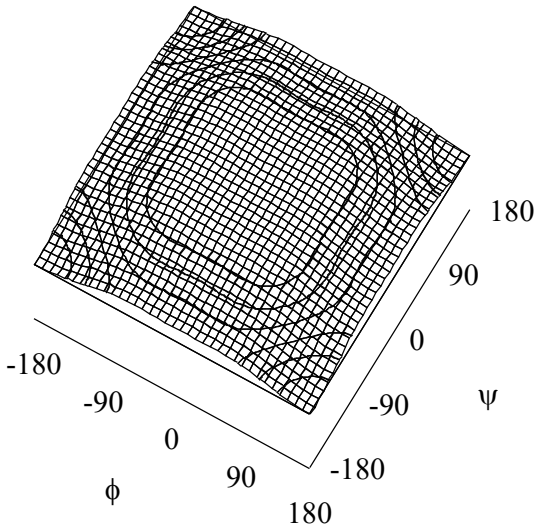


Side chain

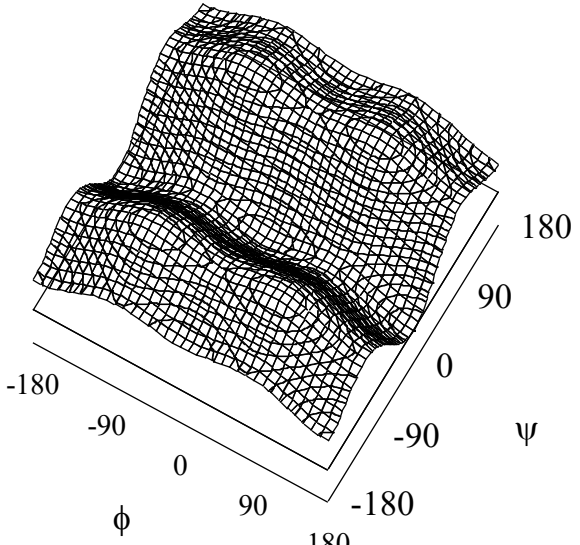
Backbone-torsion energy surfaces of some force fields



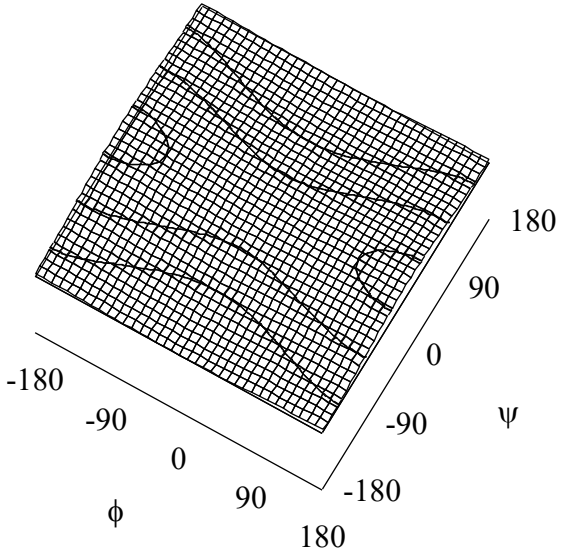
AMBER parm94



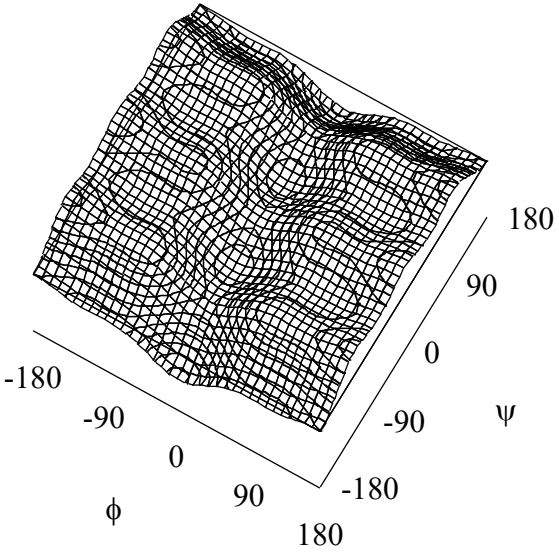
AMBER parm96



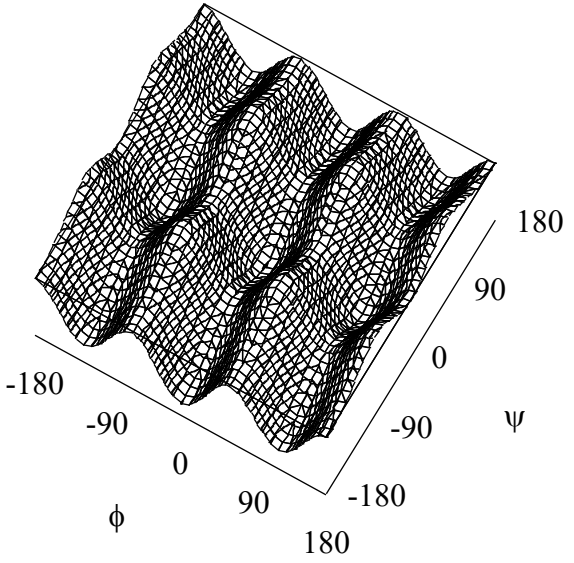
AMBER parm99



CHARMM22



OPLS-AA



OPLS-AA/L

1. Proposal of new functional form

Y. Sakae and Y.O., *J. Phys. Soc. Jpn.* **75**, 054802 (9 pages) (2006).

conventional energy term

$$E(\varphi,\psi)=\sum_m \frac{V_m}{2}[1+\cos(m\varphi-\gamma_m)]+\sum_n \frac{V_n}{2}[1+\cos(n\psi-\gamma_n)]$$

New backbone-torsion-energy term

$$\varepsilon(\varphi,\psi)=a+\sum_{m=1}^{\infty}(b_m\cos m\varphi+c_m\sin m\varphi)+\sum_{n=1}^{\infty}(d_n\cos n\psi+e_n\sin n\psi)$$

$$+\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}(f_{mn}\cos m\varphi\cos n\psi+g_{mn}\cos m\varphi\sin n\psi$$

$$+h_{mn}\sin n\varphi\cos n\psi+i_{mn}\sin m\varphi\sin n\psi)$$

$a, b_m, c_m, d_n, e_n, f_{mn}, g_{mn}, h_{mn}, i_{mn}$: Fourier coefficient

New torsion energy term

$$E_{torsion}=\varepsilon(\varphi,\psi)+E_{rest}$$

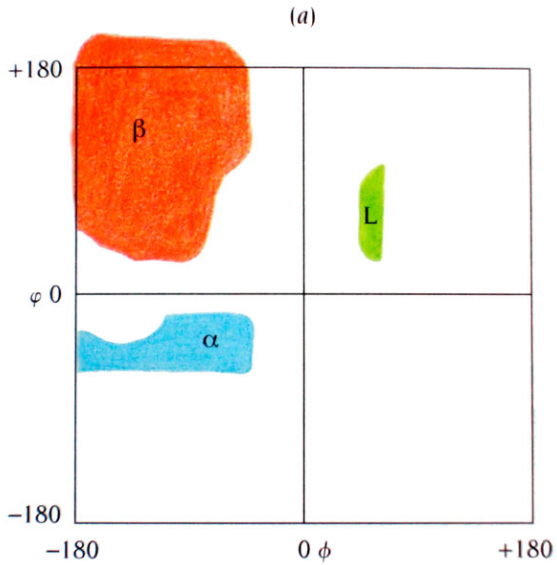
We consider the truncated Fourier series:

$$\begin{aligned}\varepsilon(\varphi,\psi) = a &+ b_1\cos\varphi + c_1\sin\varphi + b_2\cos 2\varphi + c_2\sin 2\varphi \\ &+ d_1\cos\psi + e_1\sin\psi + d_2\cos 2\psi + e_2\sin 2\psi \\ &+ f_{11}\cos\varphi\cos\psi + g_{11}\cos\varphi\sin\psi \\ &+ h_{11}\sin\varphi\cos\psi + i_{11}\sin\varphi\sin\psi\end{aligned}$$

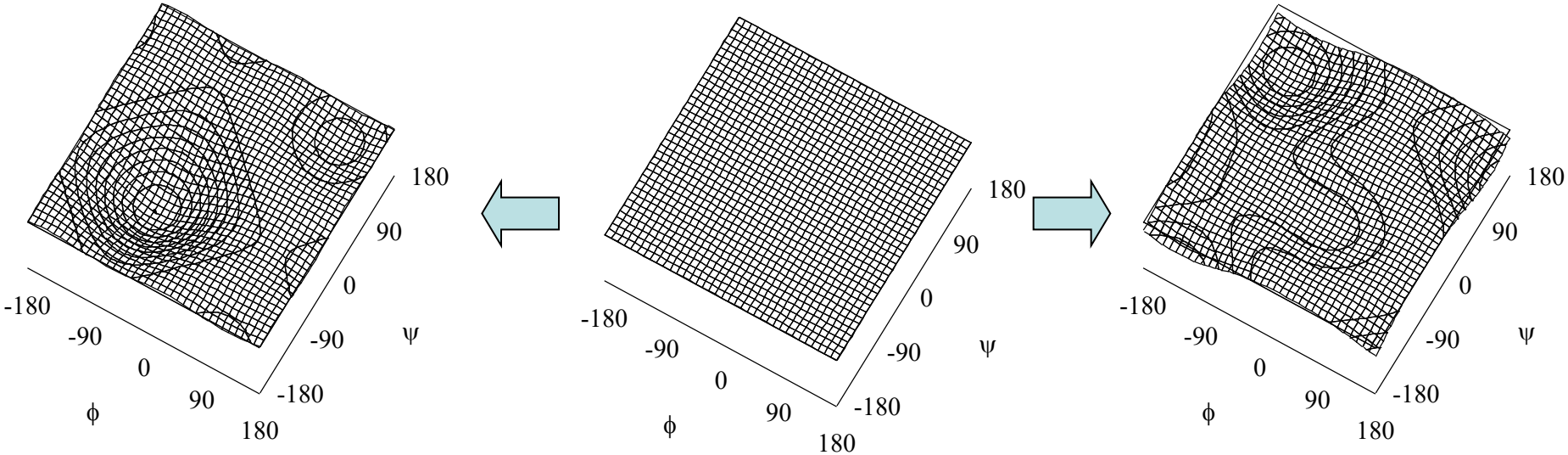
This function has 13 Fourier-coefficient parameters.

Example of the application of new backbone-torsion energy term

Ramachandran plot



Carl Branden and John Tooze著
（勝部幸輝ほか訳）
「タンパク質の構造入門第2版」
ニュートンプレス（2000）



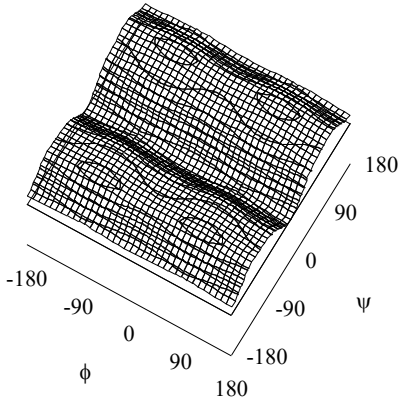
α-helix region

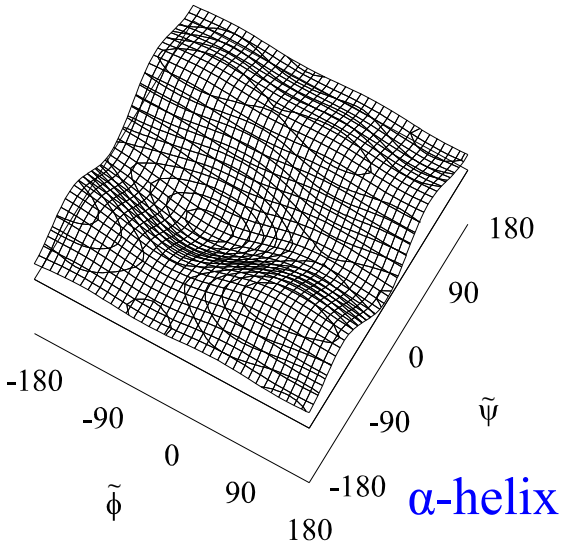
β-structure region

Energy surface of new energy term can represent Ramachandran space directly

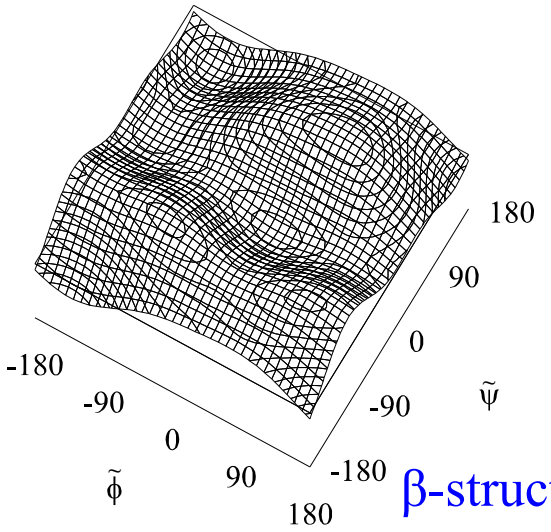
Application to AMBER parm94 and AMBER parm96

AMBER parm94



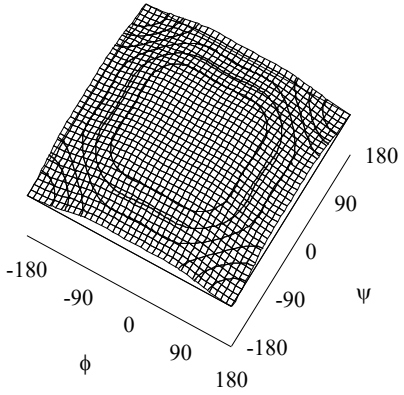


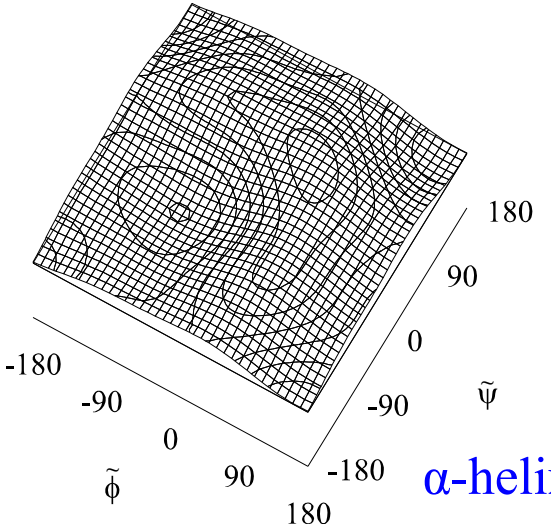
α-helix



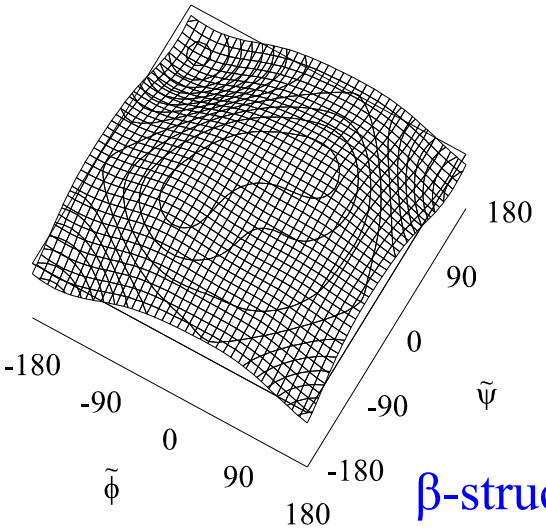
β-structure

AMBER parm96





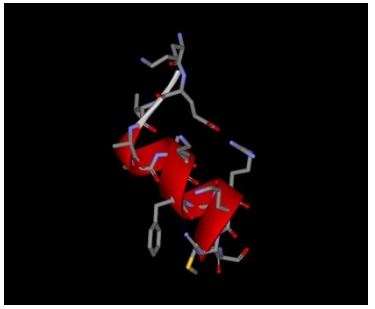
α-helix



β-structure

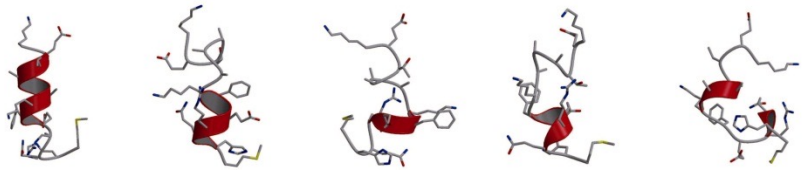
Results of folding simulations

C-peptide

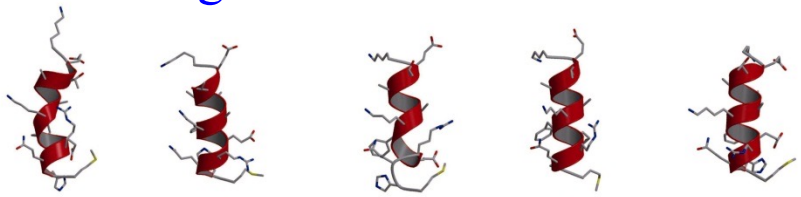


AMBER parm94

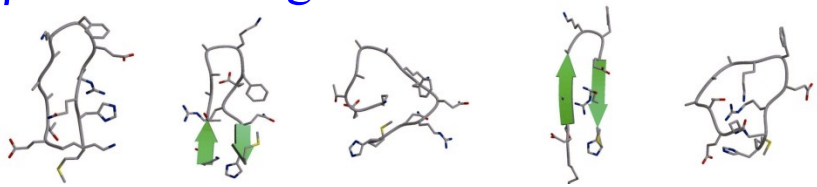
Original



α-helix region

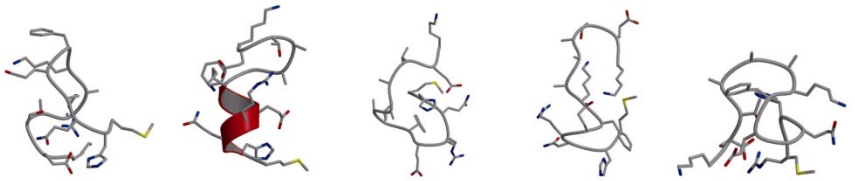


β-structure region

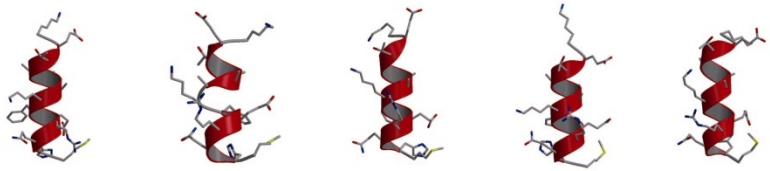


AMBER parm96

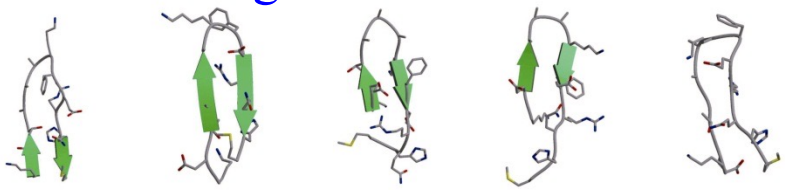
Original



α-helix region



β-structure region

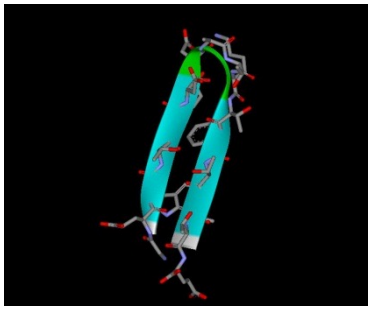


Simulated annealing simulation

Simulation time : 1ns (1,000,000 MD steps × 1.0fs × 60 times)
Temperature : 2,000K to 250K (Berendsen's method)
Solvent : GB/SA model

Results of folding simulations

G-peptide

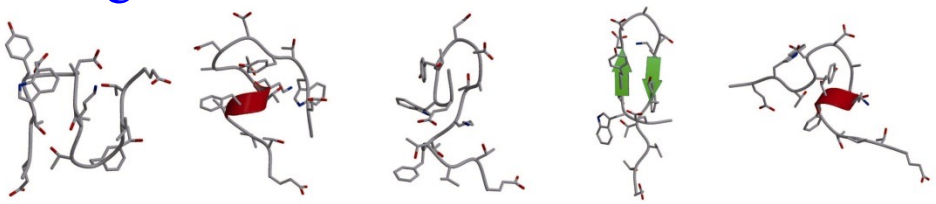


AMBER parm94

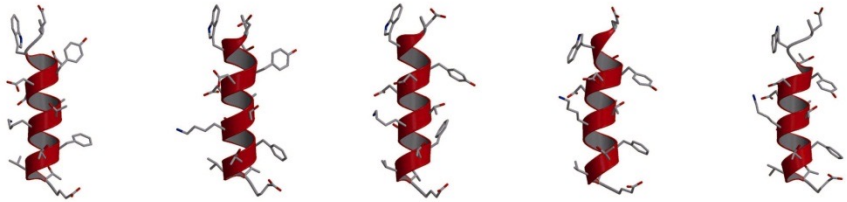
AMBER parm96

Original

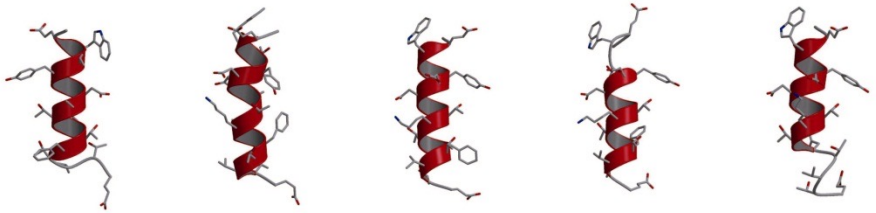
Original



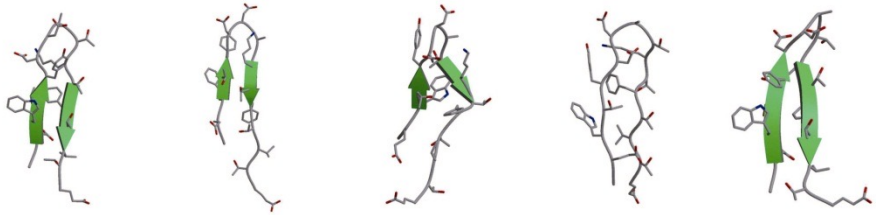
α-helix region



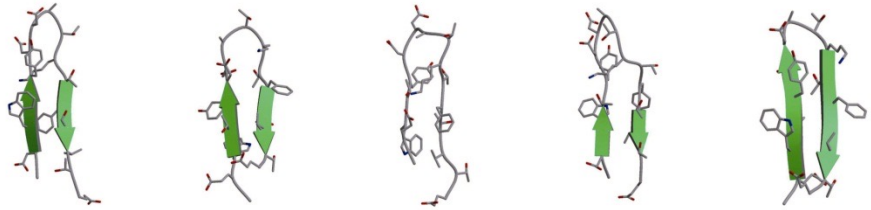
α-helix region



β-structure region



β-structure region



Simulated annealing simulation

Simulation time : 1ns (1,000,000 MD steps × 1.0fs × 60 times)
Temperature : 2,000K to 250K (Berendsen's method)
Solvent : GB/SA model

Optimization method of force-field parameters

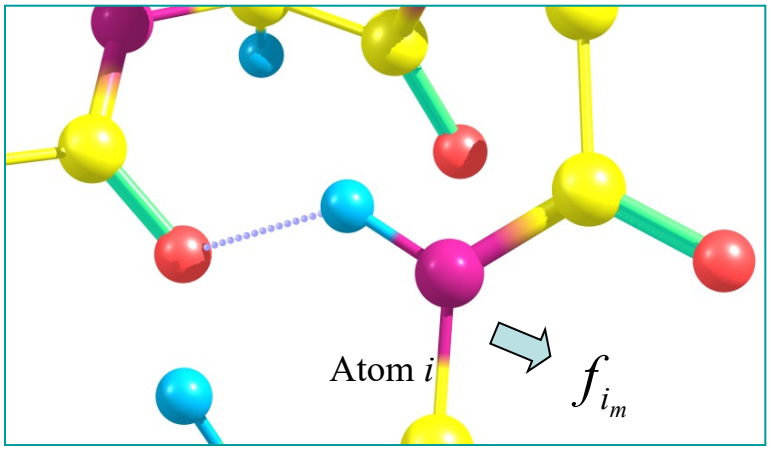
$$F = \sum_{m=1}^{N} \frac{1}{N_m} \sum_{i_m=1}^{N_m} \left| \overrightarrow{f_{i_m}} \right|^2$$

$$\vec{f}_{i_m} = -\frac{\partial E_{tot}^{\{m\}}}{\partial \vec{x}_{i_m}}$$

N_m Number of atoms in molecule m

f_{i_m} Force acting on atom i

$E_{tot}^{\{m\}}$ Total potential energy for molecule m



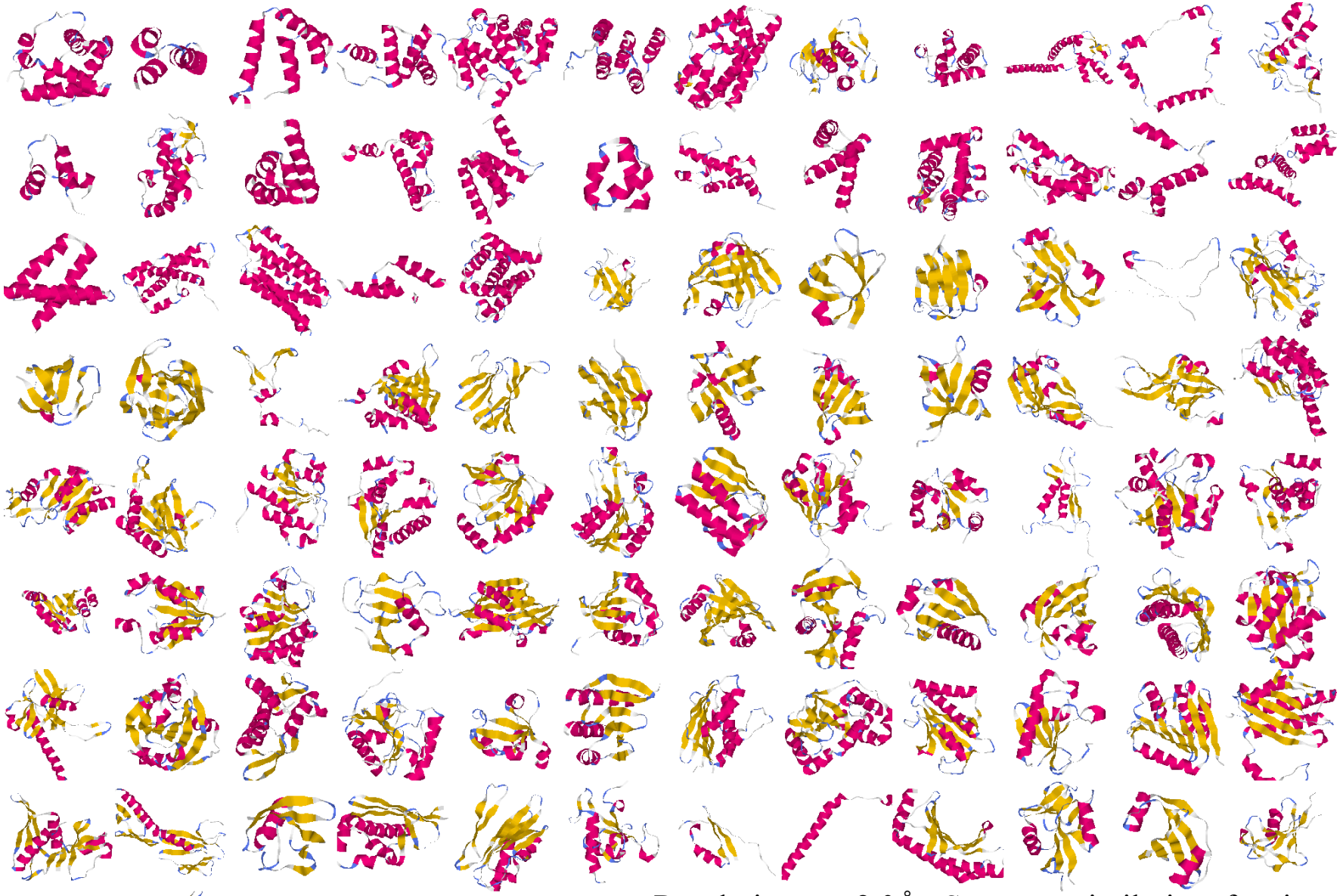
Molecule m

If force-field parameters are of ideal values, all native structures are stable without any force acting on each atom in molecules. (we expect F=0)

In reality, $F \neq 0$, and because $F \geq 0$, we can optimize the force-field parameters by minimizing F with respect to these parameters.

Y. Sakae and Y.O., *Chem. Phys. Lett.* **382**, 626-636 (2003).

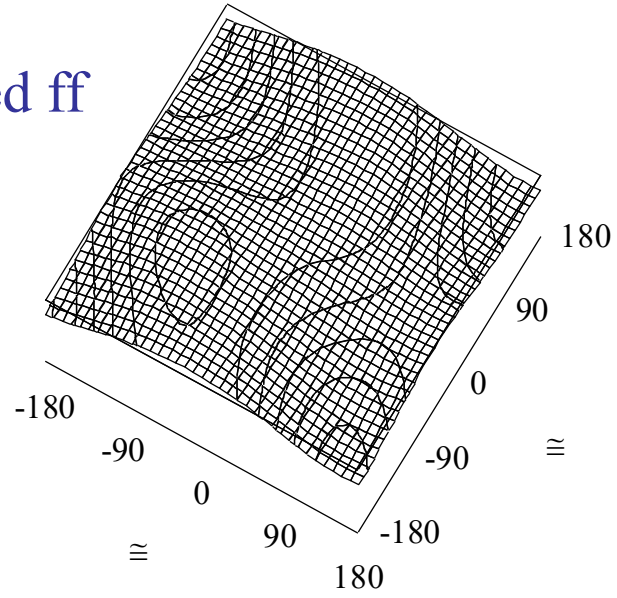
Structures of selected 100 proteins



Resolution: >= 2.0Å, Sequence similarity of amino acid: >= 30%, Number of residue: < 200

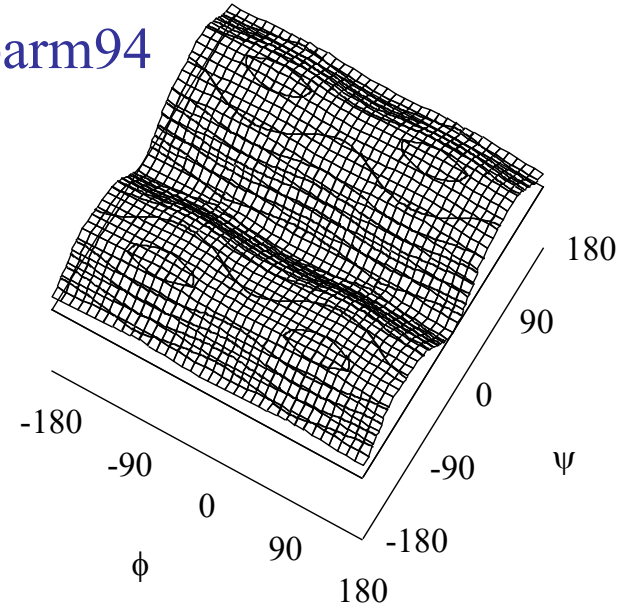
Results: Optimized force-field parameters and its energy surface

Optimized ff

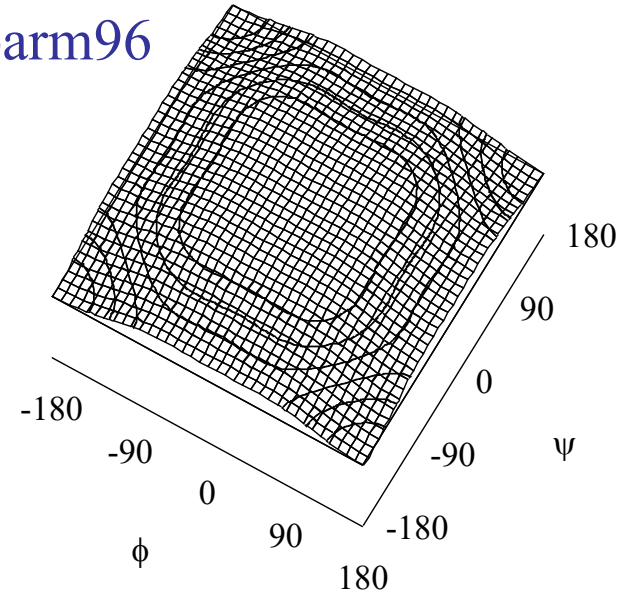


$$\varepsilon(\varphi,\psi) = 0.835\cos\varphi - 0.327\sin\varphi$$
$$- 0.088\cos 2\varphi + 0.100\sin 2\varphi$$
$$+ 0.287\cos\psi - 0.160\sin\psi$$
$$+ 0.019\cos 2\psi - 0.054\sin 2\psi$$
$$- 0.427\cos\varphi\cos\psi + 0.247\cos\varphi\sin\psi$$
$$+ 0.114\sin\varphi\cos\psi + 0.603\sin\varphi\sin\psi$$

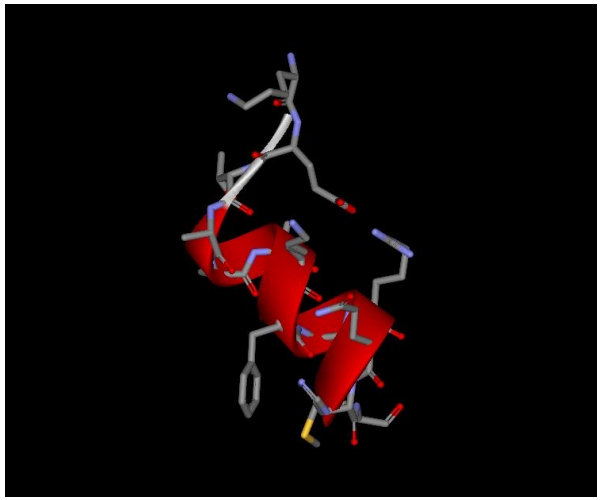
AMBER parm94



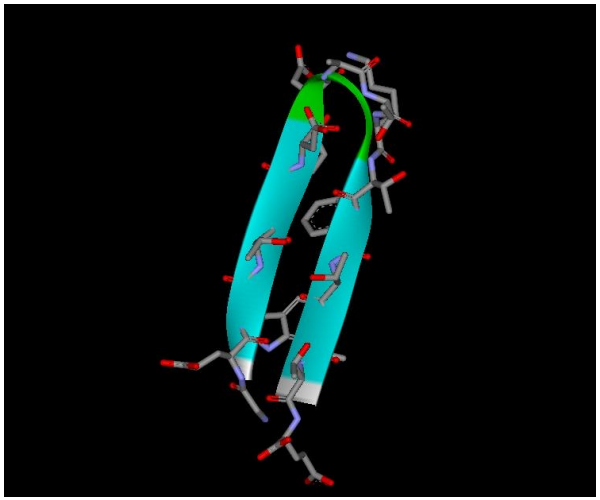
AMBER parm96



Test simulations



C-peptide
(13 residues)



G-peptide
(16 residues)

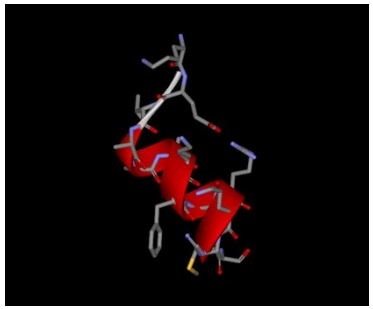
blocked by $COCH_3$- and $–NH_2$

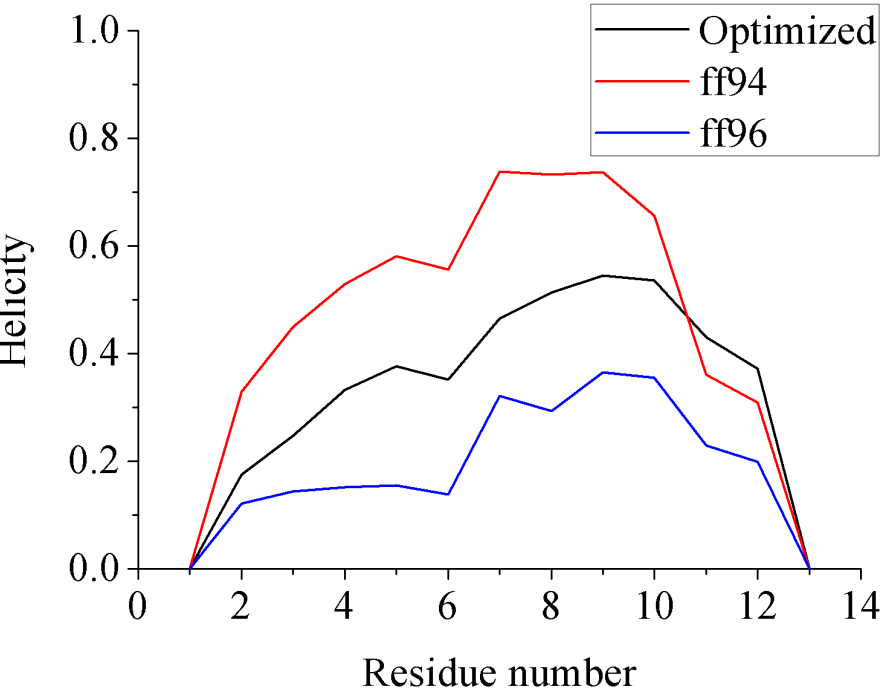
Replica-Exchange Molecular Dynamics (REMD) simulations

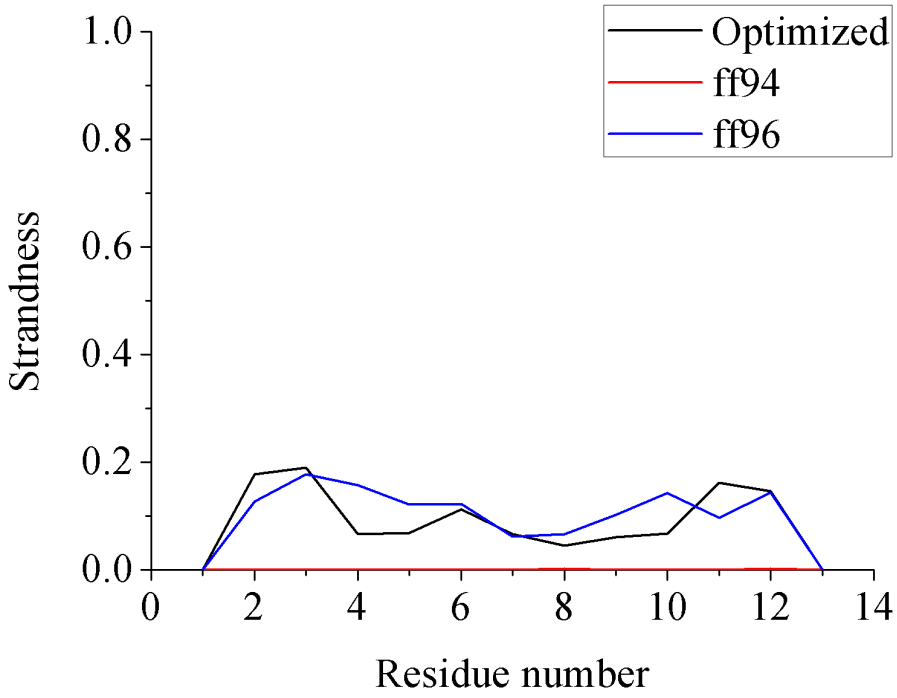
Simulation time : 5.0 ns (32 replica)
Temperature : 700K to 250K (Nosé-Hoover method)
Solvent : GB/SA model
Program : Modified TINKER program package

Comparison of helicity and strandness

C-peptide



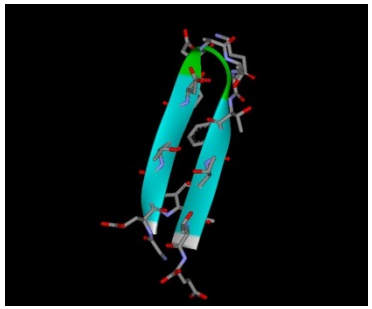


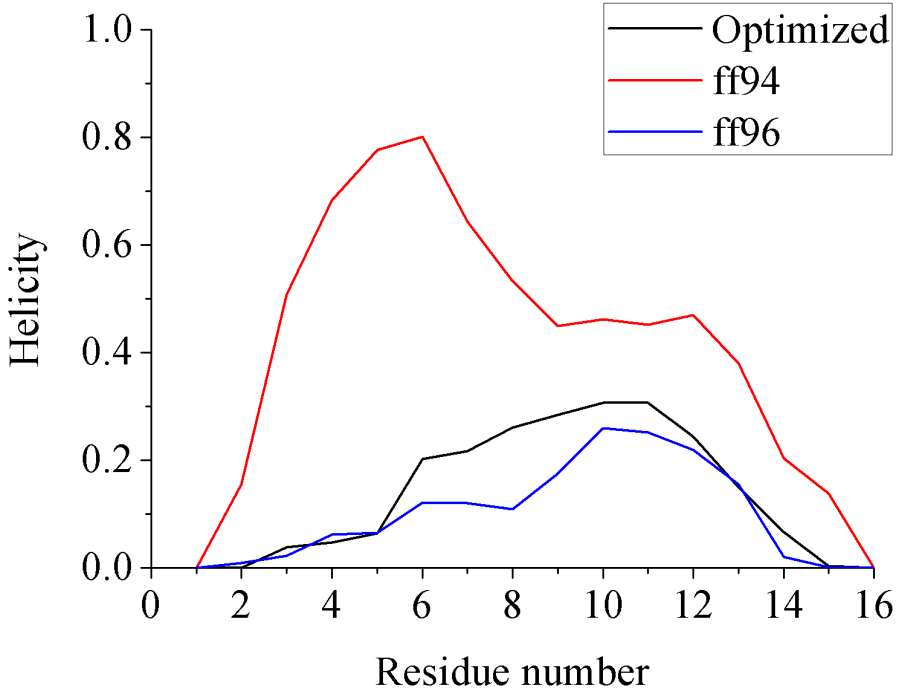


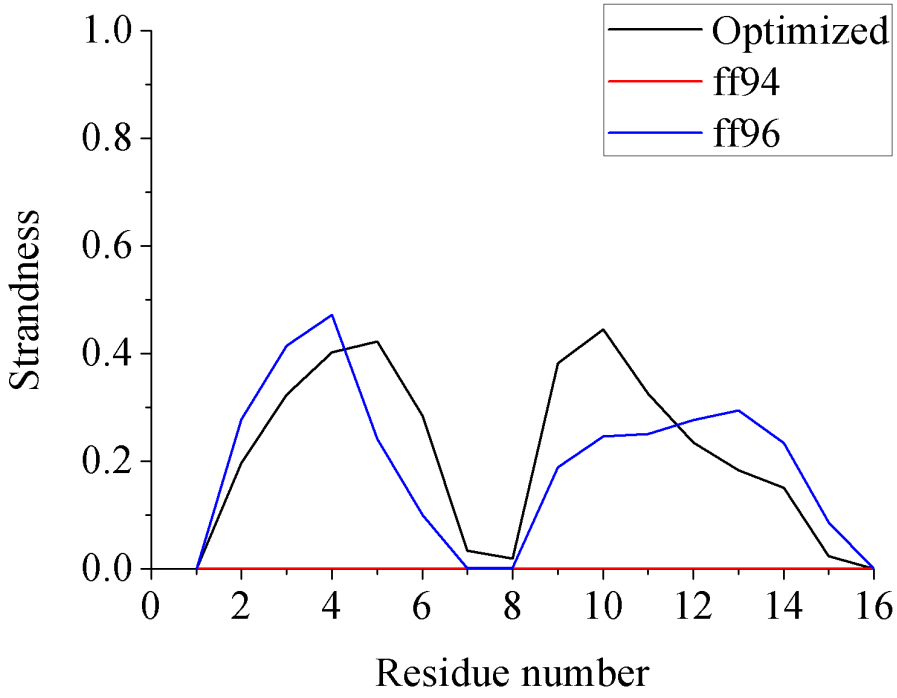
300 K

Comparison of helicity and strandness

G-peptide







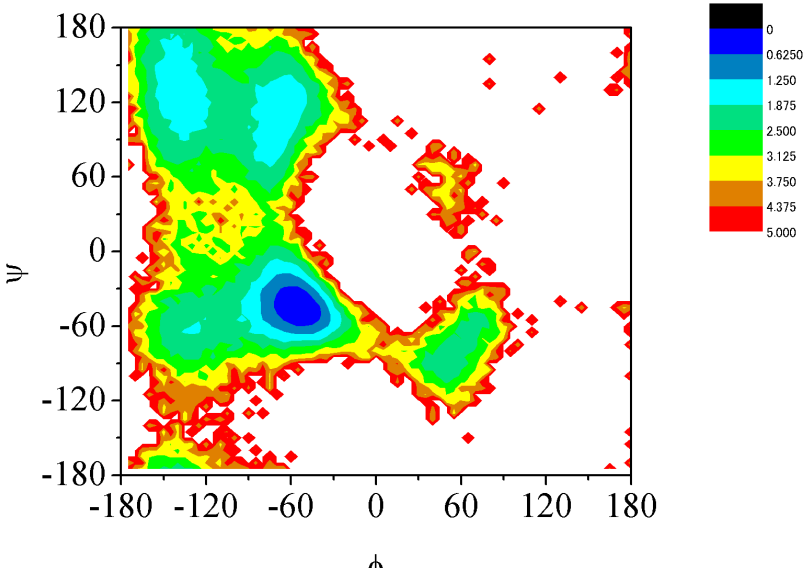
300 K

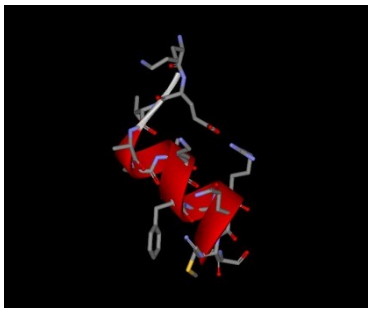
Potential of mean force

$$\Delta G(\phi,\psi) = -k_B T \ln P_B(\phi,\psi)$$

at 300 K

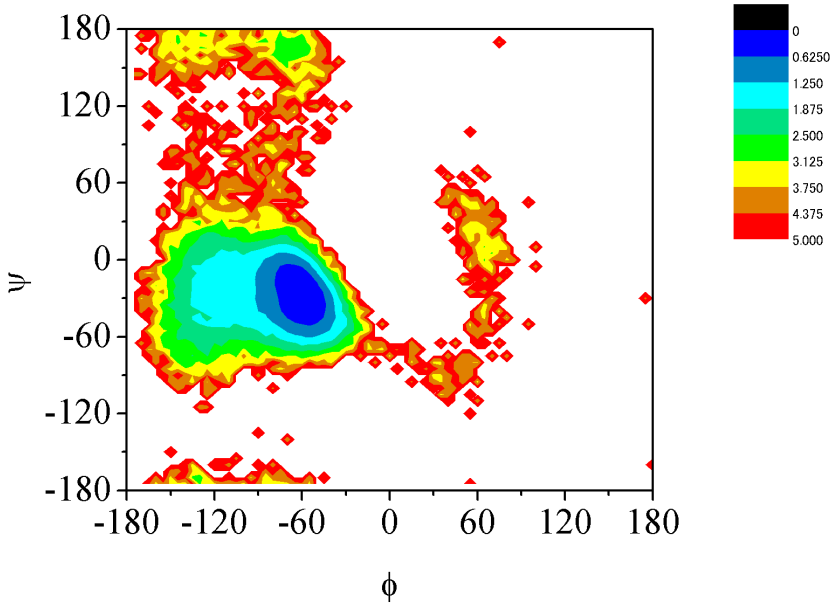
Optimized ff



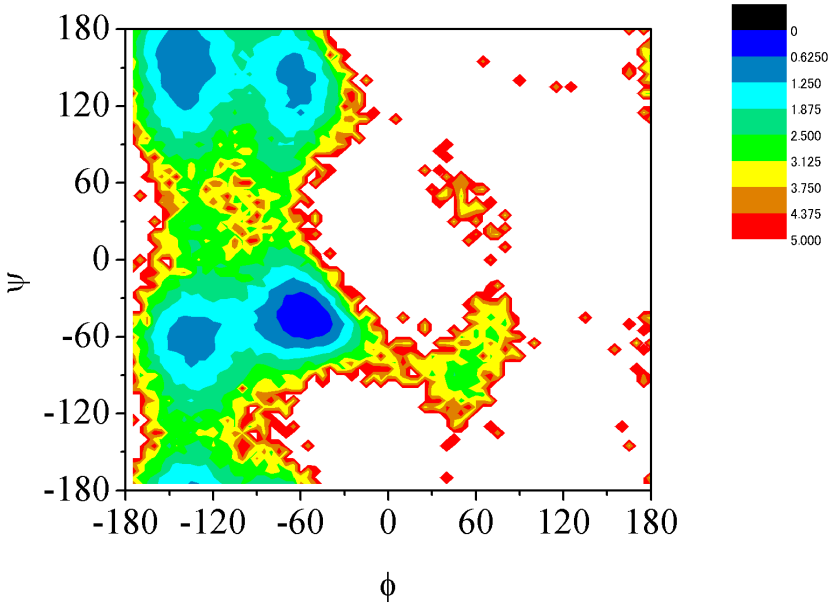


C-peptide

AMBER ff94



AMBER ff96

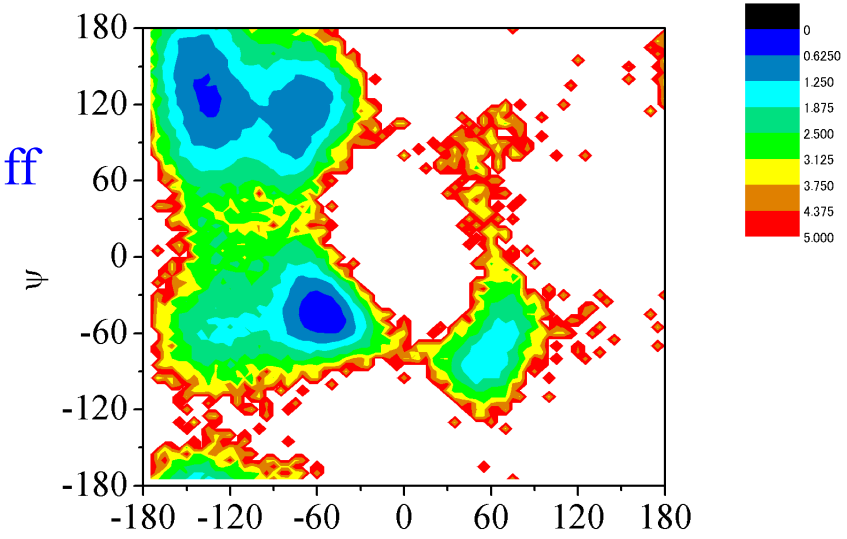


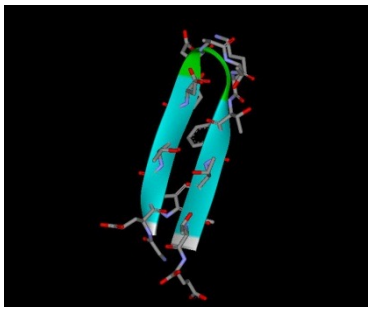
Potential of mean force

$$\Delta G(\phi,\psi) = -k_B T \ln P_B(\phi,\psi)$$

at 300 K

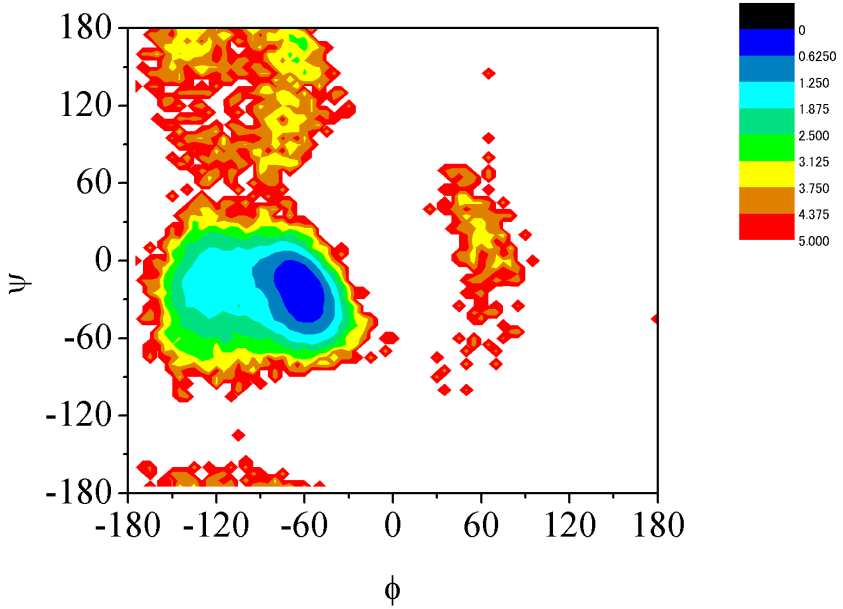
Optimized ff



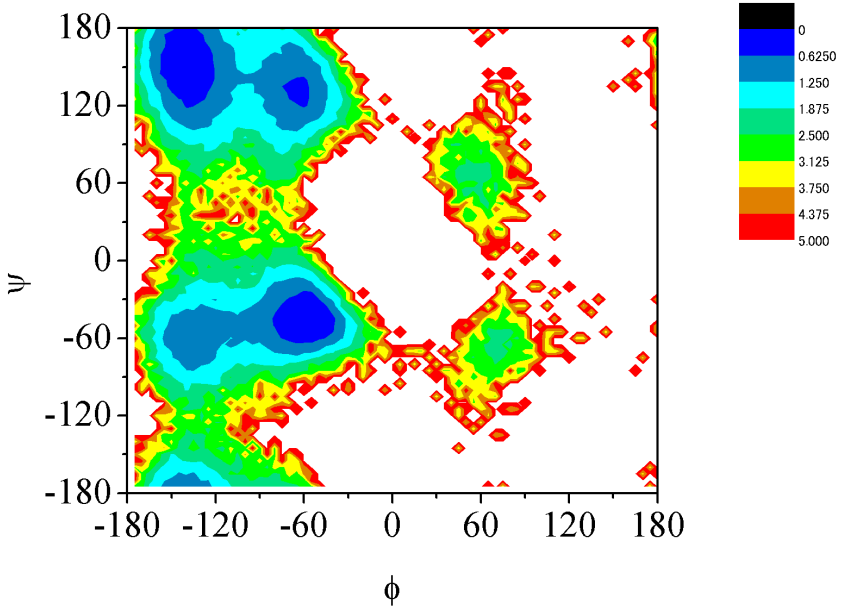


G-peptide

AMBER ff94



AMBER ff96



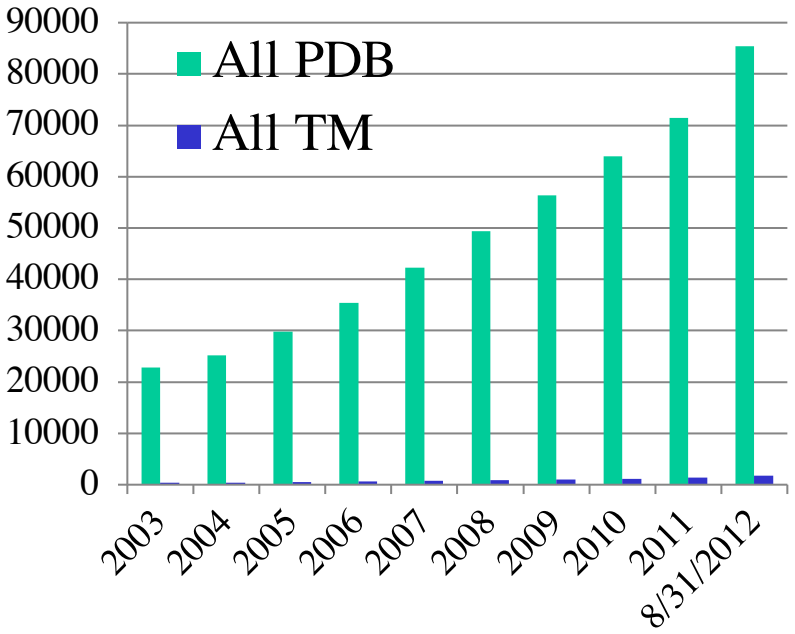
膜タンパク質の立体構造予測
Prediction of Membrane Protein Structures by Replica-Exchange Monte Carlo Simulations

H. Kokubo & Y.O., *Chem. Phys. Lett.* **383**, 397-402 (2004).
H. Kokubo & Y.O., *J. Chem. Phys.* **120**, 10837-10847 (2004).
H. Kokubo & Y.O., *J. Phys. Soc. Jpn.* **73**, 2571-2585 (2004).
H. Kokubo & Y.O., *Chem. Phys. Lett.* **392**, 168-175 (2004).
H. Kokubo & Y.O., *Biophys. J.* **96**, 765-776 (2009).
R. Urano, H. Kokubo & Y.O., *J. Phys. Soc. Jpn.* 84, 084802 (2015).
R. Urano & Y.O., *J. Chem. Phys.* **143**, 235101 (2015).

Motivation:

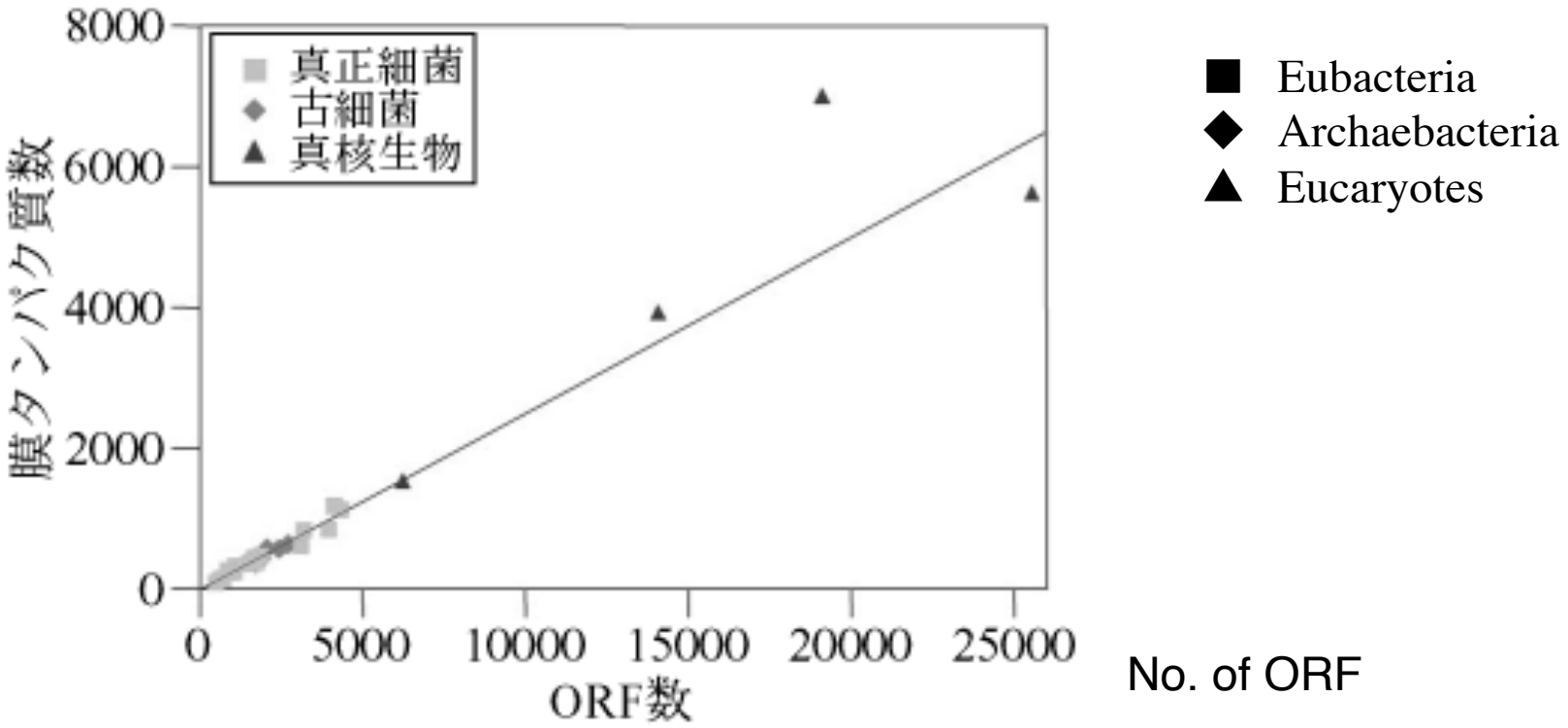
Protein Data Bank (PDB) :

194,259 entries as of August 17, 2022, but only about 2 % of them are membrane protein structures .



It is estimated that 20-30 % of all genes in most genomes encode membrane proteins. However, only a small number of detailed structures have been obtained for membrane proteins because of technical difficulties such as high quality crystal growth.

No. of Membrane Proteins



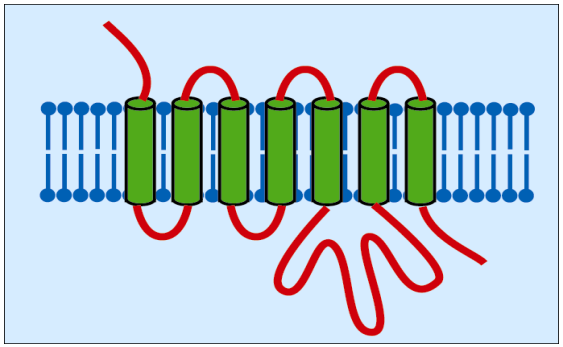
S. Mitaku, *Biophysics* **42**, (2002) 104-109 (in Japanese).
A. Krogh et al., *J. Mol. Biol*. **305** (2001) 567-580.

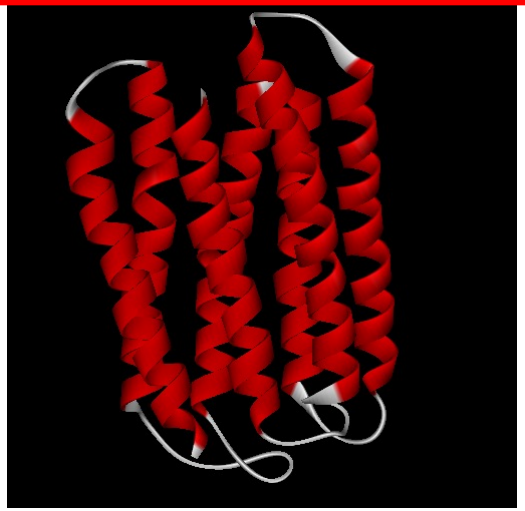
We can know **Membrane Spanning Regions** by **Prediction Tool on the Web:** SOSUI,HMMTOP,TMHMM..etc.

Predict transmembrane helix configurations by molecular simulations

Sequence

MNGTEGPNFYVP、、、、





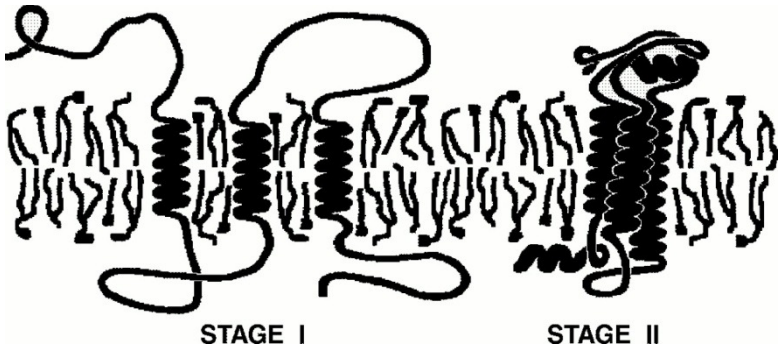
Tertiary Structure

No.	N terminal	transmembrane region	C terminal	type	length
1	10	IWLALGTALMGLGTLYFLVKGMG	32	SECONDARY	23
2	45	TTLVPAIAFTMYLSMLLGYGLTM	67	SECONDARY	23
3	77	IYWARYADWLFTTPLLLLDLALL	99	SECONDARY	23
4	106	TILALVGADGIMIGTGLVGALTK	128	PRIMARY	23
5	136	WWAISTAAMLYILYVLFFGFTSK	158	PRIMARY	23
6	176	VTVVLWSAYPVVWLIGSEGAGIV	198	PRIMARY	23
7	203	ETLLFMVLDVSAKVGFGLILLR	224	PRIMARY	22

Two-stage model

Individual helices of a protein are postulated to be stable separately as domains in a lipid bilayer. And then side-to-side helix association is driven, resulting in a functional protein.

J.-L. Popot and D.M. Engelman, *Annu. Rev. Biochem.* **69**, 881 (2000).

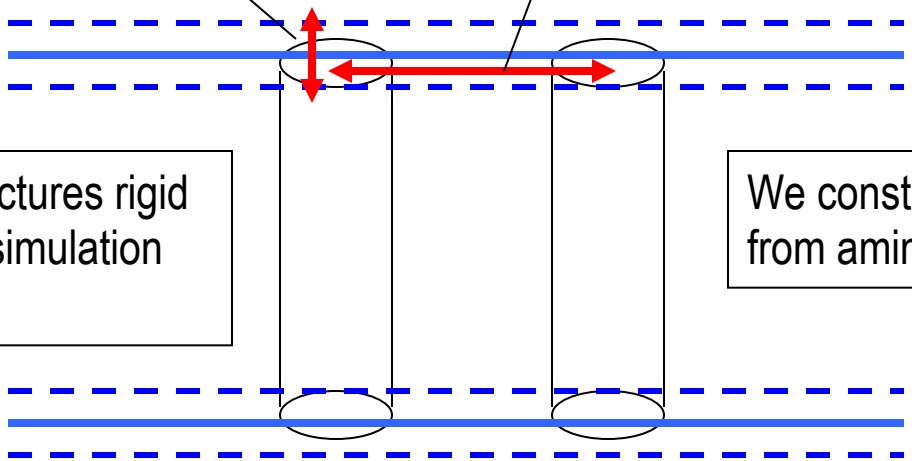


- Individual helix stability as domains is a consequence of the hydrophobic effect and main-chain hydrogen bonding.
- Specific folding energy is provided mainly by packing of the preformed helices with each other, by loop structures, and by interactions with prosthetic groups.
- Additionally, ion pairs and hydrogen bonds between helices are sometimes found, and general contributions are made by interactions with the lipid environment.

Simulation Conditions

Harmonic constraint are set only when helices are apart at some distance from membrane boundary planes.

Harmonic constraint are set only when helices are apart at some distance.



We keep backbone structures rigid in order to perform the simulation with helix structures.

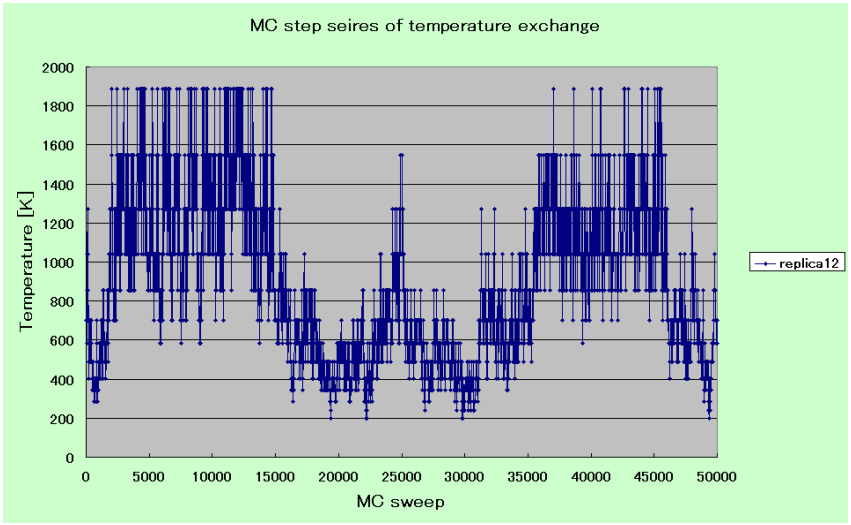
We construct ideal α-helices from amino-acid sequence.

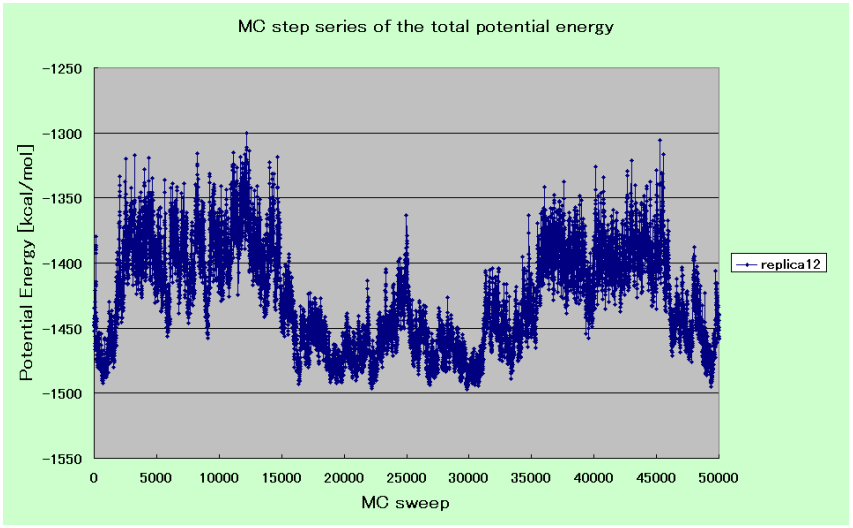
We neglect loop region amino acids, lipids, and water.

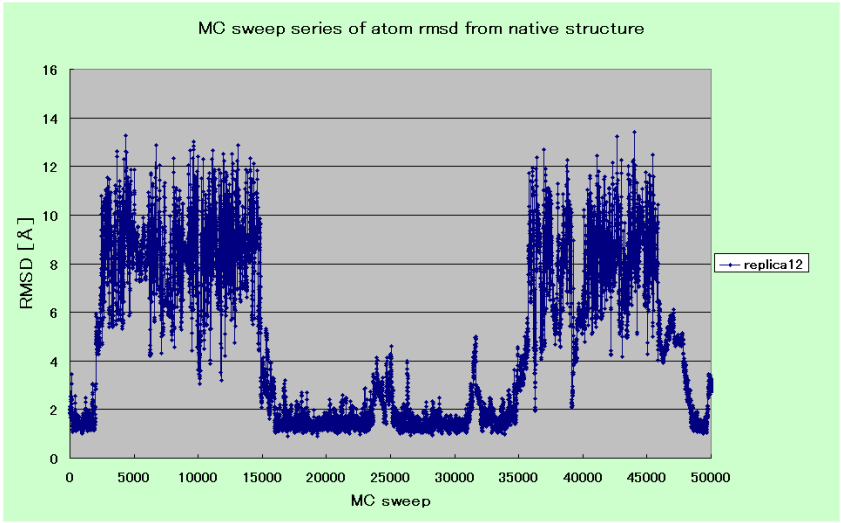
CHARMM force-field (ver. 19)
Replica-exchange Monte Carlo simulation (rigid translation, rigid rotation of each helix and torsion angle rotation of side chains)

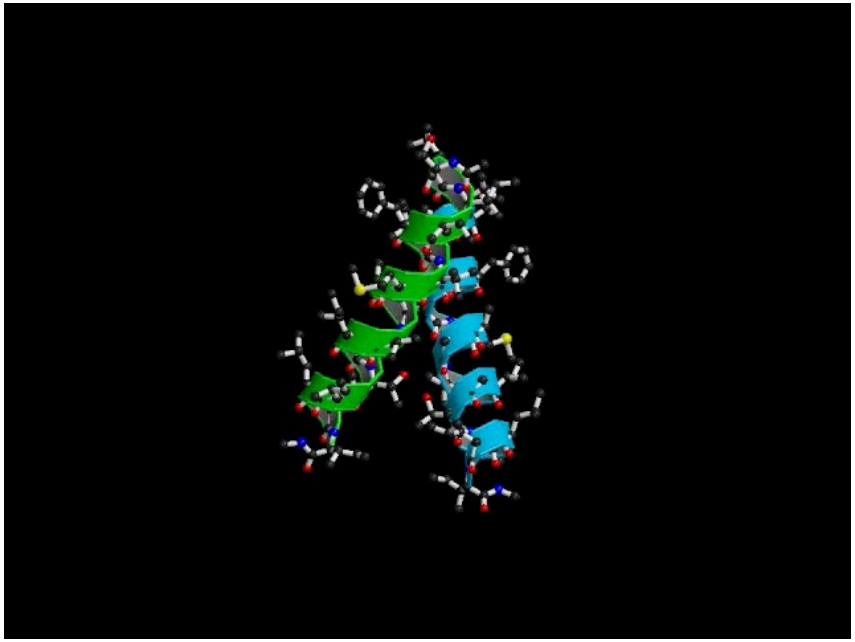
DIMERIC TRANSMEMBRANE DOMAIN OF HUMAN GLYCOPHORIN A

Simulation and movie by H. Kokubo

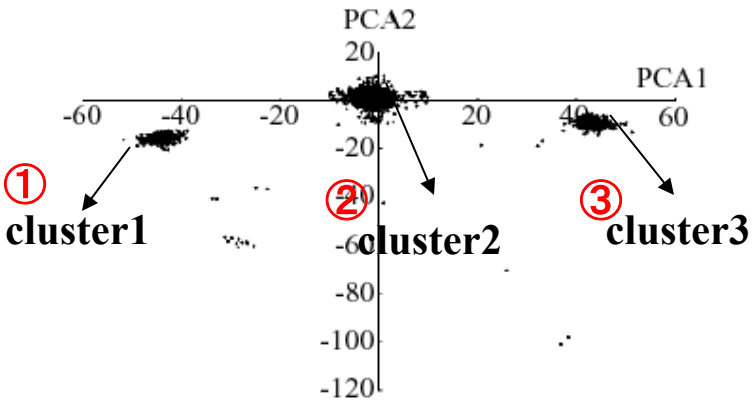




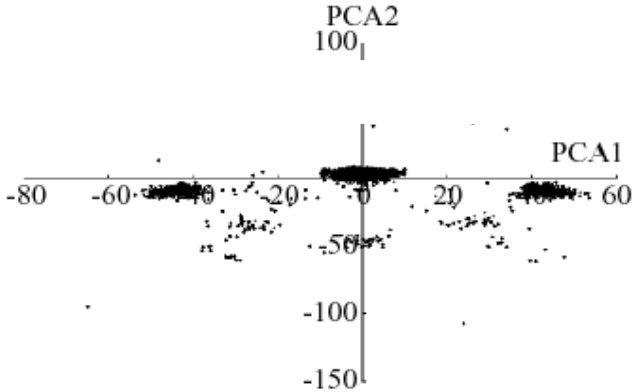




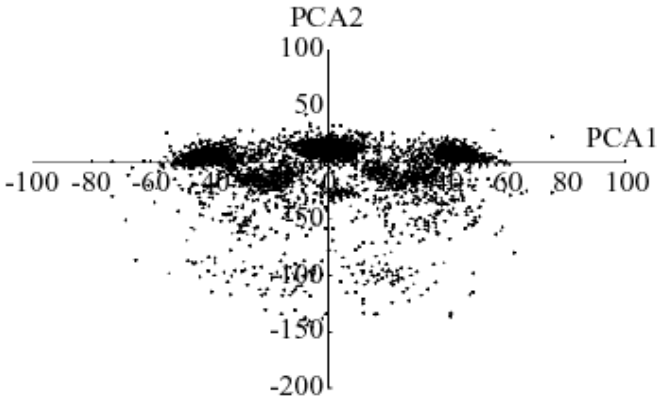
Principal Component Analysis: case for ε=1.0



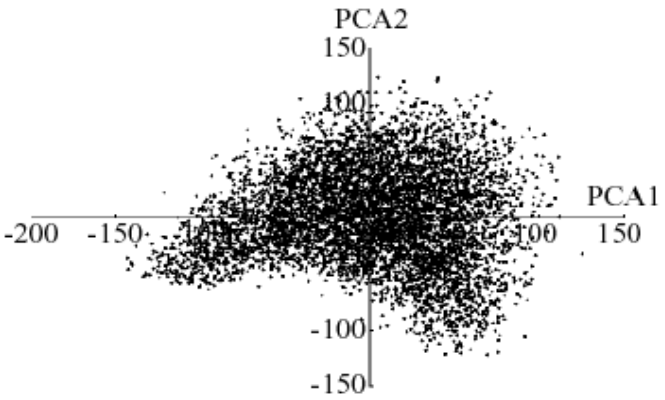
200 K



342 K

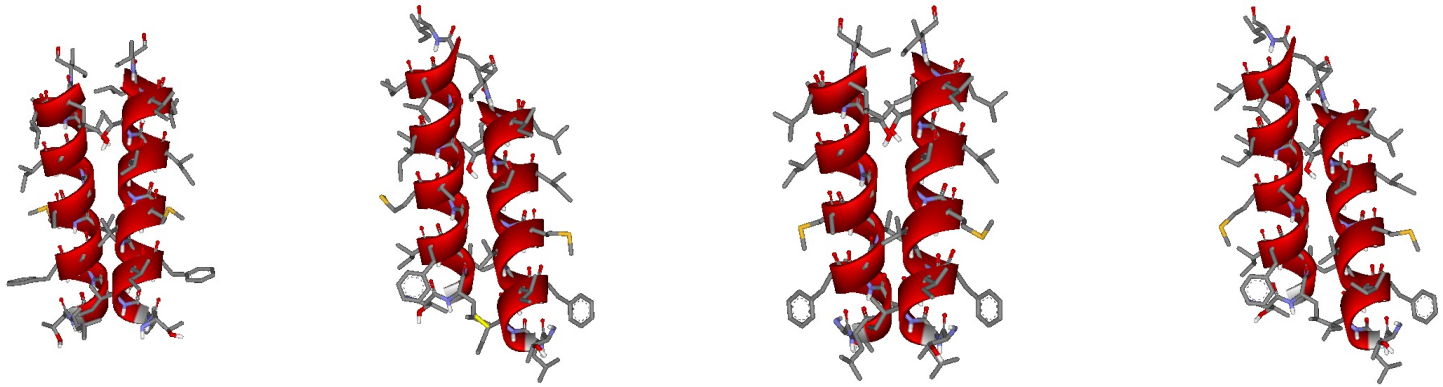


585 K

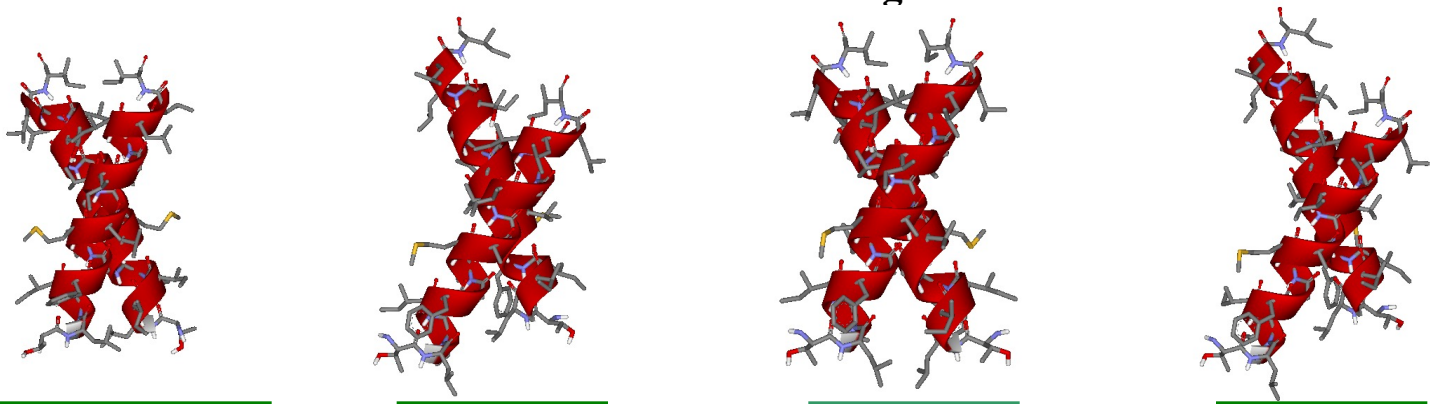


1888 K

Structure of Each Cluster: case for ε=1.0



Viewed from different angles

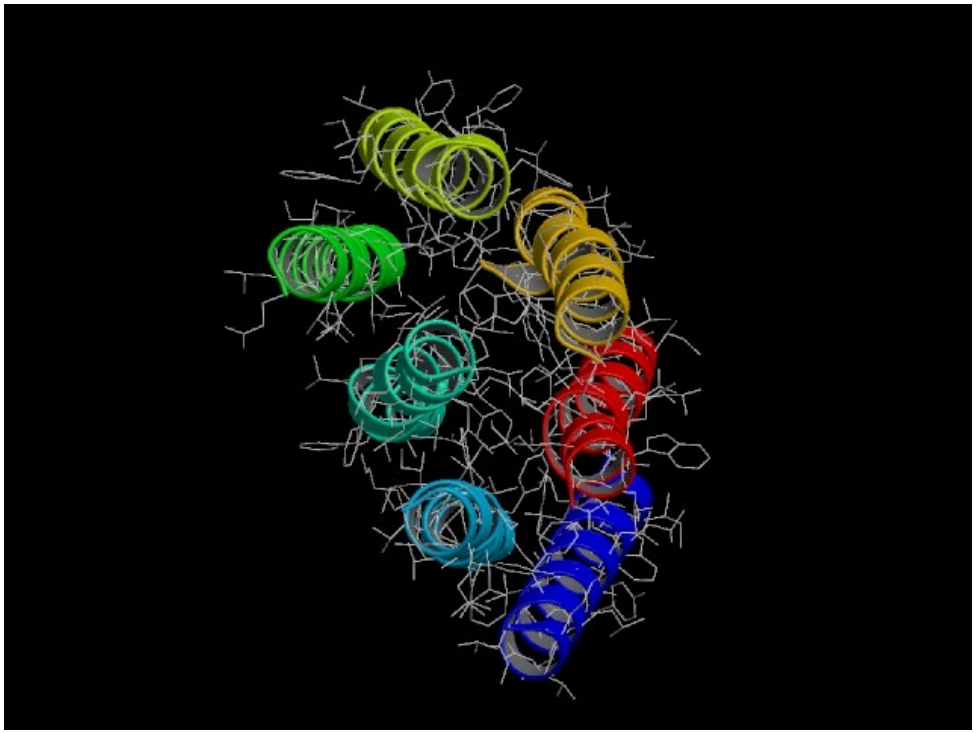


Native(NMR) | **Cluster1** | **Cluster2** | **Cluster3**

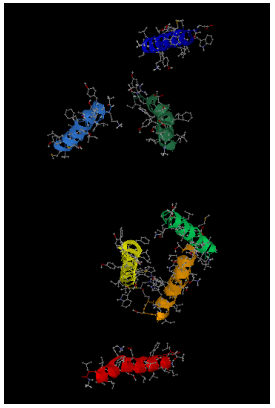
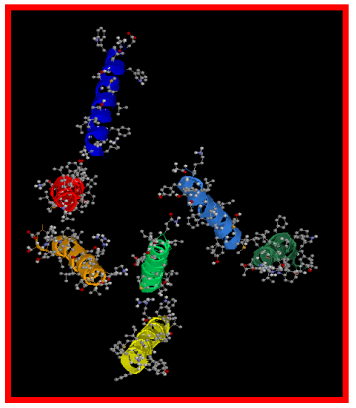
H. Kokubo & Y.O., *Chem. Phys. Lett.* **392**, 168-175 (2004).
H. Kokubo & Y.O., *Biophys. J.* **96**, 765-776 (2009).

Bacteriorhodopsin （Case of 7 Helices）

Native Structure

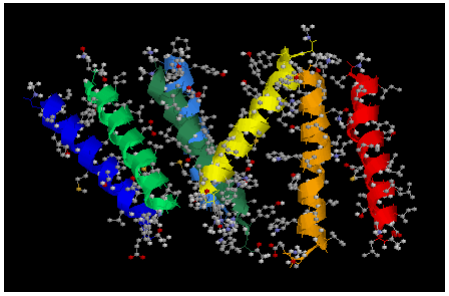


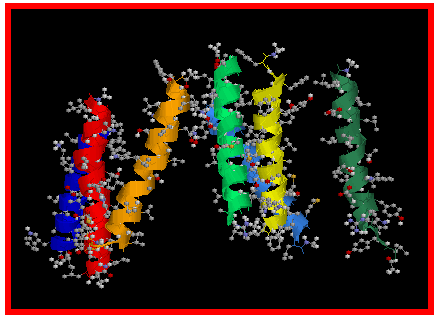
Initial structures of the simulation

From the top

From the side





Replica1 **Replica2** **Replica14**

Initial structures

32 random initial configurations obtained by simulations at the highest temperature

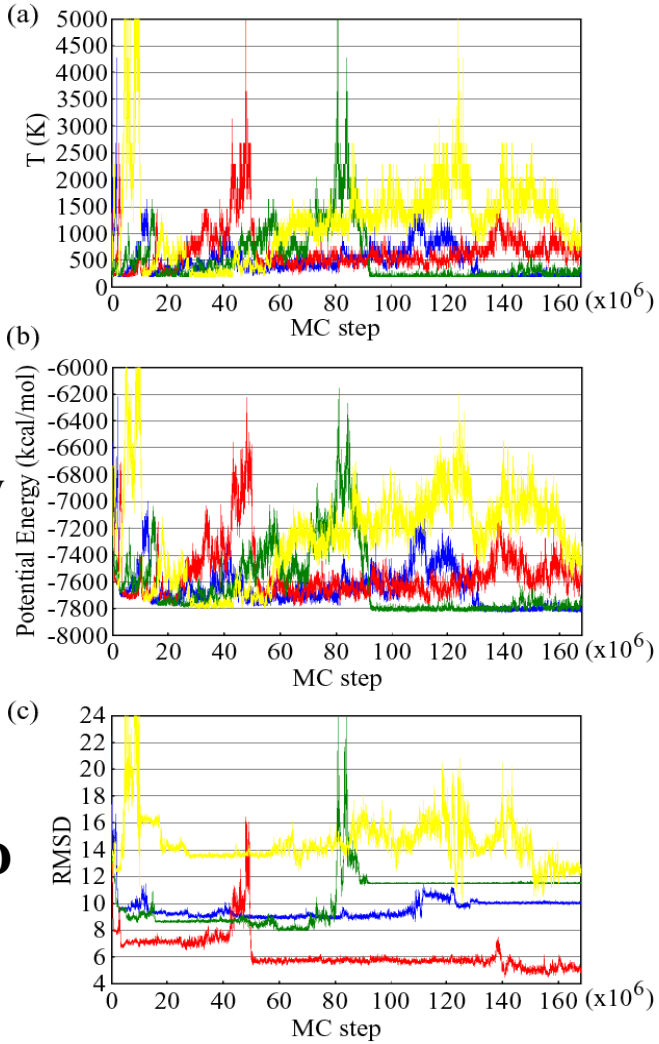
The replica-exchange MC simulation of 32 replicas

MC time series

T

Potential Energy

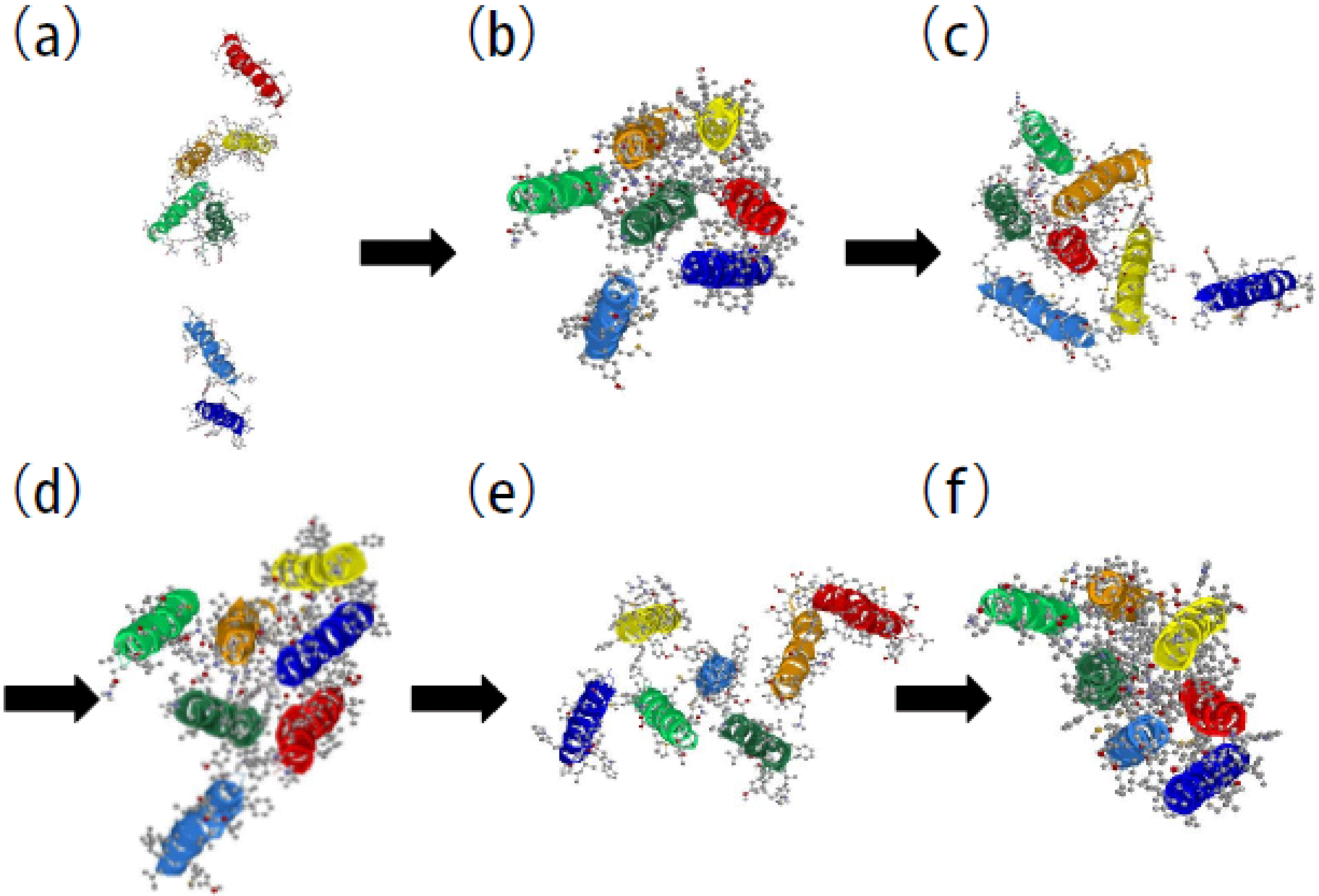
RMSD



MC time series

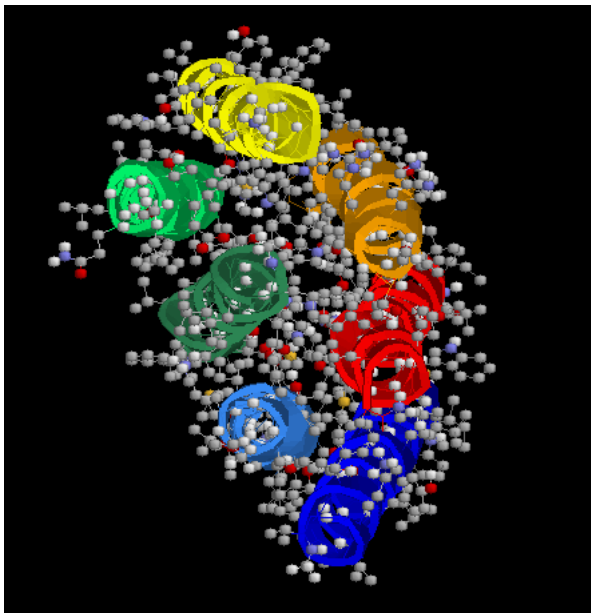
Replica6 Replica14 Replica19 Replica25

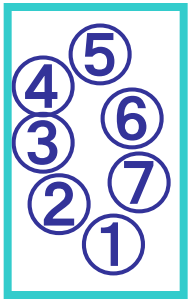
Snapshots (Replica 16)

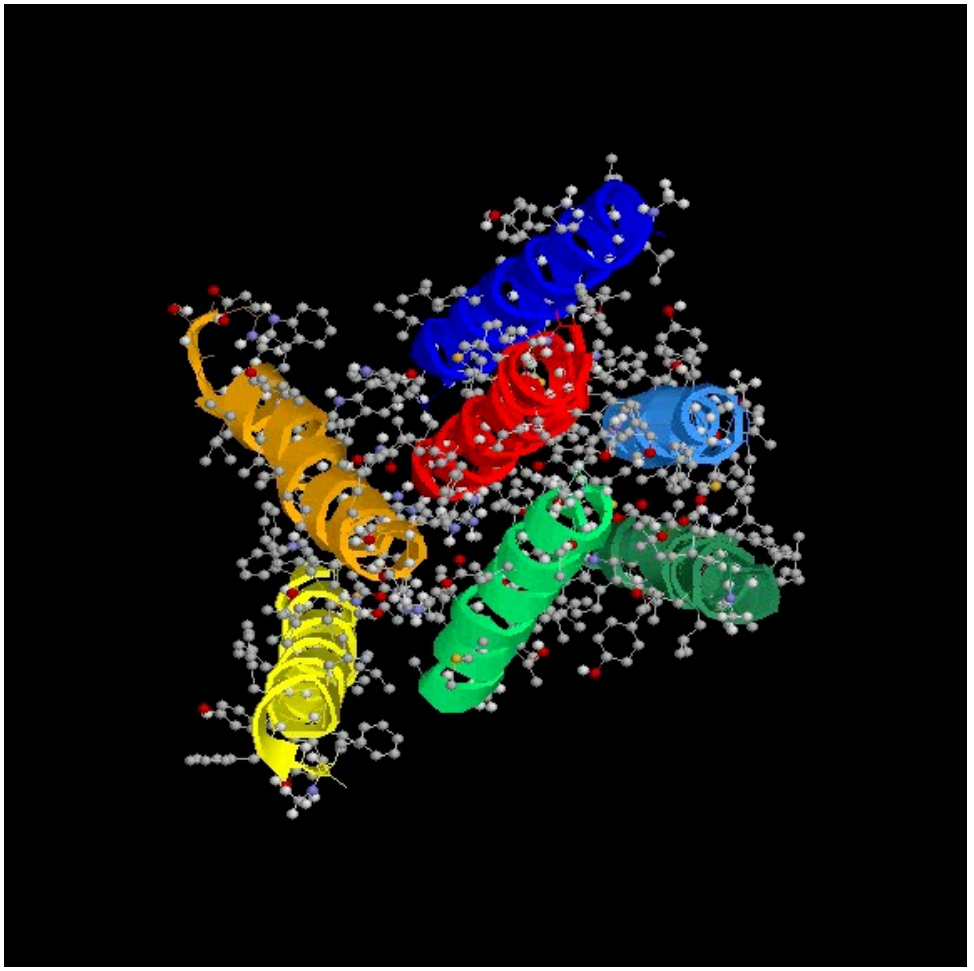


REMC (from above)

Replica 14





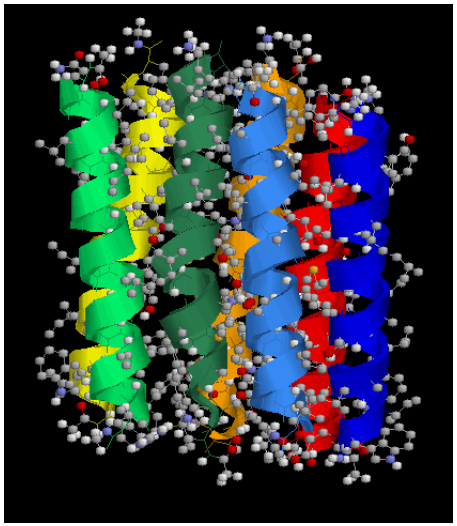


Native structure

Simulation and movie by H. Kokubo

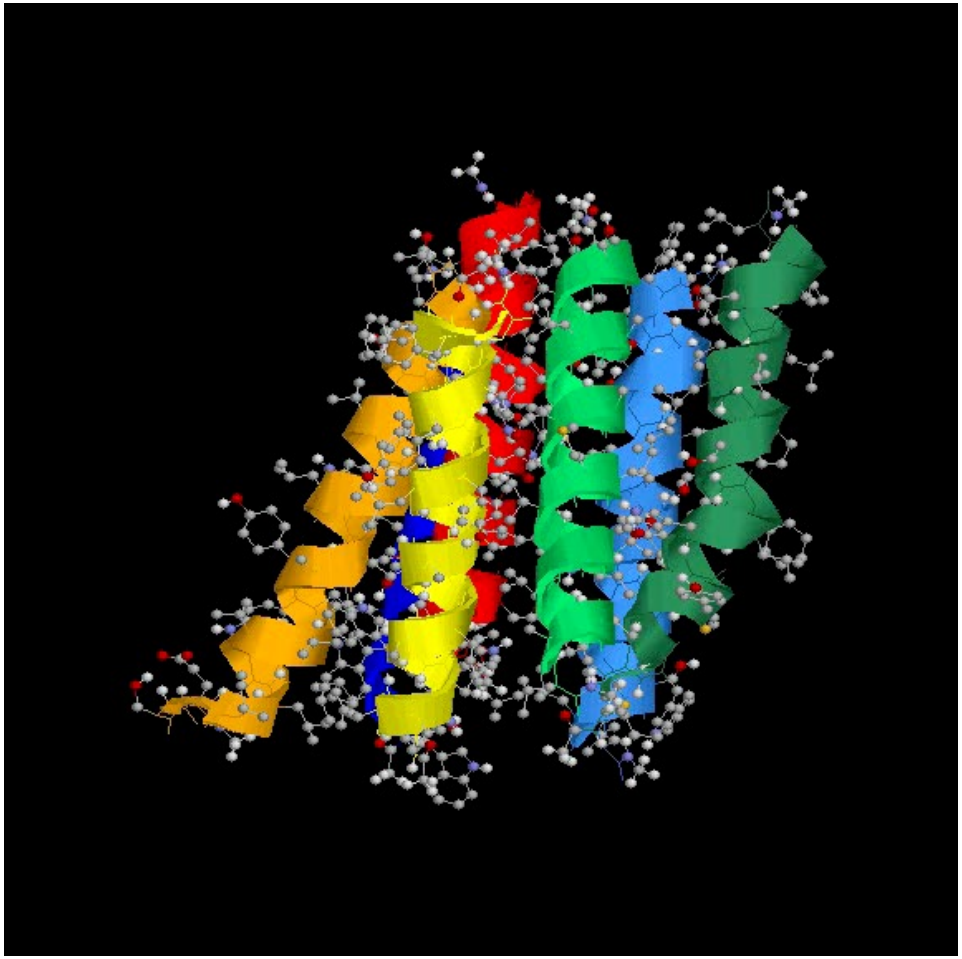
REMC
(from above)

Replica 14



5 6 7
4 3 2 1

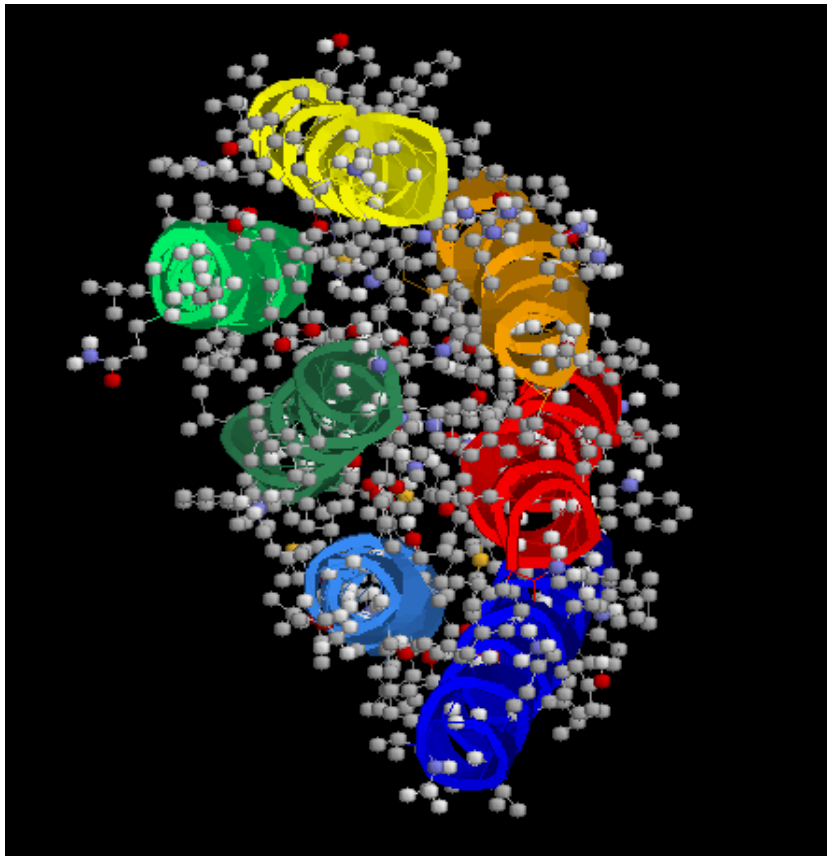
Native structure



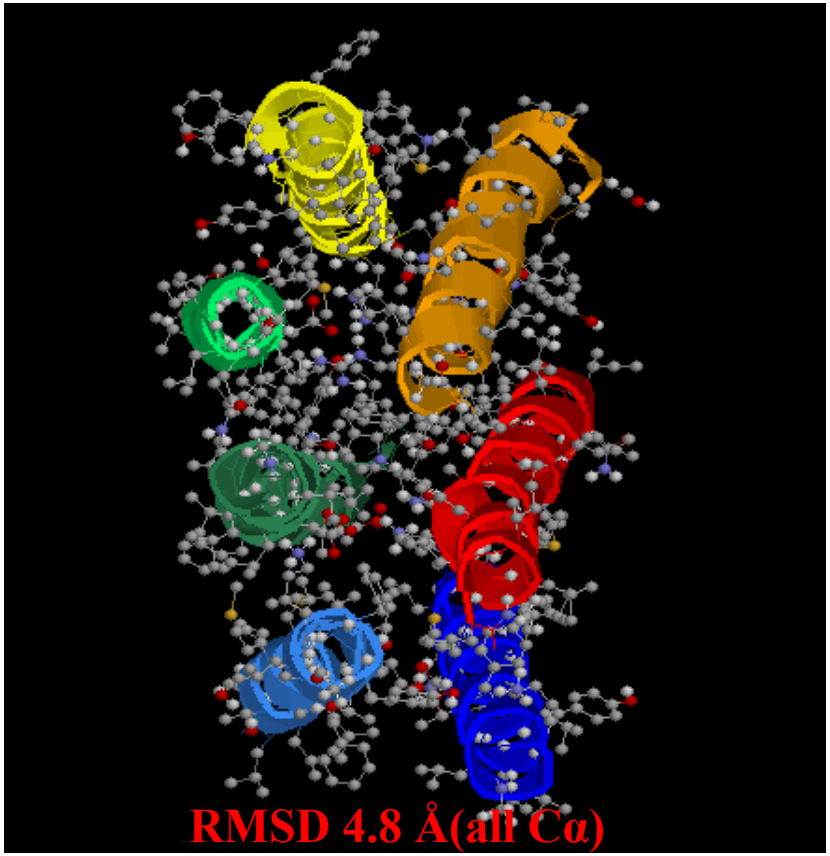
Simulation and movie by H. Kokubo

H. Kokubo & Y.O., *Biophys. J.* **96**, 765-776 (2009).

Local Minimum Structure



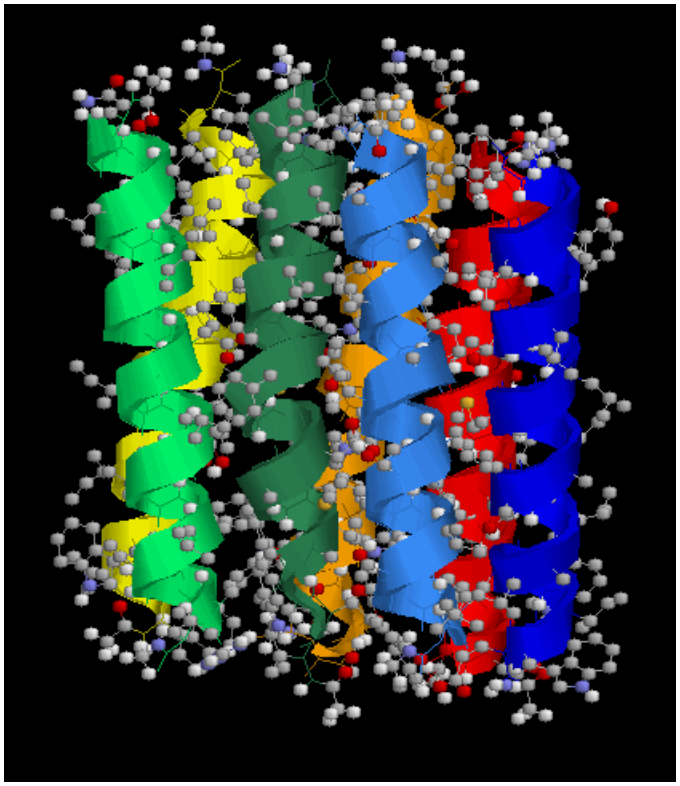
Native Structure



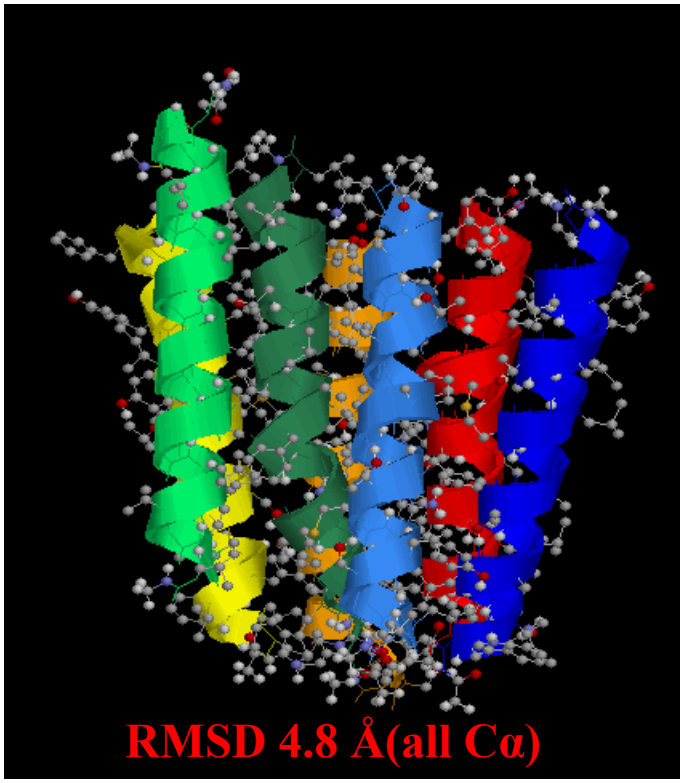
Replica 14

H. Kokubo & Y.O., *Biophys. J.* **96**, 765-776 (2009).

Local Minimum Structure



Native Structure



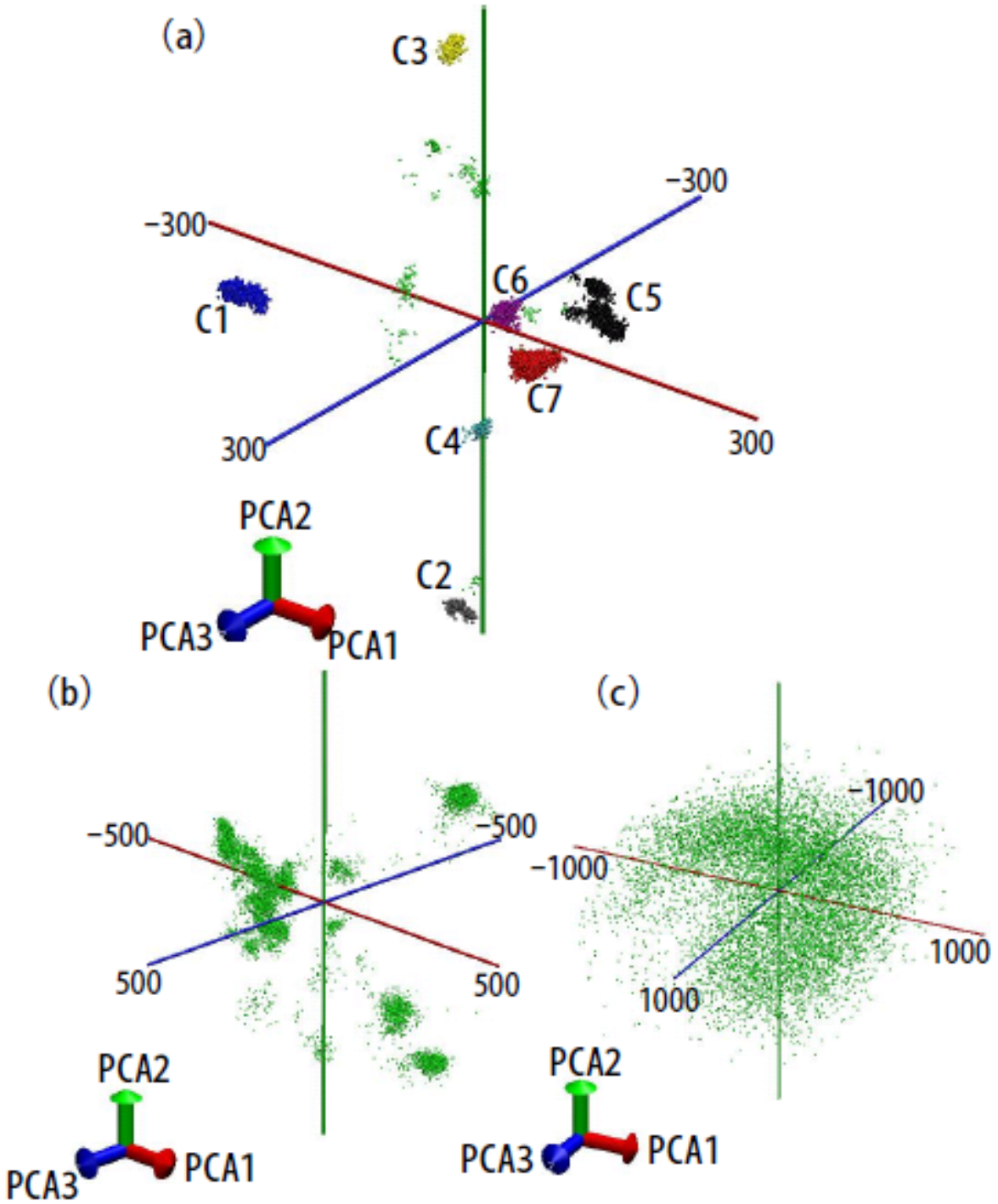
Replica 14

Principal Component Analysis at

(a) $T = 500$ K

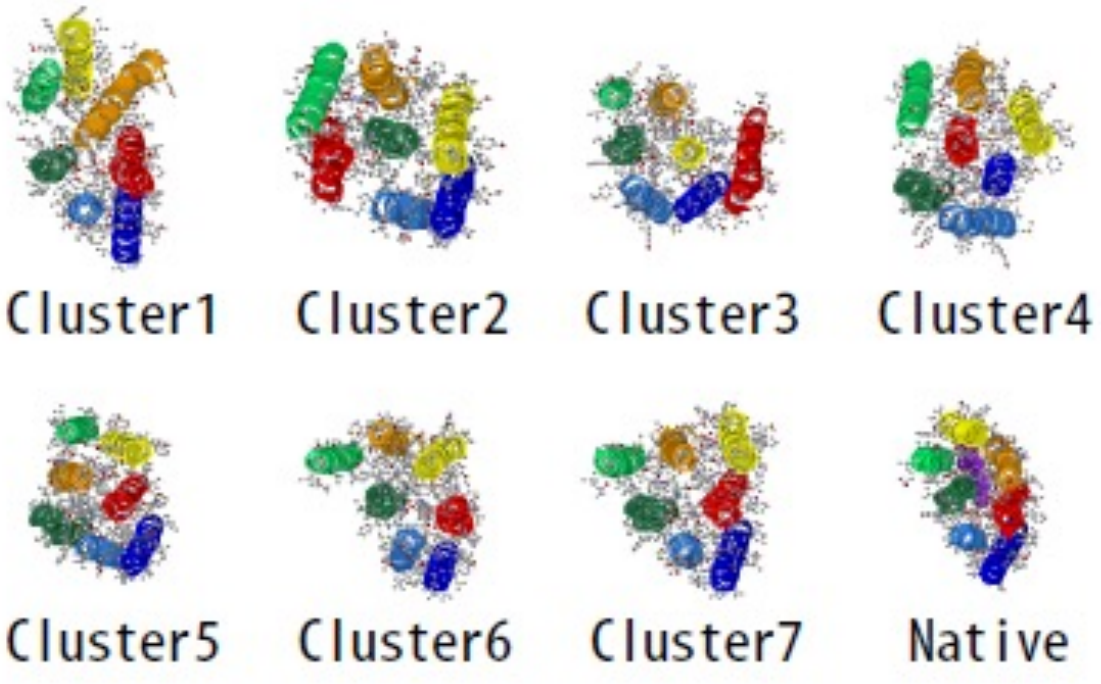
(b) $T = 976$ K

(c) $T = 5000$ K

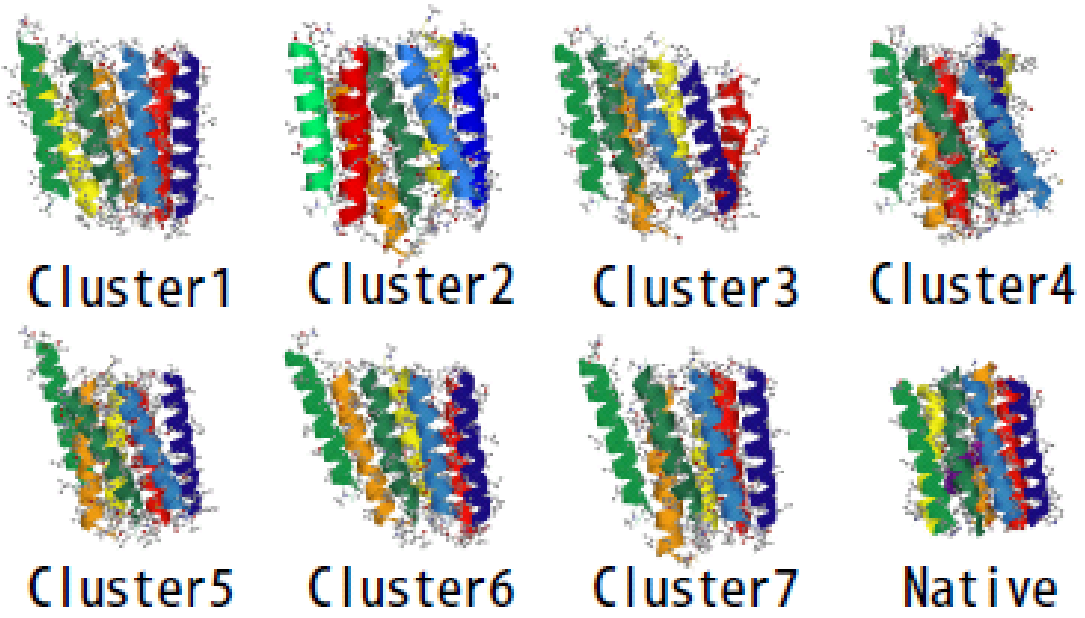


Typical Structures of the 7 Clusters

Top view



Side view



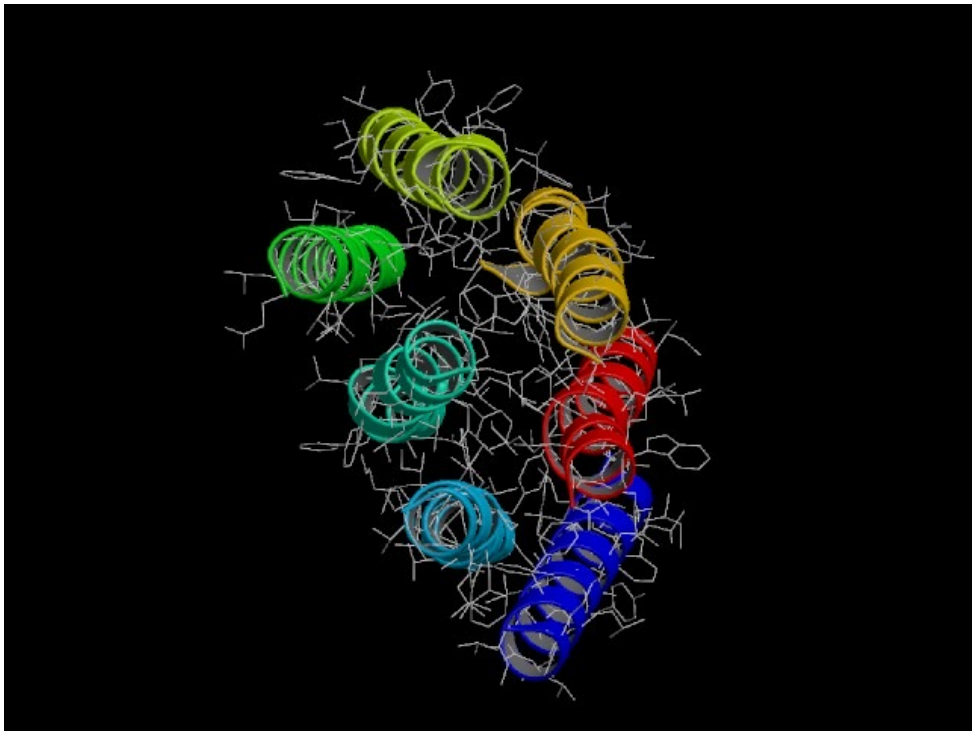
H. Kokubo & Y.O.,
Biophys. J. **96**, 765-776 (2009).

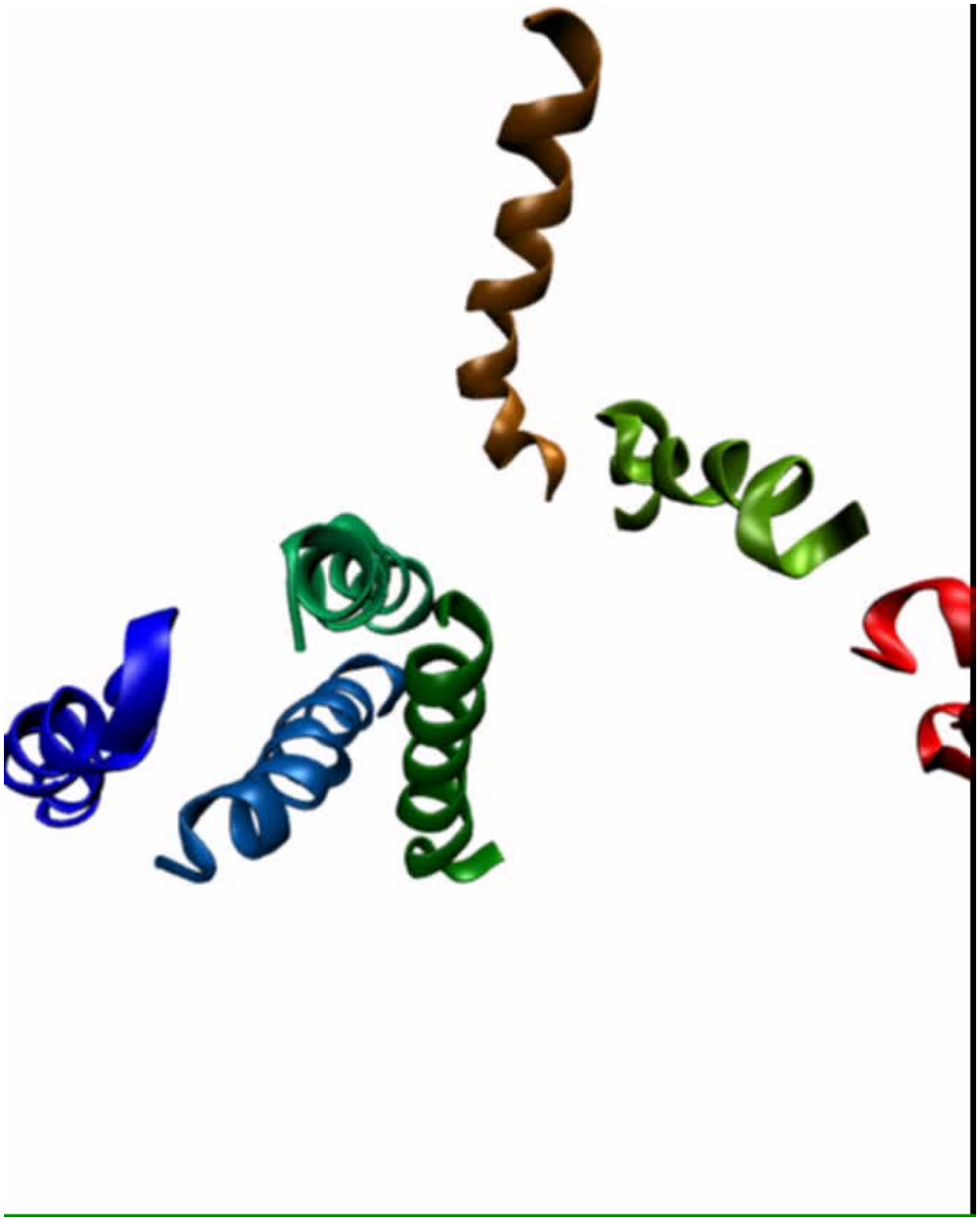
Bacteriorhodopsin （Case of 7 Helices）

R. Urano & Y.O., *J. Chem. Phys.* **143**, 235101 (2015).

Introduced bending of TM helices

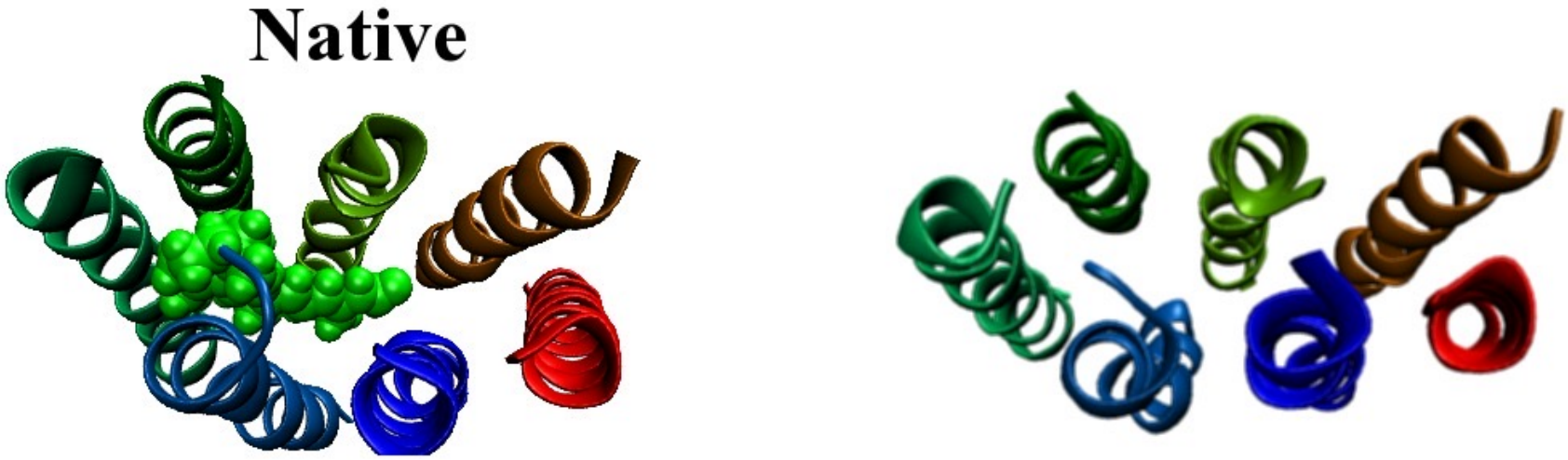
Native Structure





Simulation and movie by R. Urano

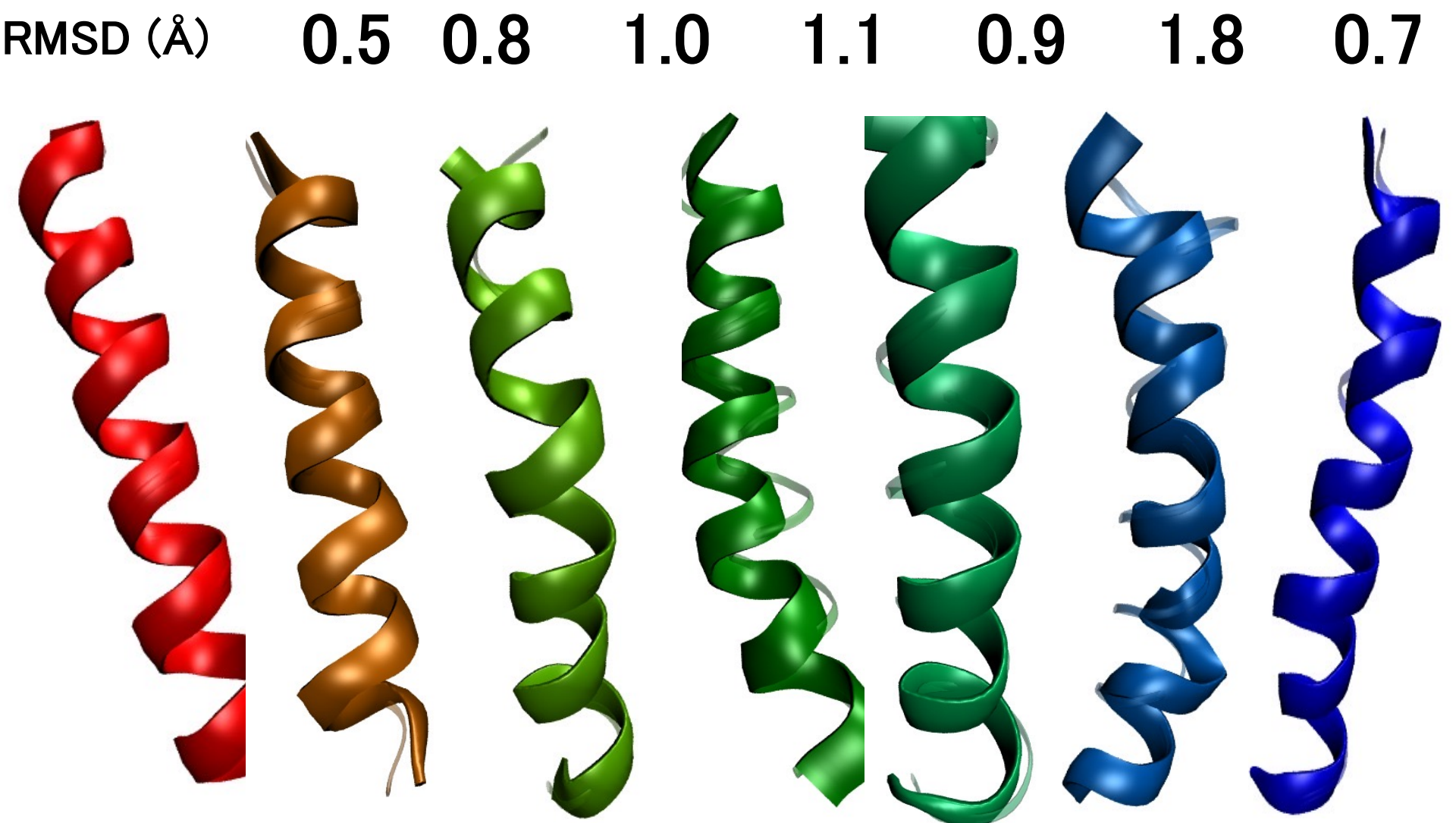
Native Structure (left) and Global-Minimum Free-Energy Structure (right) from REM Simulations



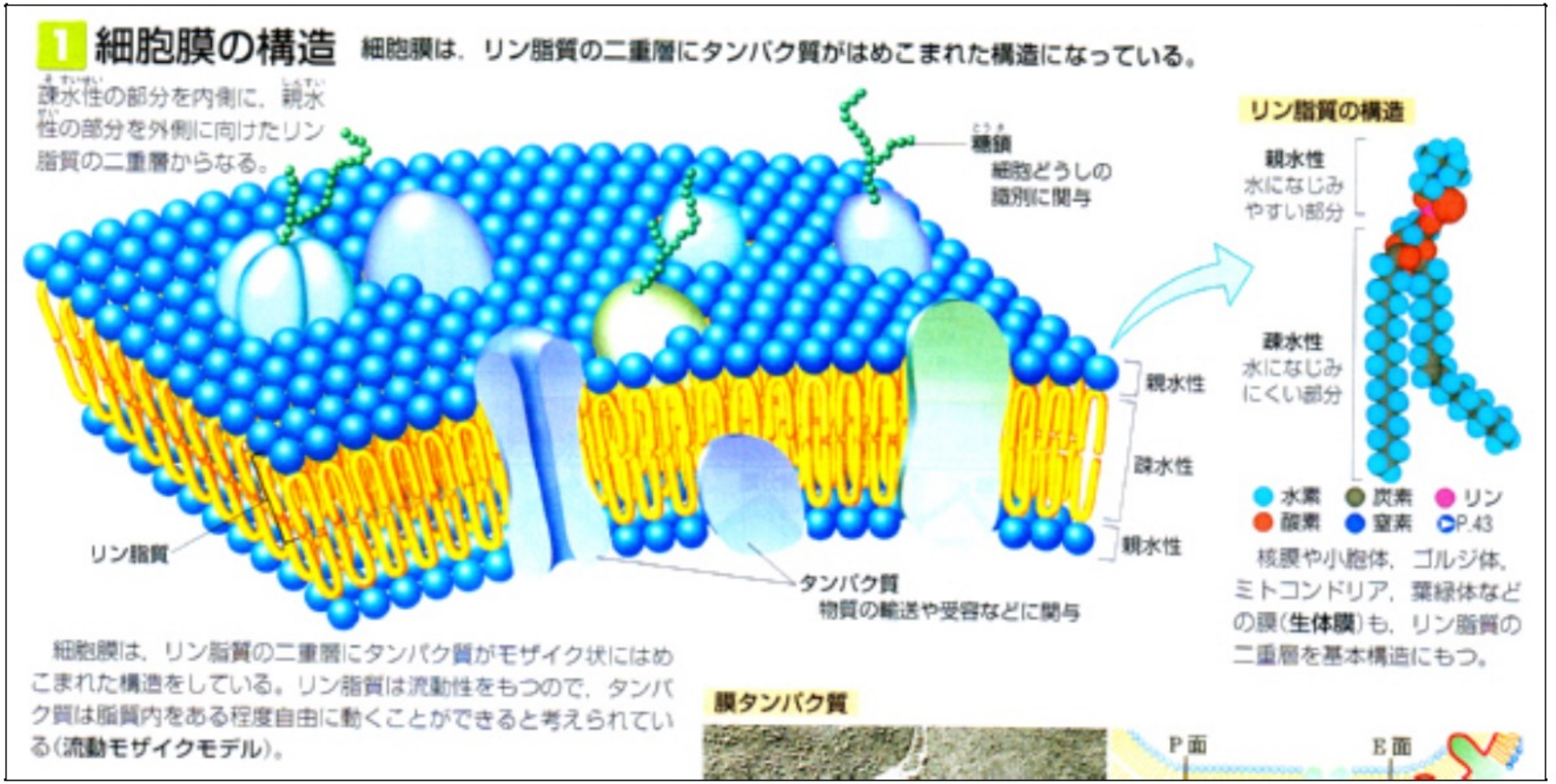
R. Urano & Y.O., *J. Chem. Phys*. **143**, 235101 (2015).

Bending of TM helices reproduced.

R. Urano & Y.O., *J. Chem. Phys.* **143**, 235101 (2015).



【細胞膜の構造】(細胞膜の構造図は浜島書店の生物図表より)



リン脂質による二重膜構造。膜貫通タンパク質がある。

http://spider.art.coocan.jp/biology2/cell2012.html

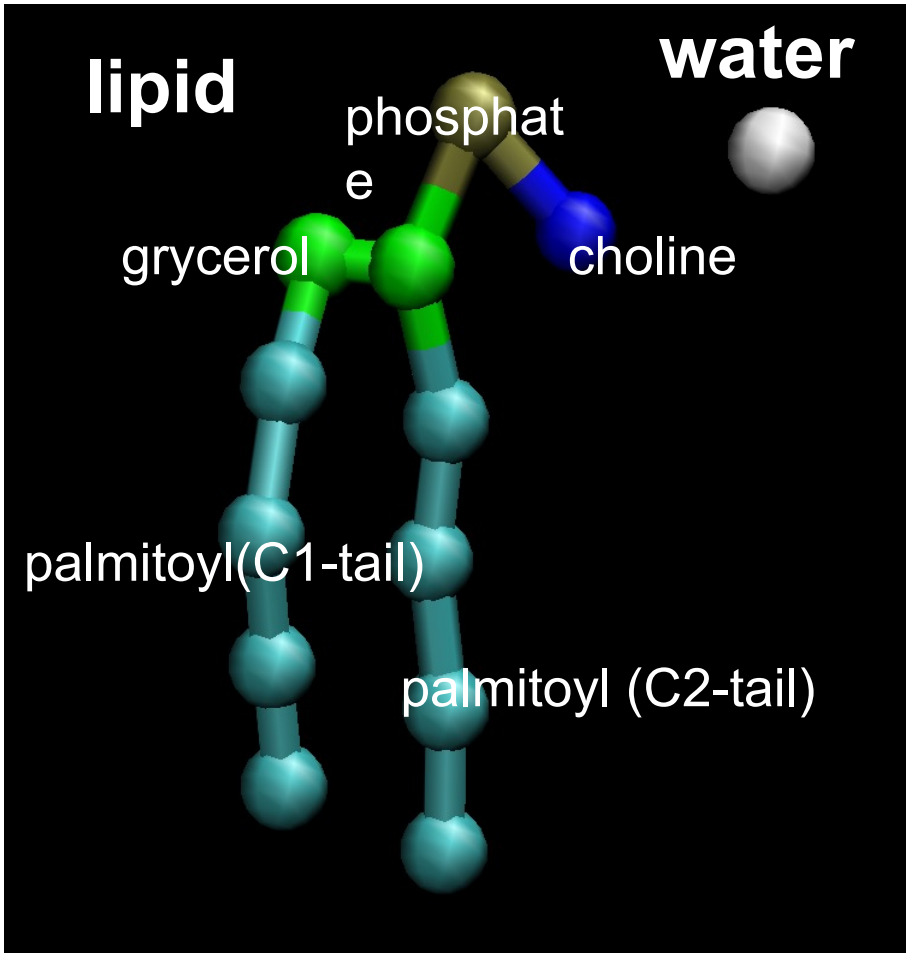
T. Nagai, R. Ueoka & Y.O., *J. Phys. Soc. Jpn.* **81**, 024002 (2012).

Phase behavior of DPPC bilayer studied by REMD with a coarse-grained model

MARTINI 2.0

a coarse-grained model.

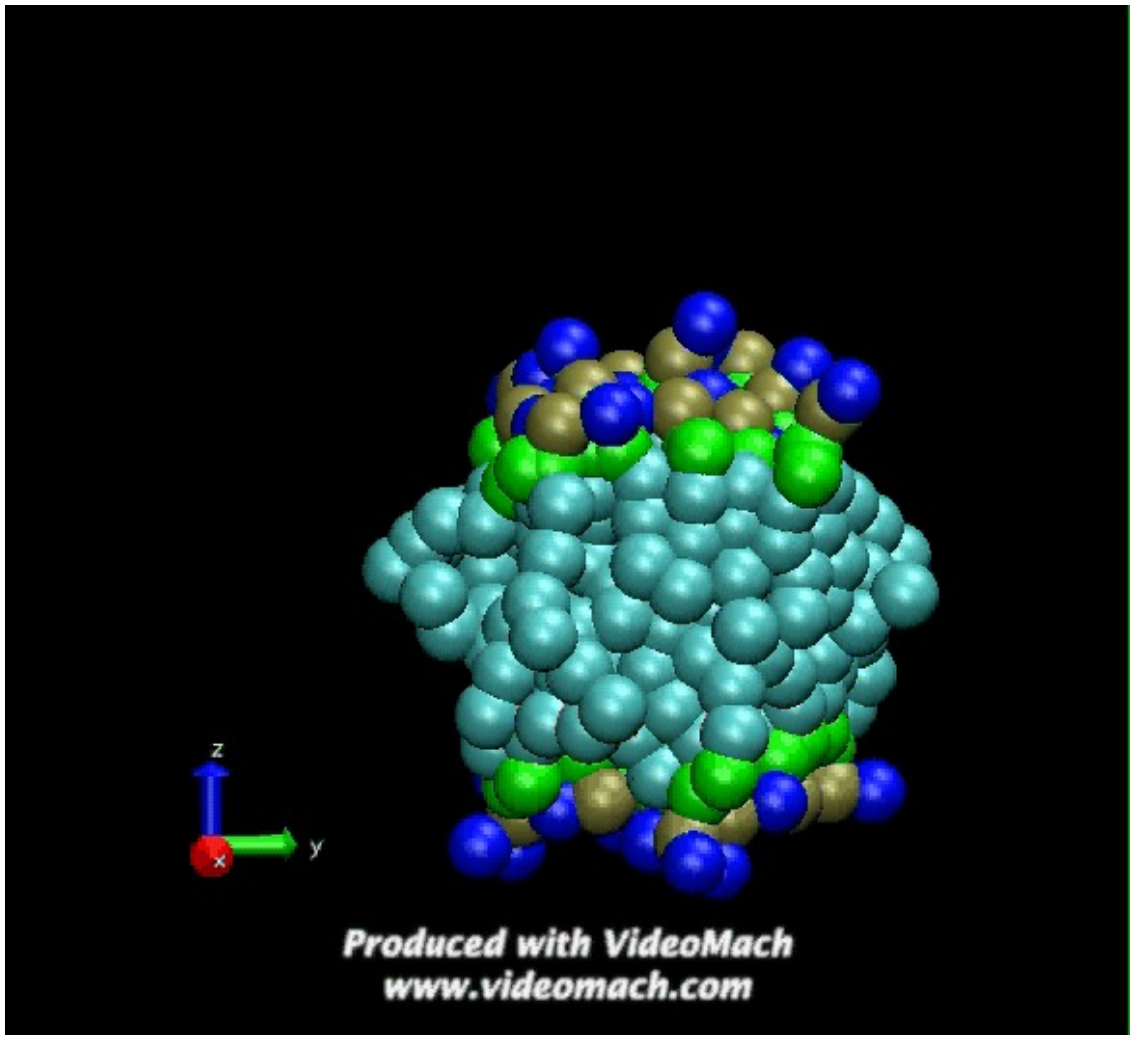
four atoms are treated as one site, so that larger system size and/or time scale can be studied.



Simulation of a 32-lipid system

- system
  - 32 DPPC and 500 water particles (including 50 anti-freeze particles)
- temperature
  - Thermostat: Nose-Hoover method in the production run
  - Temperature distribution: 127 points between283 K and 390 K
- pressure
  - Reference pressure: 1atom
  - Barostat: Parrinello-Raman method in the production run
- Frequency of replica exchange
  - Every 100 step
- Time step
  - 20 fs

Results
Animation



A trajectory of one of the replicas.

Water is suppressed for clarity.

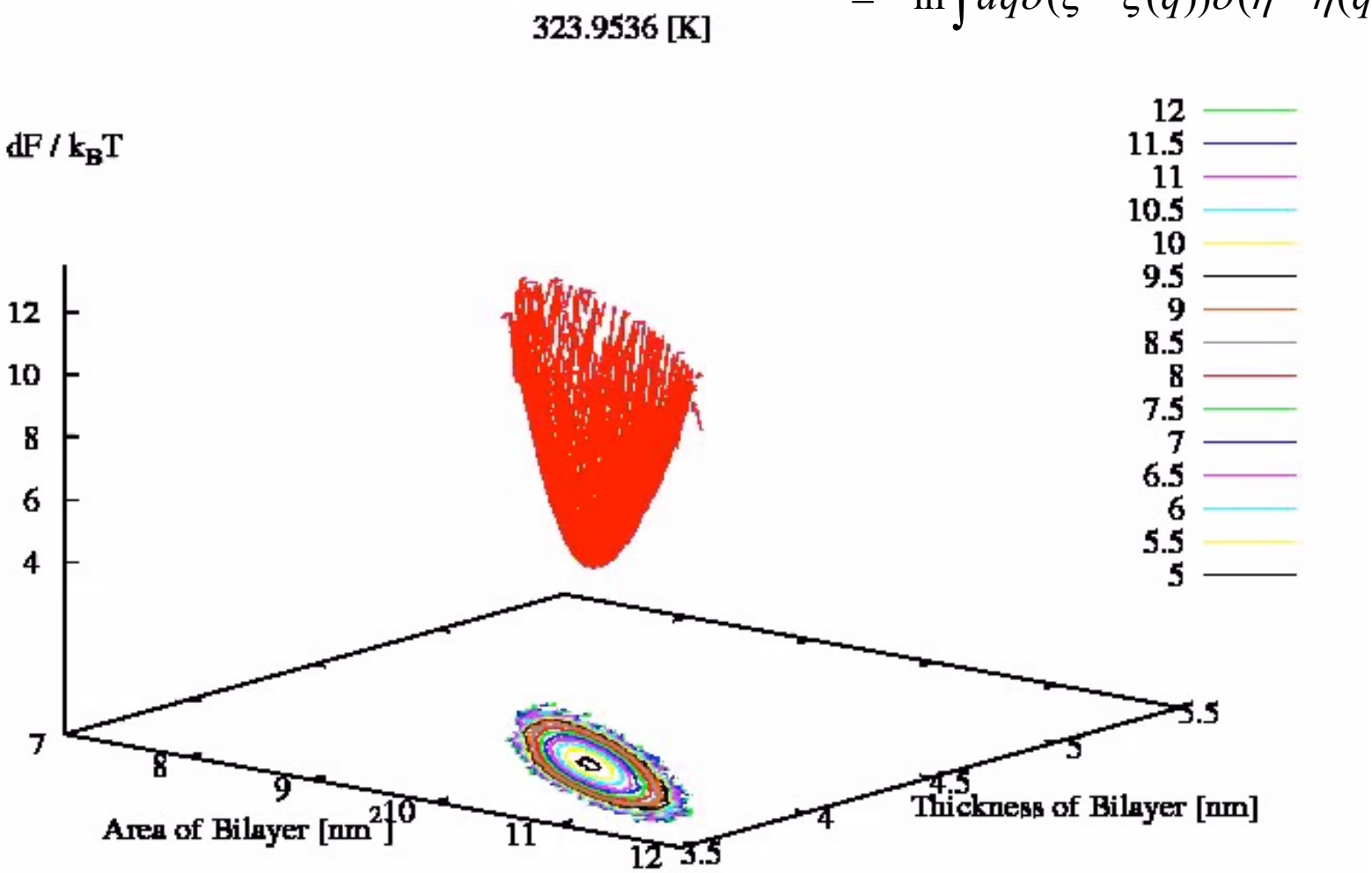
Simulation and movie by T. Nagai

Results

T. Nagai, R. Ueoka & Y.O., *J. Phys. Soc. Jpn.* **81**, 024002 (2012).

Change of Potential of Mean Force (PMF)

$$\beta f(\xi,\eta) = -\ln \int dq\,\delta(\xi-\xi(q))\delta(\eta-\eta(q))e^{-\beta E(q)}$$



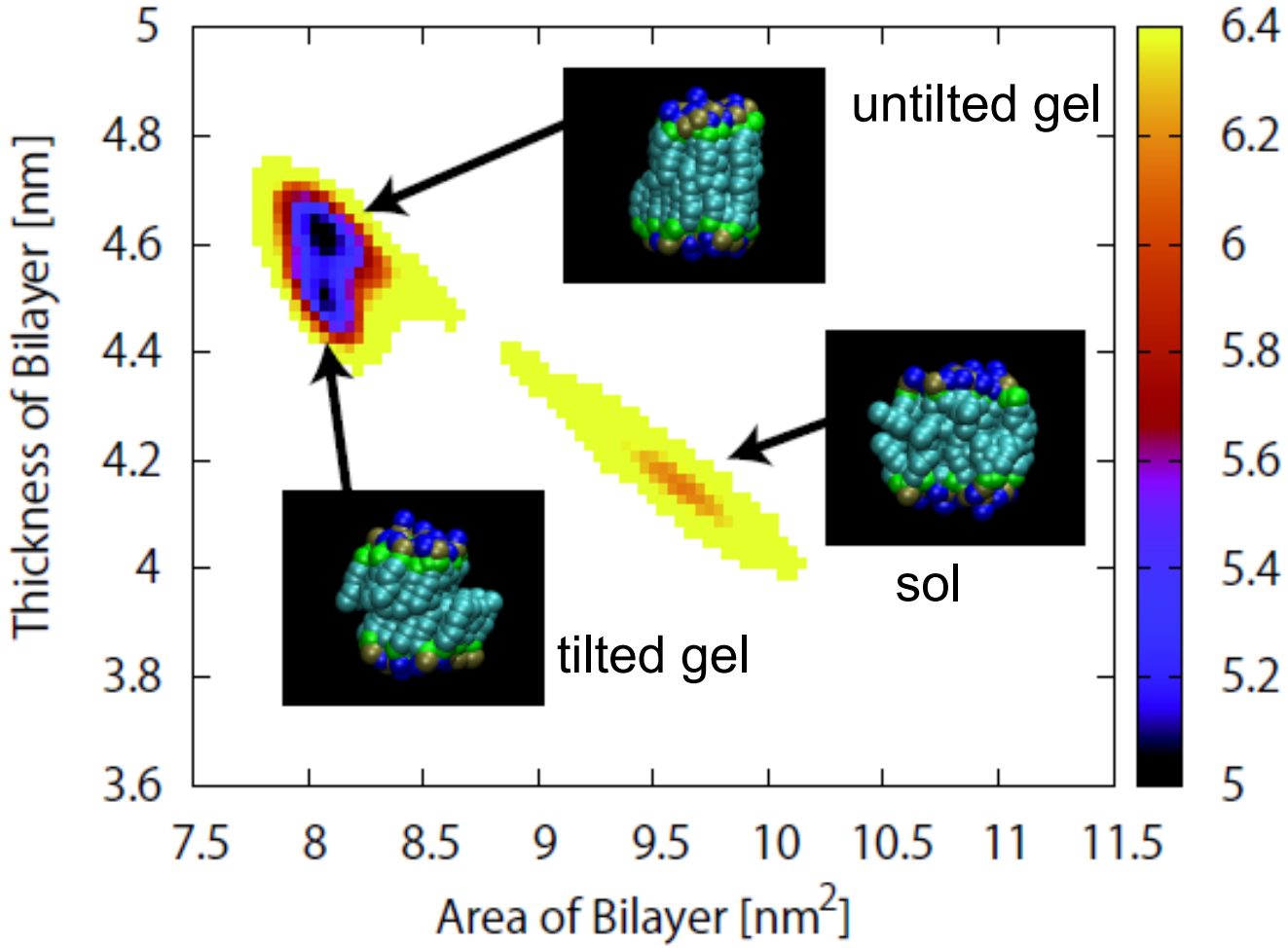
Simulation and movie by T. Nagai

Results

T. Nagai, R. Ueoka & Y.O., *J. Phys. Soc. Jpn.* **81**, 024002 (2012).

Potential of Mean Force

Three states are found.

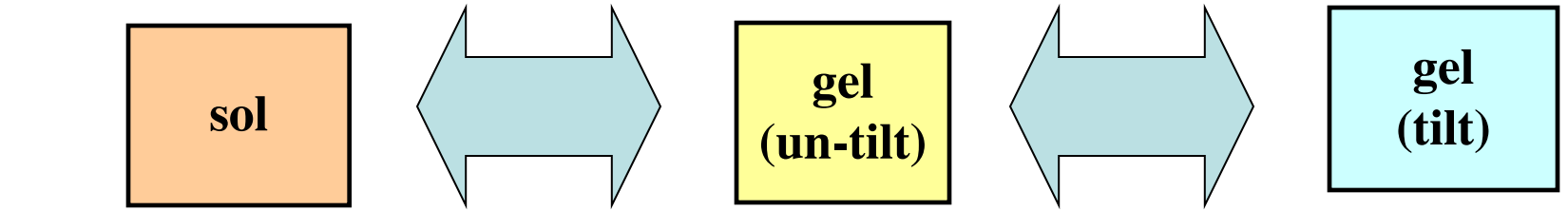


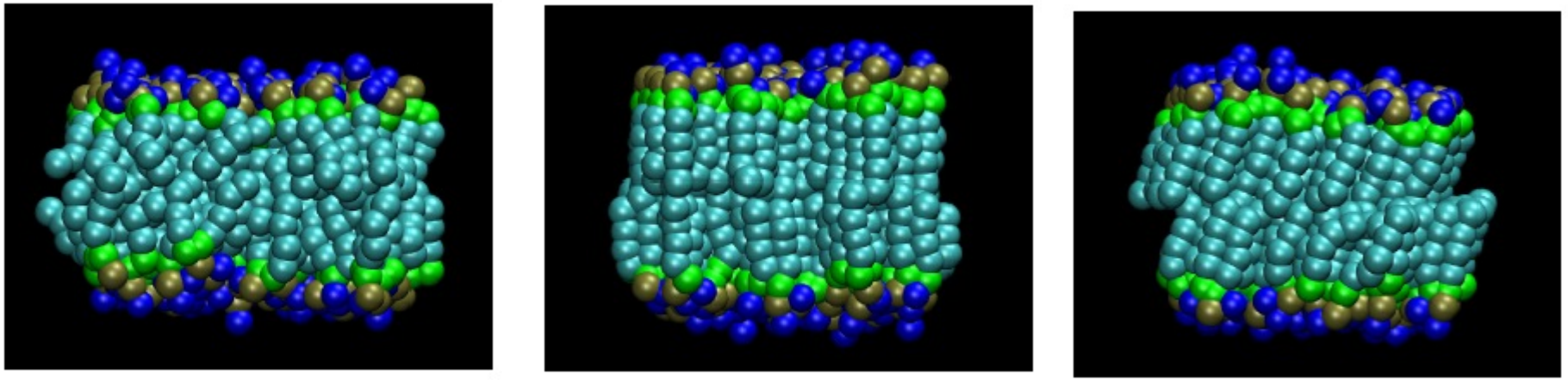
295.76 K

T. Nagai, R. Ueoka & Y.O., *J. Phys. Soc. Jpn.* **81**, 024002 (2012).

Results

Case for a bigger system: 128 lipids





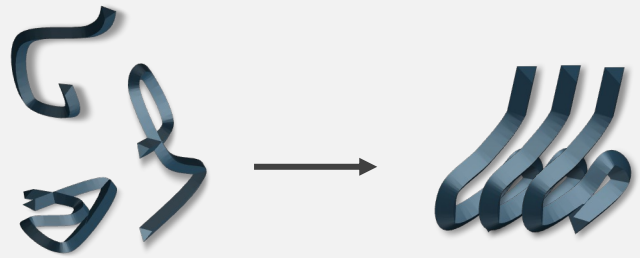
Structural Characteristics of Monomeric Aβ42 on Fibril in the Early Stage of Secondary Nucleation Mechanism

K. Noda, Y. Tachi, & Y.O.,

ACS Chemical Neuroscience **11**, 2989-2998 (2020).

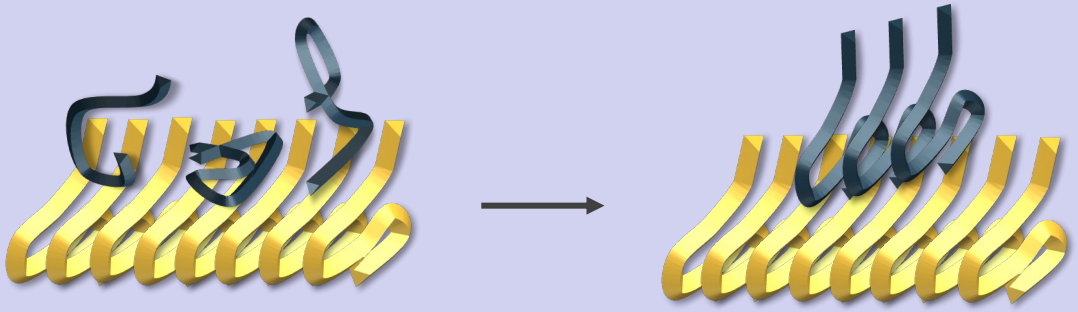
Path for amyloid β fibril formation

A Primary nucleation



B Elongation

C Secondary nucleation



[1] M. Törnquist et al., *Chem. Commun.* **54**, 8667 (2018).
[2] S. I. A. Cohen et al., *PNAS* **110**, 9758 (2013).

3. 方法

REST2 (Replica Exchange with Solute Tempering 2)

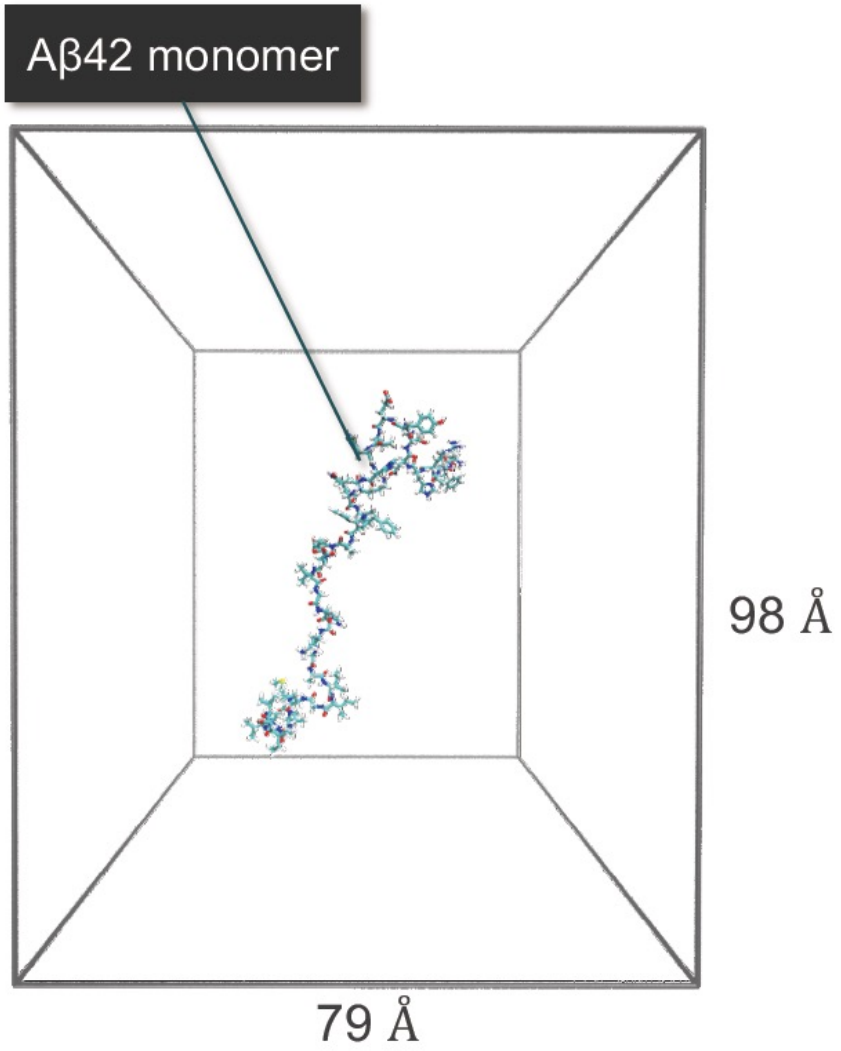
REST2:

$$E_m^{\mathrm{REST2}}(X) = \frac{\beta_m}{\beta_0} E_{\mathrm{ss}}(X) + \sqrt{\frac{\beta_m}{\beta_0}} E_{\mathrm{sw}}(X) + E_{\mathrm{ww}}(X)$$

[1] P. Liu et al., *PNAS* **102**, 13749 (2005).
[2] T. Terakawa, T. Kameda, and S. Takada, *J. Comput. Chem.* **32**, 1228 (2011).
[3] L. Wang et al., *J. Phys. Chem.* B **115**, 9431 (2011).
[4] S. Jo and W. Jiang, *Comp. Phys. Commun.* **197**, 304 (2015).

Simulation 1 | Aβ 42 monomer in solution

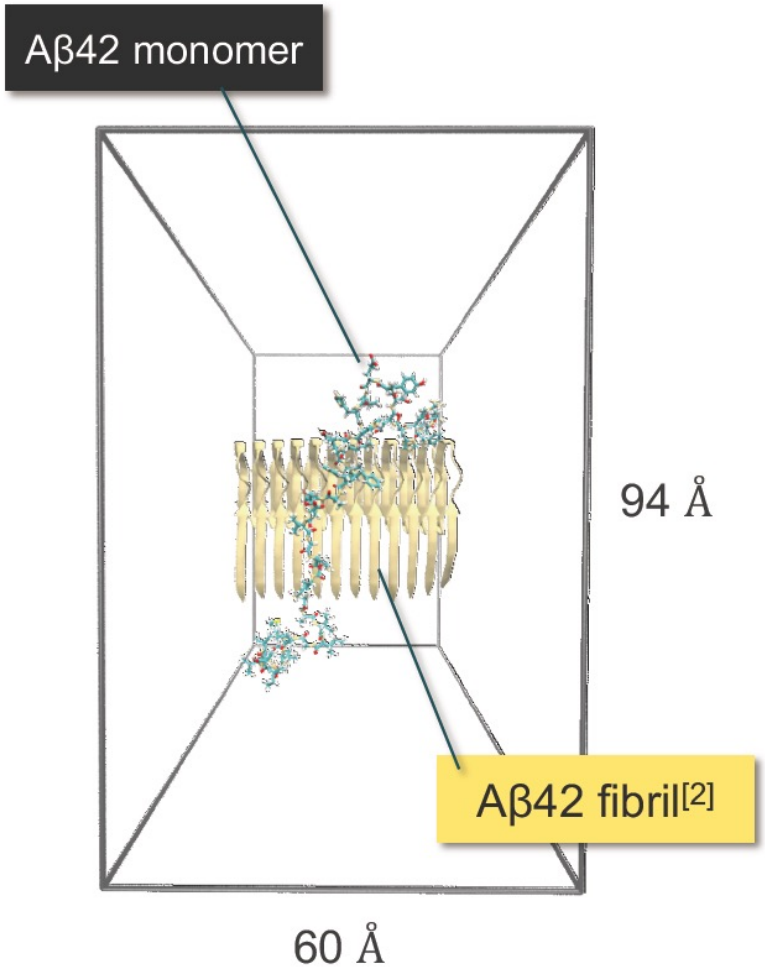
Program package	NAMD
Force field	CHARMM36m
Number of atoms	60,890
Ensemble	NPT (1 atm)
Solvent	TIP3P[1] (20,074)
Solute for REST2	monomer Aβ42
Solvent for REST2	Water
Thermostat	Langevin
Temperature (T_0)	310 K
Temperature for REST2 (T_m)	310 – 570 K
Time step	2 fs
The number of replicas	12 replicas
Simulation time	550 ns/replica
Restraint	SHAKE
Barostat	Nosé-Hoover Langevin piston
Box Size	79×98×79 Å^3



[1] W. L. Jorgensen et al., *J. Chem. Phys.* **79**, 926 (1983).

Simulation 2 | Aβ 42 fibril and monomer in solution

Program package	NAMD
Force field	CHARMM36m
Number of atoms	59,605
Ensemble	NVT
Solvent	TIP3P[1] (17,711)
Solute for REST2	monomer Aβ42
Solvent for REST2	Water + fibril (=12 Aβ42)
Thermostat	Langevin
Temperature (T_0)	310 K
Temperature for REST2 (T_m)	310 – 510 K
Time step	2 fs
The number of replicas	12 replicas
Simulation time	550 ns/replica
Restraint	SHAKE
Constraint	Cα atoms in fibril
Fibril structure (PDB)	2MXU [2]
Box Size	60×94×104 Å^3

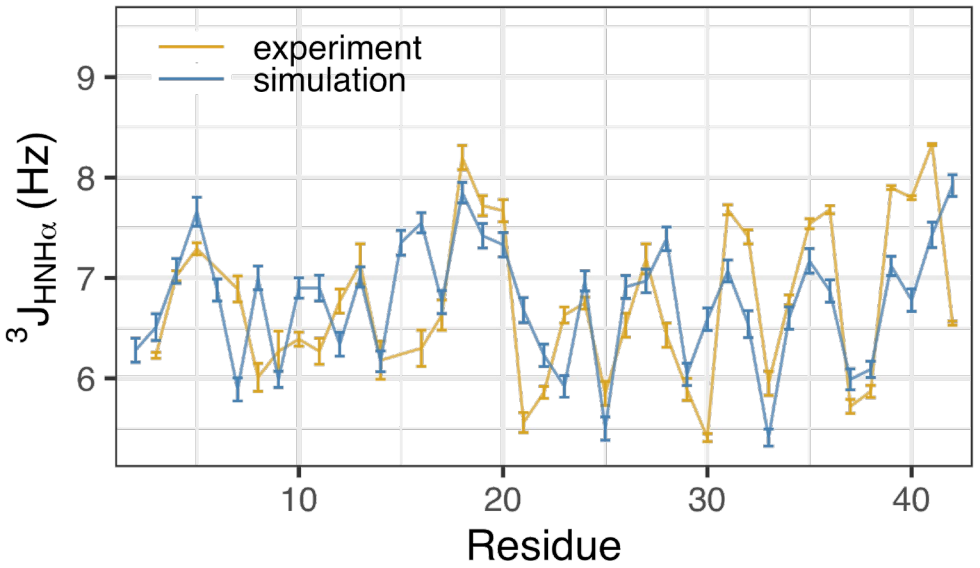


[1] W. L. Jorgensen et al., *J. Chem. Phys.* **79**, 926 (1983).
[2] Y. Xiao et al., *Nat. Struct. Mol. Biol.* **22**, 499 (2015).

Test of simulations

J coupling constant

$$^3J_{\mathrm{HNH}\alpha} = A\cos^2(\phi - 60) + B\cos(\phi - 60) + C$$



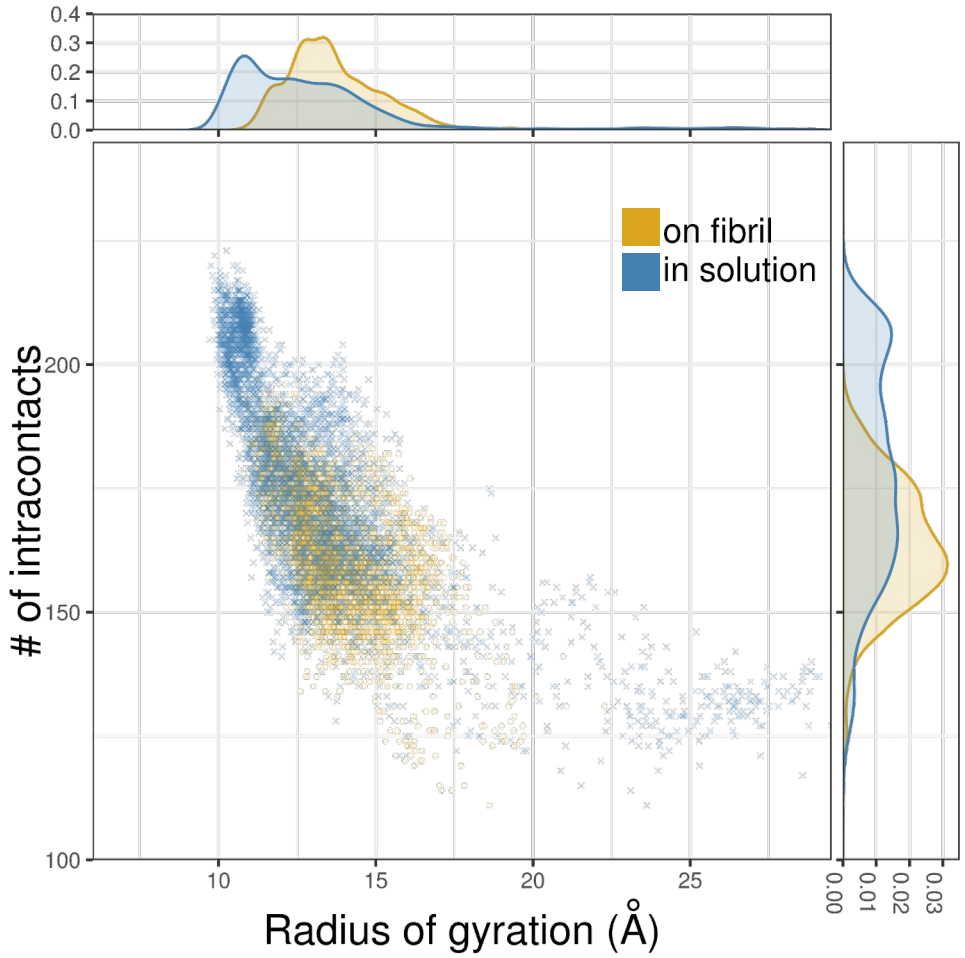
RMSD from experiment

This study	0.66 Hz
Sugourakis et al. [2]	0.73 Hz
Song et al. [3]	0.82 Hz

J-coupling is in good agreement

[1] J. Roche, et al., *Biochemistry* **55**, 762 (2016).
[2] N. G. Sugourakis, et al., *J. Mol. Bio.* **405**, 570 (2011).
[3] W. Song, et al., *Sci. Rep.* **5**, 11024 (2015).

Radius of gyration and number of contacts



Radius of Gyration

In solution < on fibril

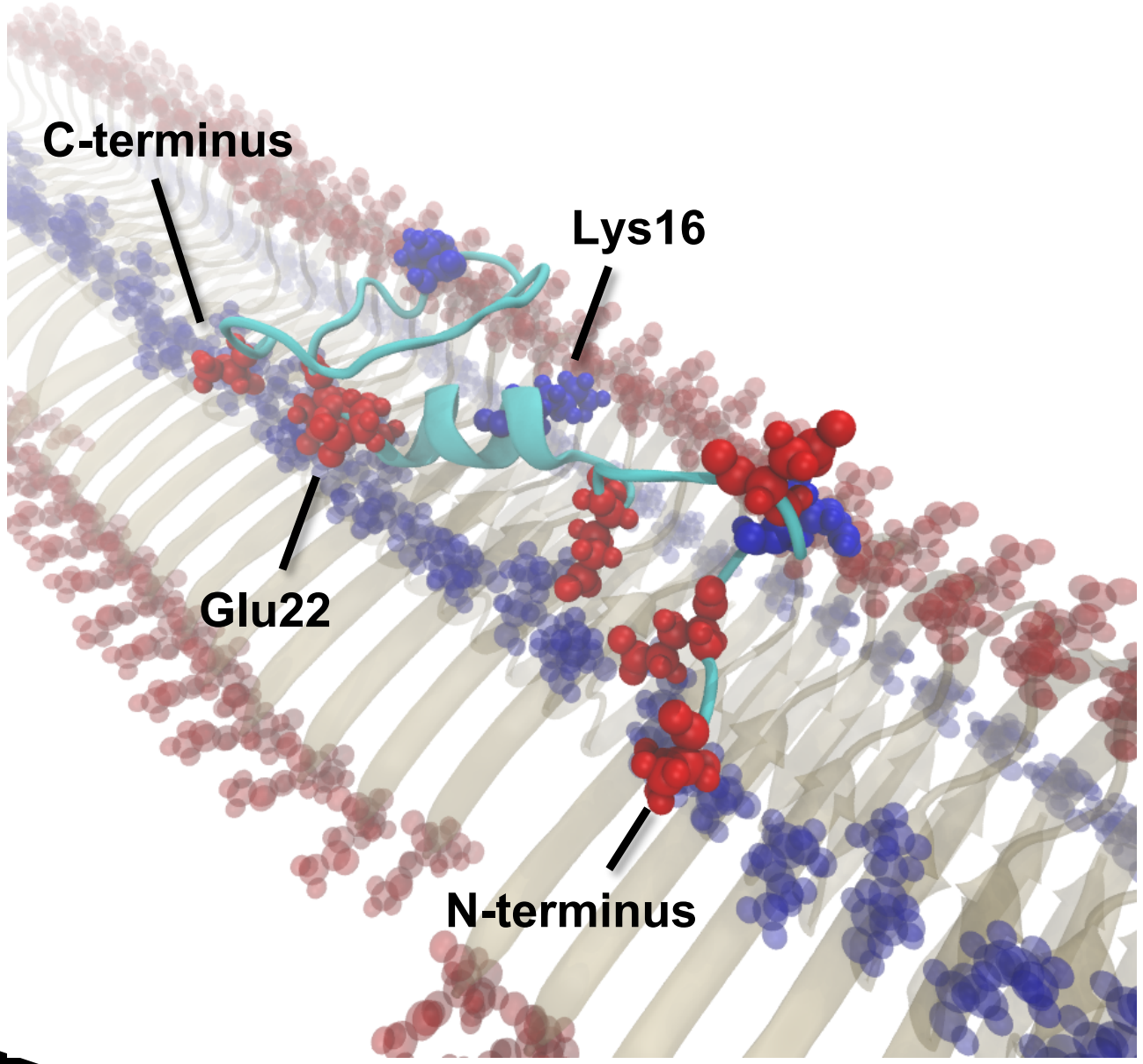
Intracontacts

In solution > on fibril

▼

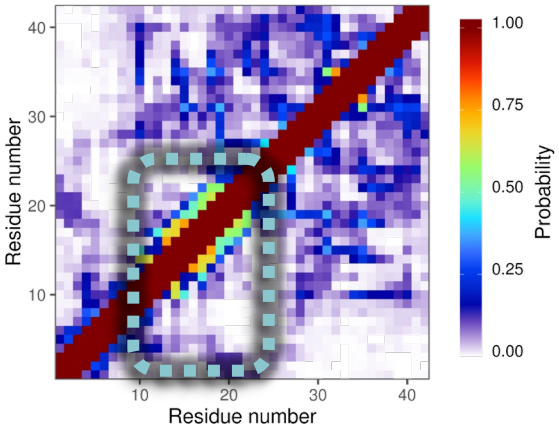
Unfolded on fibril

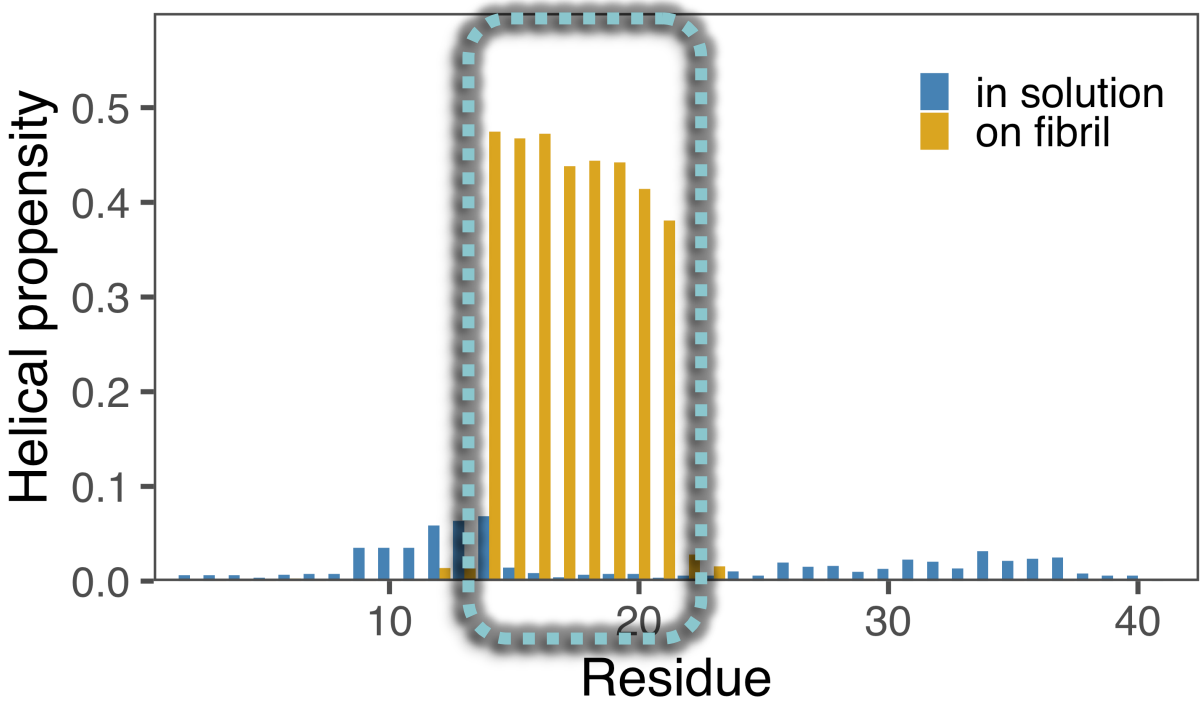
線維上でのヘリックス構造



Helix formation

Monomer on fibril





Helix formation : in solution ＜ on fibril

Development of New Generalized-Ensemble Algorithms

REVIEWS: A. Mitsutake, Y. Sugita, & Y.O., *Biopolymers* **60**, 96 (2001);
Y.O., *J. Mol. Graphics Modell.* **22**, 425 (2004);
Y. Sugita, A. Mitsutake, & Y.O., in *Lecture Notes in Physics*,
W. Janke (ed.) (Springer-Verlag, 2008) pp. 369-407.

Combination of Replica-Exchange Method (REM) and
Multicanonical Algorithm (MUCA)

1. **Replica-Exchange Multicanonical Algorithm (REMUCA)**

Y. Sugita & Y.O., *Chem. Phys. Lett.* 329, 261 (2000).

A. Mitsutake, Y. Sugita, & Y.O., *J. Chem. Phys.* 118, 6664; 6676 (2003).

- Multicanonical weight factor is determined from a short REM simulation by multiple histogram techniques

2. **Multicanonical Replica-Exchange Method (MUCAREM)**

Y. Sugita & Y.O., *Chem. Phys. Lett.* 329, 261 (2000).

A. Mitsutake, Y. Sugita, & Y.O., *J. Chem. Phys.* 118, 6664; 6676 (2003).

- Multicanonical simulations are performed for each replica and a pair of replicas are exchanged

Development of New Generalized-Ensemble Algorithms

REVIEWS: A. Mitsutake, Y. Sugita, & Y.O., *Biopolymers* **60**, 96 (2001);
Y.O., *J. Mol. Graphics Modell.* **22**, 425 (2004);
Y. Sugita, A. Mitsutake, & Y.O., in *Lecture Notes in Physics*,
W. Janke (ed.) (Springer-Verlag, 2008) pp. 369-407.

Combination of Replica-Exchange Method (REM) and
Simulated Tempering (ST)

3. **Replica-Exchange Simulated Tempering (REST)**

A. Mitsutake & Y.O., *Chem. Phys. Lett.* 332, 131 (2000).

- Simulated tempering weight factor is determined from a short REM simulation by multiple histogram techniques

4. **Simulated Tempering Replica-Exchange Method (STREM)**

A. Mitsutake & Y.O., *J. Chem. Phys.* 121, 2491 (2004).

- Simulated tempering simulations are performed for each replica and a pair of replicas are exchanged

See also:

M.K. Fenwick & F.A. Escobedo, *J. Chem. Phys.* 119, 11998 (2003).

KEY ELEMENT: Multiple-Histogram Reweighting Techniques (WHAM)

A. Ferrenberg & R. Swendsen, *Phys. Rev. Lett.* **63**, 1195 (1989); S. Kumar, D. Bouzida, R. Swendsen, P. Kollman & J. Rosenberg, *J. Comput. Chem.* **13**, 1011 (1992).

Given *M* set of histograms $N_m(E)$, the following WHAM equations are solved iteratively for density of states $n(E)$ and dimensionless Helmholtz free energy f_m:

$$n(E) = \frac{\sum_{m=1}^{M} N_m(E)}{\sum_{m=1}^{M} n_m e^{f_m - \beta_m E}}, \text{ where } e^{-f_m} = \sum_E n(E) e^{-\beta_m E}.$$

Estimating Density of States of Complex Systems by REWL-MUCAREM Simulations

* T. Hayashi and Y.O.,
"An Efficient Simulation Protocol for Determining the Density of States: Combination of Replica-Exchange Wang-Landau Method and Multicanonical Replica-Exchange Method,"
Phys. Rev. E **100**, 043304 (2019).

* T. Hayashi, C. Muguruma, and Y.O.,
"Calculation of the Residual Entropy of Ice Ih by Monte Carlo Simulation with the Combination of the Replica-Exchange Wang-Landau Algorithm and Multicanonical Replica-Exchange Method,"
J. Chem. Phys. **154**, 044503 (11 pages) (2021).

MUCAREM and REWL

* MUCAREM (Multicanonical Replica-Exchange Method)
  Y. Sugita and Y.O., *Chem. Phys. Lett.* **329**, 261 (2000).
  A. Mitsutake, Y. Sugita, and Y.O.,
  J. Chem. Phys. **118**, 6664 (2003).
  A. Mitsutake, Y. Sugita, and Y.O.,
  J. Chem. Phys. **118**, 6676 (2003).

  In MUCAREM, we use the multiple-histogram reweighting techniques extensively.

* REWL (Replica-Exchange Wang-Landau)
  T. Vogel, Y. W. Li, T. W¨ust, and D. P. Landau,
  Phys. Rev. Lett. **110**, 210603 (2013).
  T. Vogel, Y. W. Li, T. W¨ust, and D. P. Landau,
  Phys. Rev. E **90**, 023302 (2014).

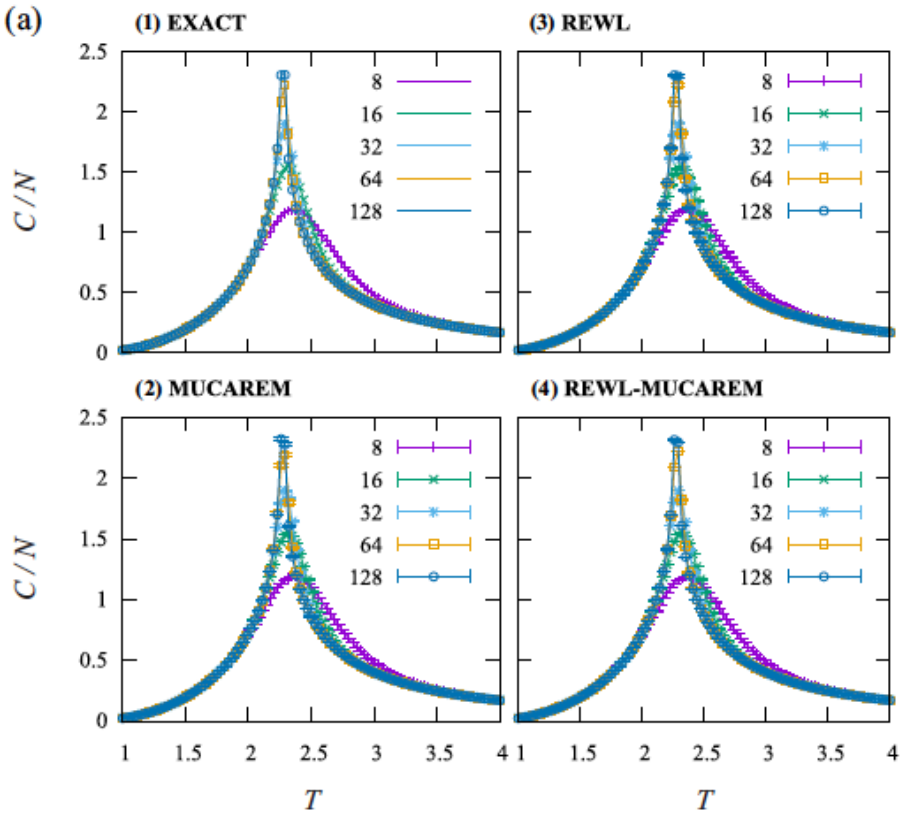
REWL-MUCAREM protocol:

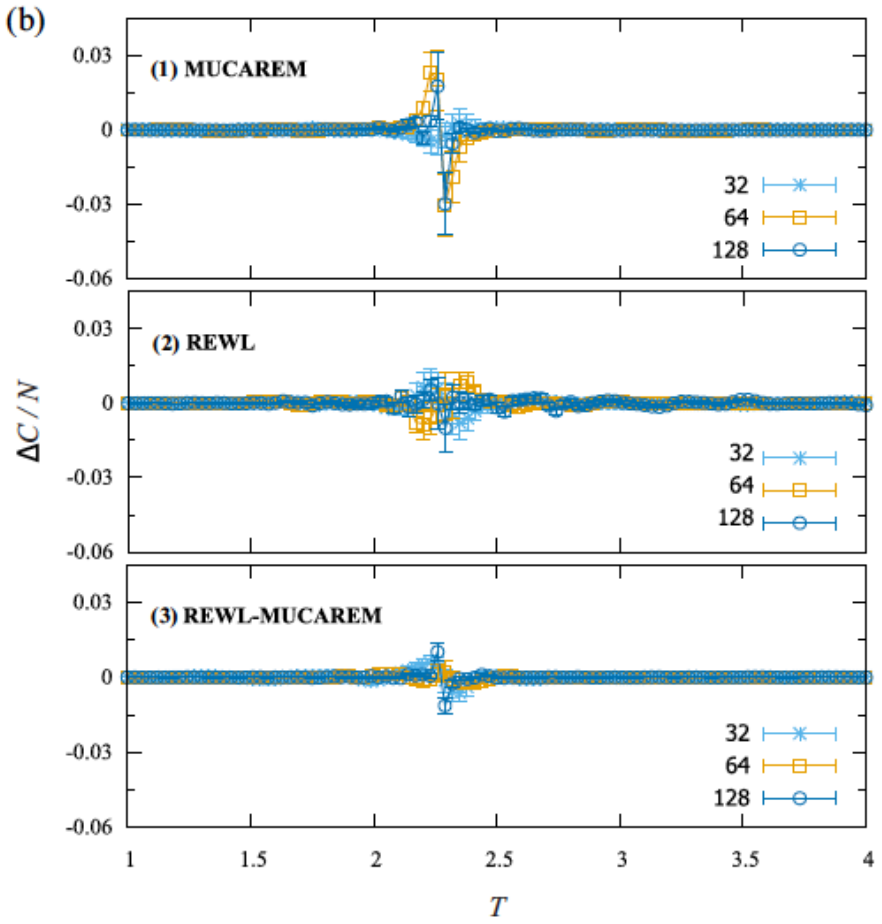
T. Hayashi and Y.O., *Phys. Rev. E* **100**, 043304 (2019).

We first perform REWL, then MUCAREM.
We tested the performances of MUCAREM only, REWL only, and REWL-MUCAREM with the 2-dimensional Ising Model.

REWL-MUCAREM protocol:

T. Hayashi and Y.O., *Phys. Rev. E* **100**, 043304 (2019).





Exact Solutions:

P.D. Beale, *Phys. Rev. Lett.* **76**, 78 (1996);

A.E. Ferdinand and M.E. Fisher, *Phys. Rev.* **185**, 832 (1969).

Calculation of residual entropy of ice by multicanonical simulations

B.A. Berg, C. Muguruma & Y.O., *Phys. Rev. B* **75**, 092202 (2007);
B.A. Berg & W. Yang, *J. Chem. Phys.* **127**, 224502 (2007);
C. Muguruma, Y.O. & B.A. Berg, *Phys. Rev. E* **78**, 041113 (2008);
B.A. Berg, C. Muguruma & Y.O., *Mol. Sim.* **38**, 856 (2012);
T. Hayashi, C. Muguruma & Y.O., *J. Chem. Phys.* **154**, 044503 (2021).

$S_0^{\text{experiment}} = 0.82(5)$ cal/deg/mole

W.F. Giauque and M. Ashley, *Phys. Rev.* **43** 81 (1933).

$S_0^{\text{experiment}} = 0.815(26)$ cal/deg/mole

O. Haida, T. Matsuo, H. Suga, and S. Seki, *J. Chem. Thermodynamics* **6**, 815 (1977).

$S_0 = k \ln(W)$ where $W = (W_0)^N$

$S_0^{\text{Pauling}} = 0.80574...$ cal/deg/mole $\qquad W_0^{\text{Pauling}} = \frac{3}{2}$

L. Pauling, *J. Am. Chem. Soc.* **57** 2680 (1935).

Onsager pointed out this is just a lower bound.

L. Onsager and M. Dupuis, *J. Rend. Sc. Inst. Fis. "Enrico Fermi"* **10**, 294 (1960).

Onsager's student Nagle improved by series expansion:

$S_0^{\text{Nagle}} = 0.81480(20)$ cal/deg/mole $\qquad W_0^{\text{Nagle}} = 1.50685(15)$

J.F. Nagle, *J. Mat. Phys.* **7**, 1484 (1966).

$S_0^{\text{MUCA}} = 0.81550(21)$ cal/deg/mole $\qquad W_0^{\text{MUCA}} = 1.50738(16)$

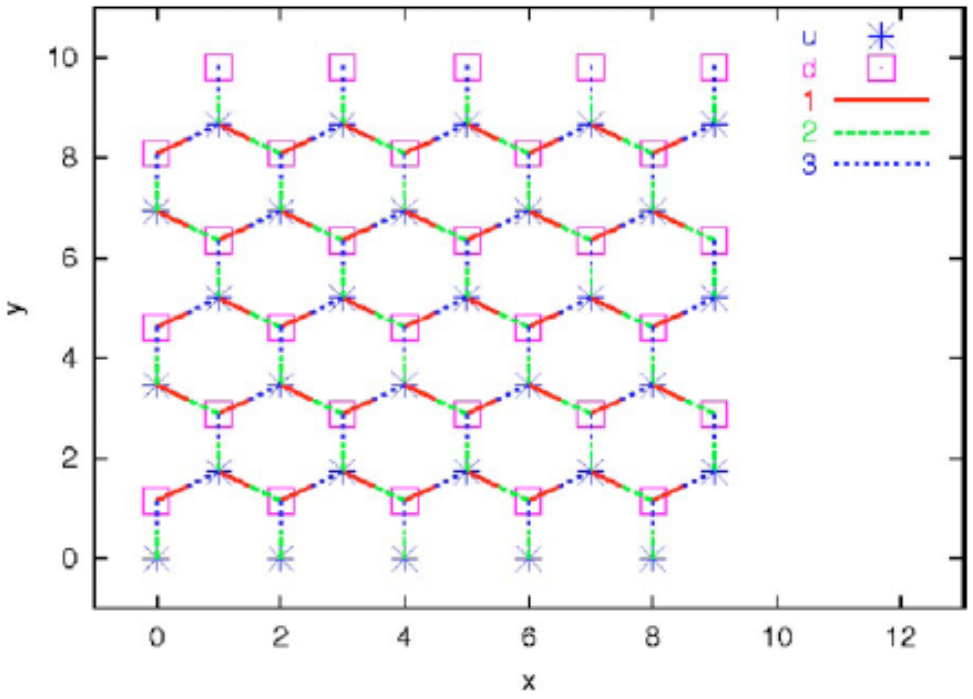
B.A. Berg, C. Muguruma & Y.O., *Phys. Rev. B* **75**, 092202 (2007).

$S_0^{\text{MUCA}} = 0.81527(17)$ cal/deg/mole $\qquad W_0^{\text{MUCA}} = 1.50721(13)$

C. Muguruma, Y.O. & B.A. Berg, *Phys. Rev. E* **78**, 041113 (2008).

Ice Rules

1. There is one hydrogen atom on each bond
   (then called hydrogen bond)
2. There are two hydrogen atoms near each oxygen atom
   (these constitute a water molecule)



Lattice structure of one layer of ice I

B.A. Berg, C. Muguruma & Y.O., *Phys. Rev. B* **75**, 092202 (2007).

We proposed the following two Hamiltonians in which only the groundstate satisfies both of the Ice Rules.

$$E = -\sum_b h(b, s_b^1, s_b^2)$$

6-state H_2O molecule model
[Ice Rule (2) is always enforced]

where $$h(b, s_b^1, s_b^2) = \begin{cases} 1, & \text{for H-bond} \\ 0, & \text{otherwise} \end{cases}$$

$$E = -\sum_s f(s, b_s^1, b_s^2, b_s^3, b_s^4)$$

2-state H-bond model
[Ice Rule (1) is always enforced]

where

$$f(s, b_s^1, b_s^2, b_s^3, b_s^4) = \begin{cases} 2, & \text{for 2 H nuclei close to } s \\ 1, & \text{for 1 or 3 H nuclei close to } s \\ 0, & \text{for 0 or 4 H nuclei close to } s \end{cases}$$

Single-Histogram Reweighting Techniques

A. Ferrenberg & R. Swendsen, *Phys. Rev. Lett.* **61**, 2635 (1988).

The number of states *n(E)* is obtained from the histogram of the energy distribution $N_{mu}(E)$ that was obtained from the production run of the multicanonical simulation:

$$n(E) = \frac{N_{mu}(E)}{W_{mu}(E)}, \quad \text{where} \quad N_{mu}(E) = n(E) W_{mu}(E) \ .$$

Then the residual entropy is obtained by

$$S_0 = k \ln(n(E_0)), \quad \text{where } E_0 \text{ is the energy of the groundstate}$$

where the normalization for *n(E)* is set by

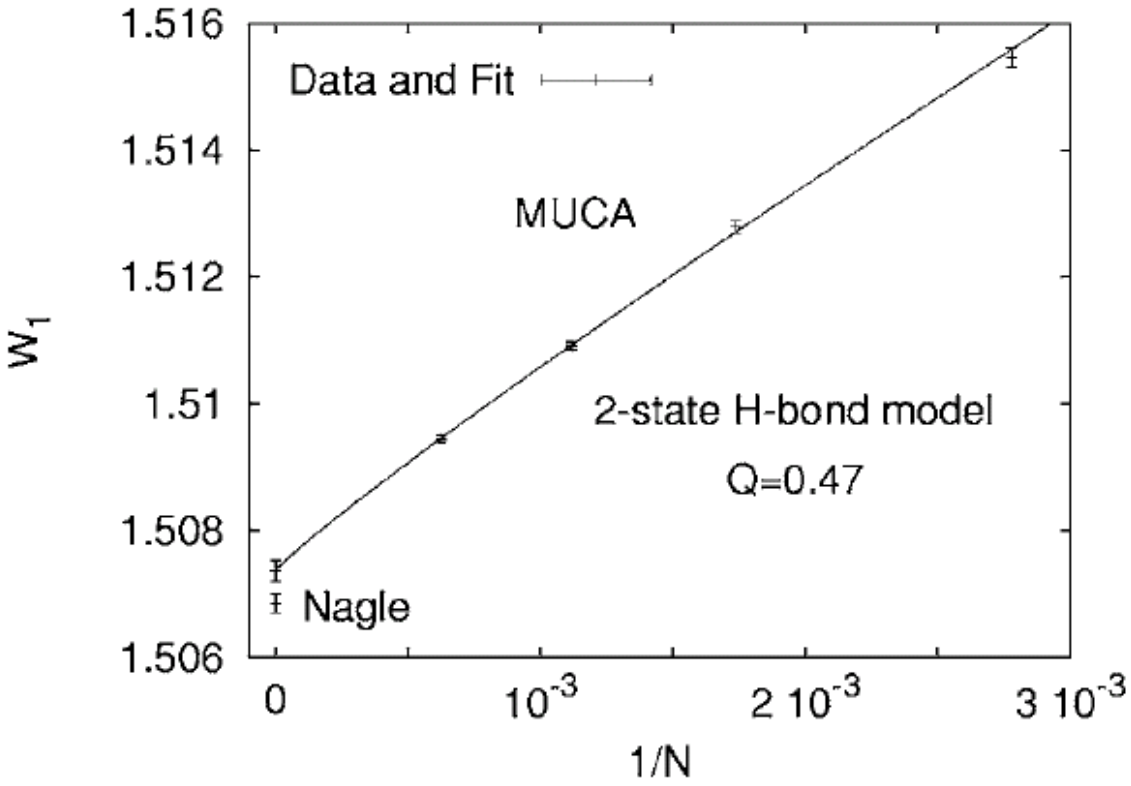
$$\sum_E n(E) = \begin{cases} 6^N & \text{for the 6-state model} \\ 2^{2N} & \text{for the 2-state model} \end{cases} \qquad \text{at } \beta = 0$$

Cf. B.A. Berg and T. Celik, *Phys. Rev. Lett.* **69**, 2292 (1992).

B.A. Berg, C. Muguruma & Y.O., *Phys. Rev. B* **75**, 092202 (2007).

$$S_0 = k \ln(W) \qquad \text{where} \qquad W = (W_0)^N$$

$$S_0^{\mathrm{MUCA}} = 0.81550(21) \quad \text{cal/deg/mole} \qquad W_0^{\mathrm{MUCA}} = 1.50738(16)$$



SUMMARY up to 2012

B.A. Berg, C. Muguruma & Y.O., *Mol. Sim.* **38**, 856 (2012).

We have improved the accuracy of residual entropy of ordinary ice to be

$$S_0 = k \ln(W) \qquad \text{where} \qquad W = (W_0)^N$$

(1935) $W_0^{\text{Pauling}} = \frac{3}{2}$ $\qquad S_0^{\text{Pauling}} = 0.80574...$

(1966) $W_0^{\text{Nagle}} = 1.50685(15)$ $\qquad S_0^{\text{Nagle}} = 0.81480(20)$

(2008) $W_0^{\text{MUCA}} = 1.50721(13)$ $\qquad S_0^{\text{MUCA}} = 0.81527(17)$

But the accuracy of experiment is just one digit.

$$S_0^{\text{experiment}} = 0.815(26)$$

O. Haida, T. Matsuo, H. Suga, and S. Seki, *J. Chem. Thermodynamics* **6**, 815 (1977).

We would like the calorimetry experimentalists to improve the values.

New Calculations

In 2012, we used up to $N = 2880$ water molecules.

For more water molecules, Wang-Landau did not work any more.

B.A. Berg, C. Muguruma & Y.O., *Mol. Sim.* **38**, 856 (2012).

Our new method, REWL-MUCAREM (combination of Replica-Exchange Wang-Landau method

T. Vogel, Y.W. Li, T. Wust & D.P. Landau, *Phys. Rev. Lett.* **110**, 21 (2013).

and Multicanonical Replica-Exchange Method)

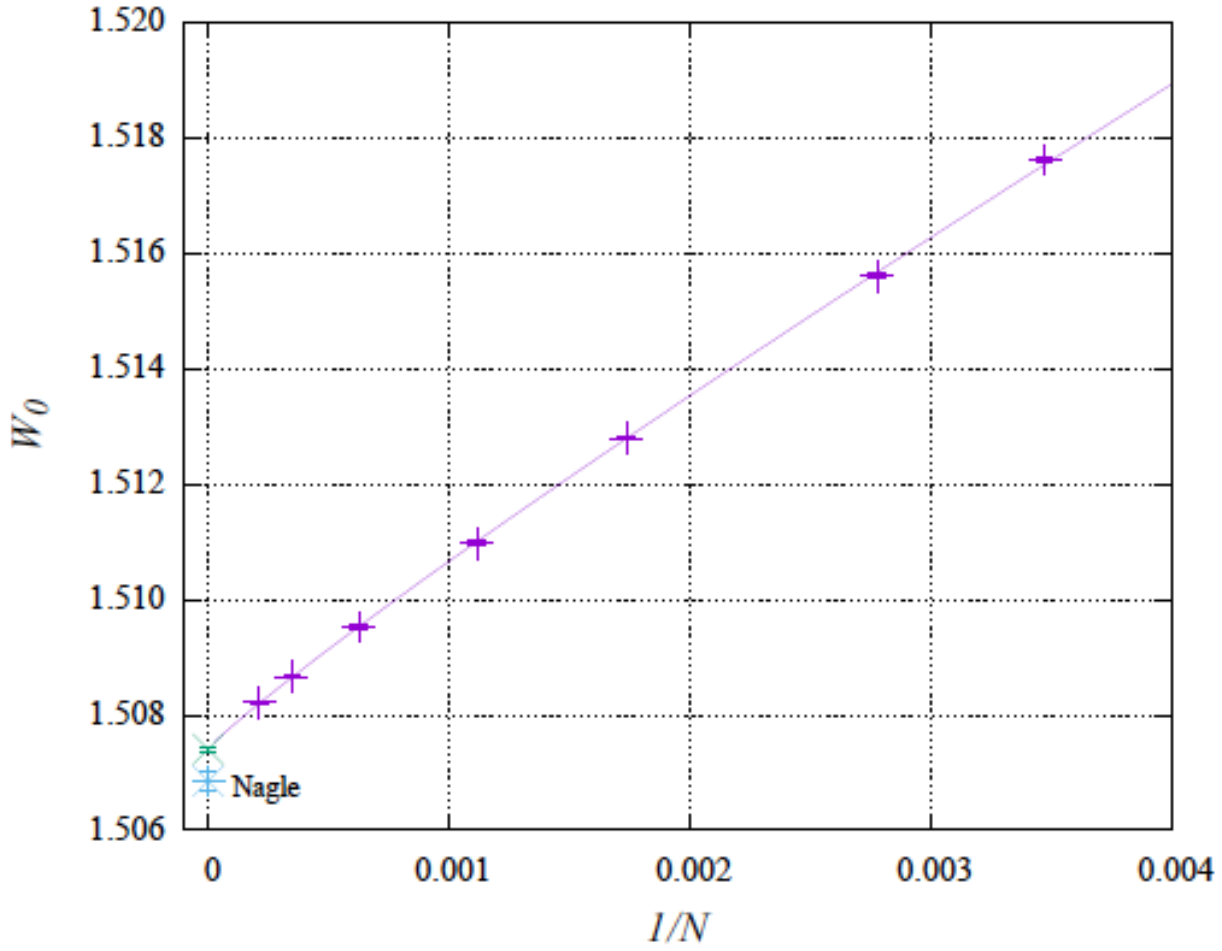
Y. Sugita & Y.O., *Chem. Phys. Lett.* **329**, 261 (2000).

worked for $N = 4704$.

T. Hayashi, C. Muguruma & Y.O., *J. Chem. Phys.* **154**, 044503 (2021).

New Calculations

T. Hayashi, C. Muguruma & Y.O., *J. Chem. Phys.* **154**, 044503 (2021).



New Calculations

T. Hayashi, C. Muguruma & Y.O., *J. Chem. Phys.* **154**, 044503 (2021).

Table 2: Estimated residual entropy of Ice Ih.

N	n_x	n_y	n_z	Tunneling[a]	Flatness[b]	$W_0{}^*$	$S_0{}^*$
128	4	8	4	7612228	0.98391	1.5286054(462)	0.8432816(601)
288	4	12	6	1598145	0.97176	1.5176118(362)	0.8289382(474)
360	5	12	6	1020866	0.97870	1.5156001(402)	0.8263023(527)
576	6	12	8	404617	0.97047	1.5127892(339)	0.8226133(446)
896	7	16	8	172052	0.95956	1.5109753(276)	0.8202291(363)
1600	8	20	10	57171	0.93313	1.5095170(284)	0.8183102(373)
2880	10	24	12	17417	0.90334	1.5086586(304)	0.8171799(401)
4704	12	28	14	6318	0.83998	1.5082141(319)	0.8165944(420)
∞		fitting				1.5074123(466)	0.8155376(614)

Latest Value

Comparisons with Other Calculations

T. Hayashi, C. Muguruma & Y.O., *J. Chem. Phys.* **154**, 044503 (2021).

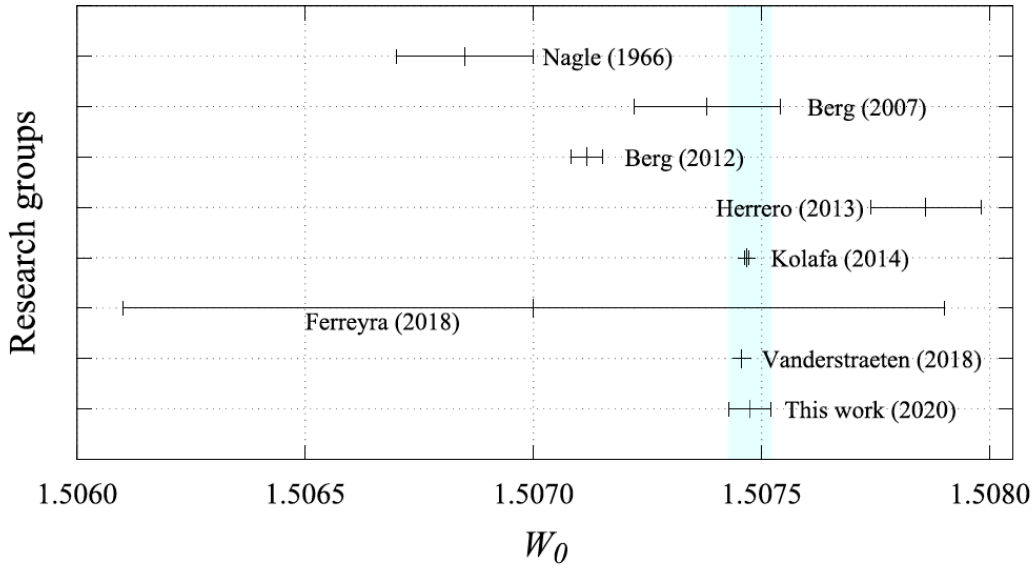


Table 3: Comparisons of the estimates by various methods.

Group	Methods	W_0	ΔW_0	S_0	ΔS_0
Nagle[5]	Series expansion	1.50685	0.00015	0.8147962	0.000198
Berg (2007)[6]	Multicanonical algorithm	1.50738	0.00016	0.81550	0.00021
Berg (2012)[9]	Multicanonical algorithm	1.507117	0.000035	0.815149	0.000046
Herrero[14]	Thermodynamic Integration	1.50786	0.00012	0.81613	0.00016
Kolafa[15]	Thermodynamic Integration	1.5074674	0.0000038	0.8156103	0.0000051
Ferreyra[17]	Wang-Landau algorithm	1.5070	0.0009	0.81478	0.00012
Vanderstraeten[18]	PEPS algorithm	1.507456		0.8155953	
This work	REWL-MUCAREM	1.507412	0.000047	0.815538	0.000062

Problem with Random Numbers:

T. Hayashi, C. Muguruma & Y.O., *J. Chem. Phys.* **154**, 044503 (2021).

Marsaglia

G. Marsaglia, A. Zaman, and W.W. Tsang, *Stat. Probab. Lett.* **9**, 35 (1990).

vs.

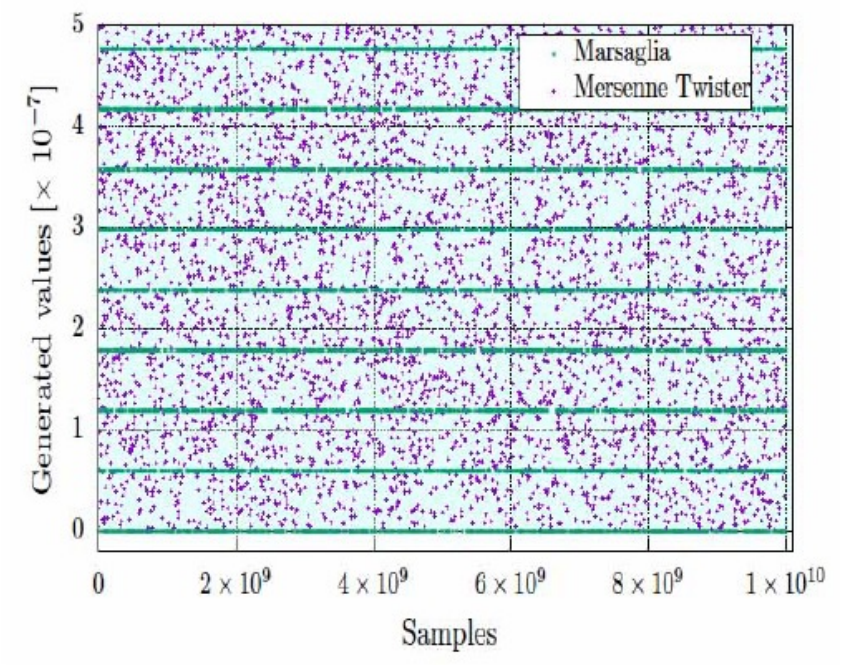
Mersenne Twister

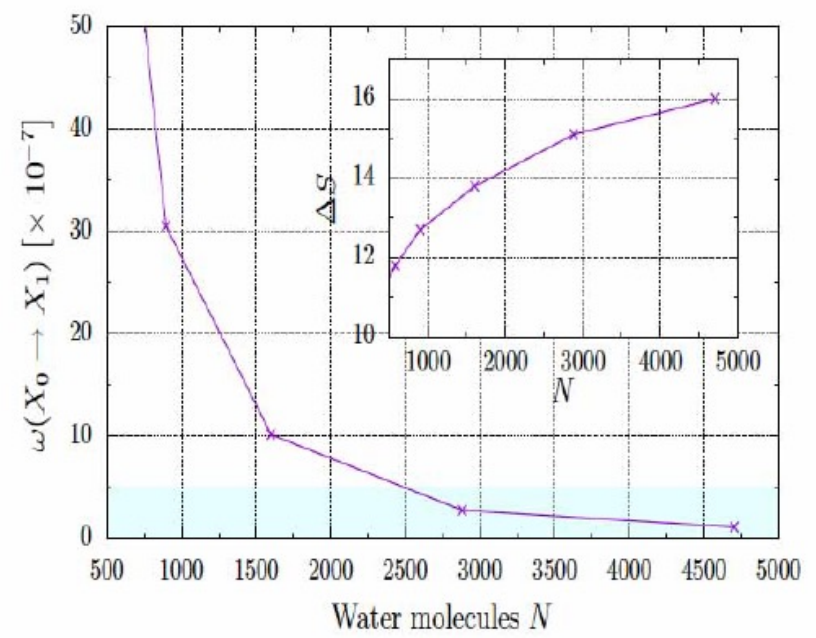
M. Matsumoto and T. Nishimura, *ACM Trans. Model. Comput. Simul.* **8**, 3 (1998).

Mersenne Twister is a better random number generator than Marsaglia.

Problem with Random Numbers:

T. Hayashi, C. Muguruma & Y.O., *J. Chem. Phys.* **154**, 044503 (2021).





In the range $(0 \sim 5) \times 10^{-7}$, Mersenne Twister gives a uniform distribution of random numbers, while Marsaglia gives only 9 values (green horizontal lines in the left figure).

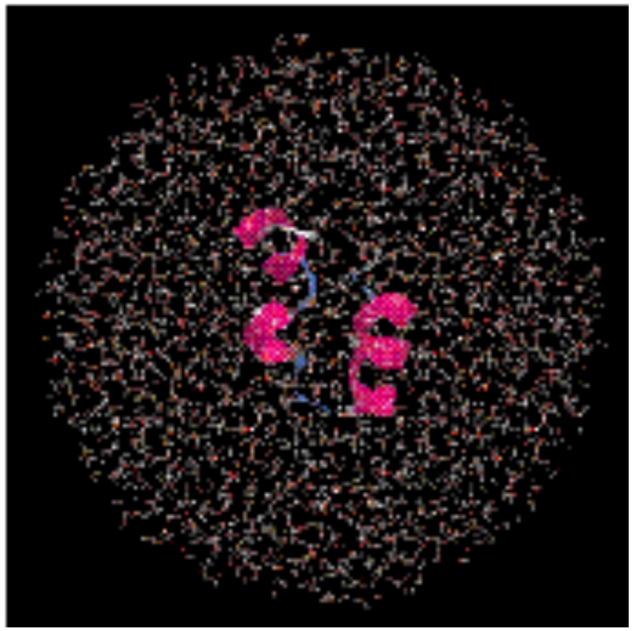
球状タンパク質の立体構造予測

T. Yoda, Y. Sugita & Y.O., *Biophys. J.* **99**, 1637 (2010).
T. Yoda, Y. Sugita & Y.O., *Proteins* **82**, 933-943 (2014).

Challenge the prediction of the 3-dimensional structure of a small protein by MUCAREM.

R = 30A, 3550wat, 11246atom

ヴィリン ヘッドピース サブドメイン
(36 残基; 596 原子)
半径 30 Åの水の球 (3513 個の水分子);
全原子数は 11,135 個



1 10 20 30
MLSDEDFKAVFGMTRSAFANLPLWKQQNLKKEKGLF
helix1 helix2 helix3

Computational Details

T. Yoda, Y. Sugita & Y.O., *Biophys. J.* **99**, 1637 (2010).

(Force Field = CHARMM22/CMAP for protein
& TIP3P for water)

(1) REMD with 96 replicas in implicit solvent (GB/SA); initial conformation was fully extended

(2) Unfolded protein w/o any secondary structures was soaked in a sphere of radius 30Å (with 3513 TIP3P water molecules)

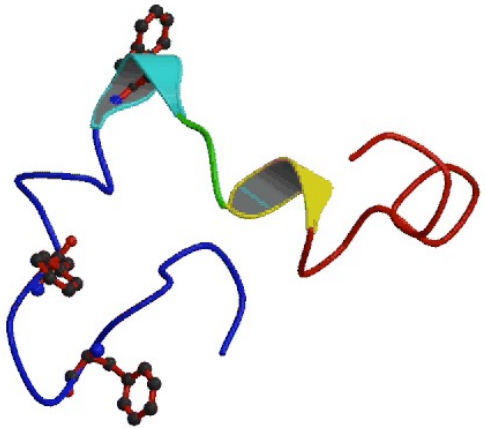
(3) REMD with 128 replicas (T = 250 K ~ 700 K)

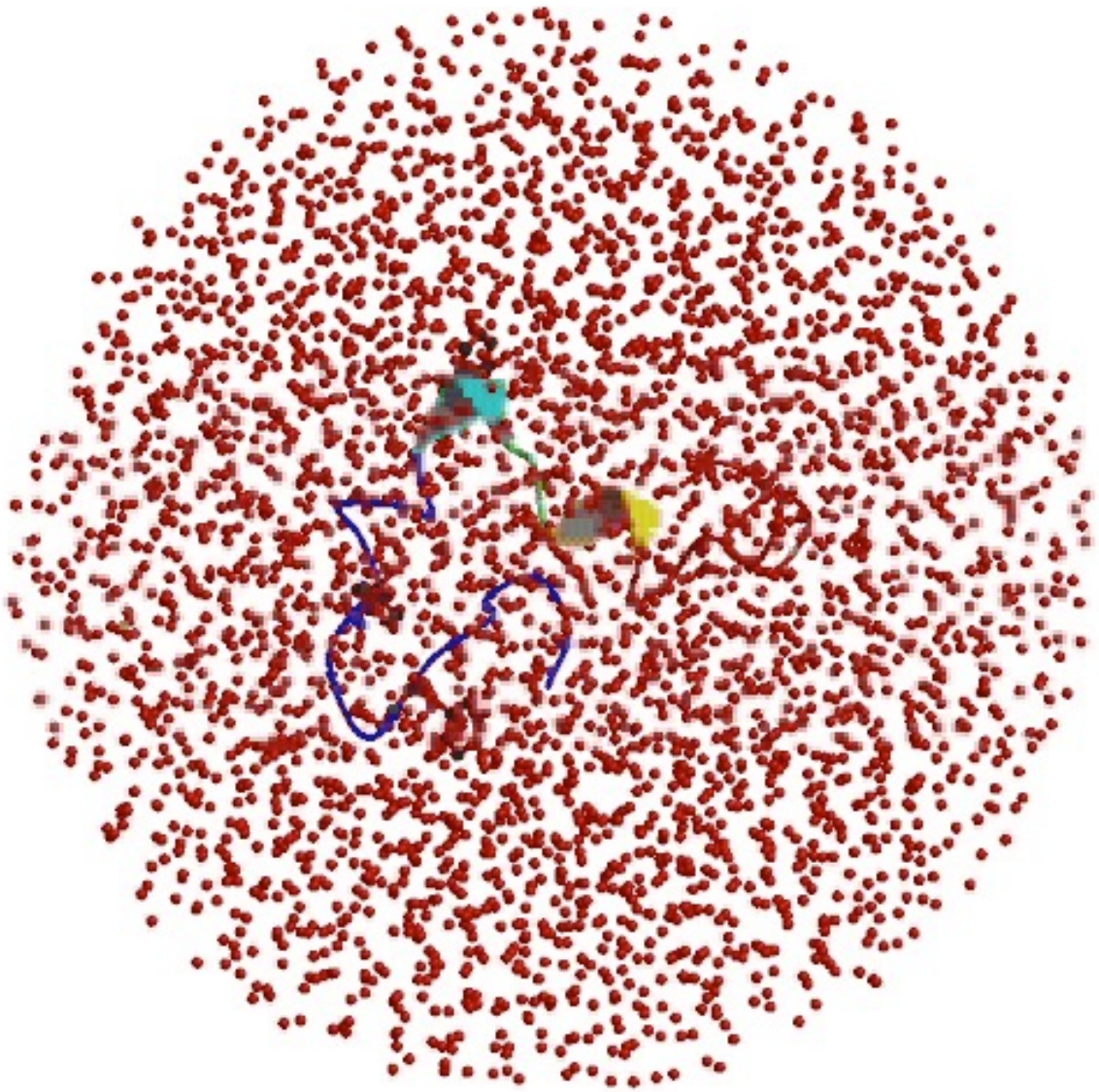
(4) Determine multicanonical weight factors by WHAM (iterate several times to refine weight)

(5) Two production runs of MUCAREM with 8 replicas

(MUCAREM1: 1.127 μs in total covering T = 269 K ~ 699 K
MUCAREM2: 1.157 μs in total covering T = 289 K ~ 699 K)

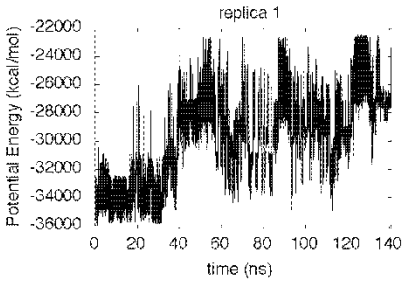
Villin headpiece subdomain
(36 amino acids; 596 atoms)
in sphere of water of radiu
(3513 water molecules);
altogether 11,135 atoms

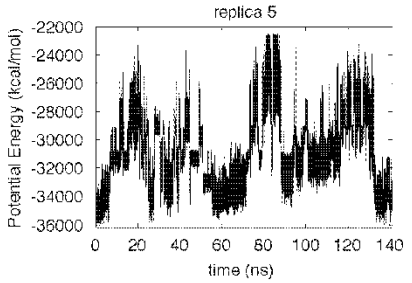


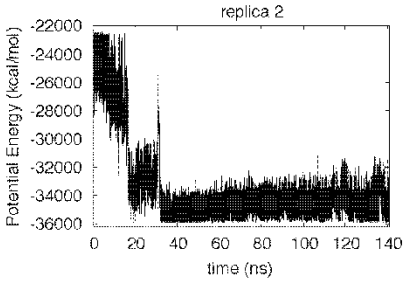


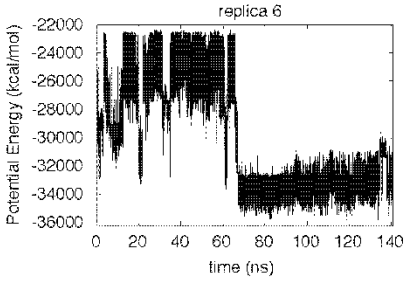
MUCAREM simulation

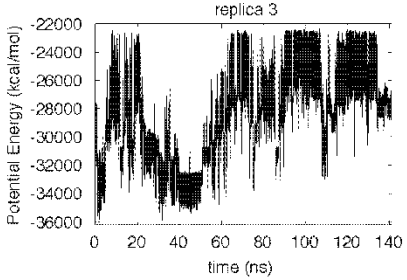
Potential Energy Time Series

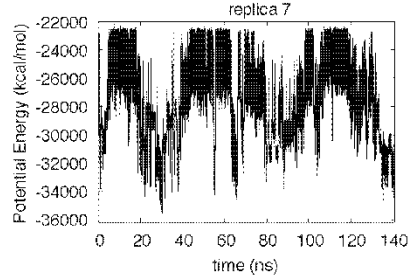


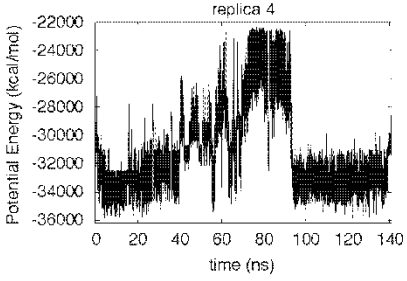


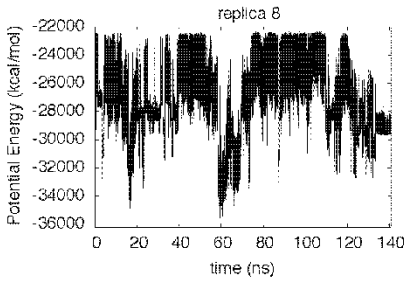


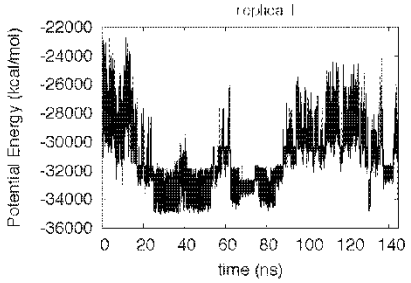


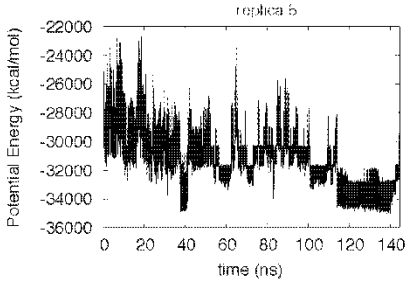


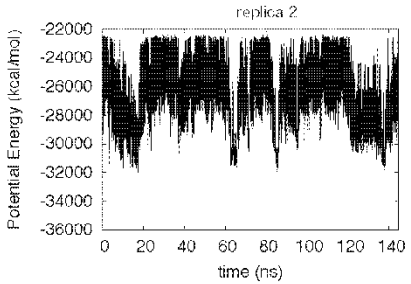


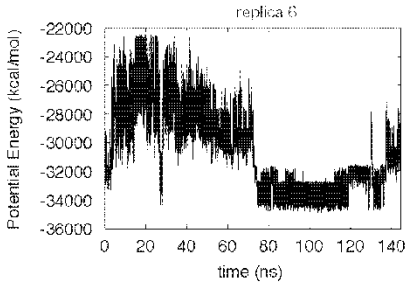


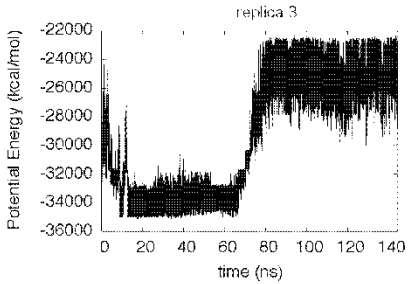


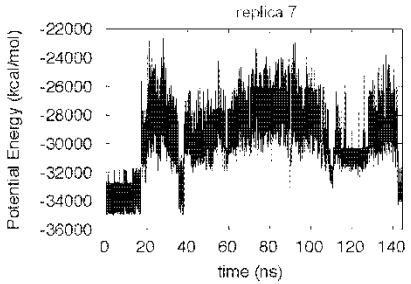


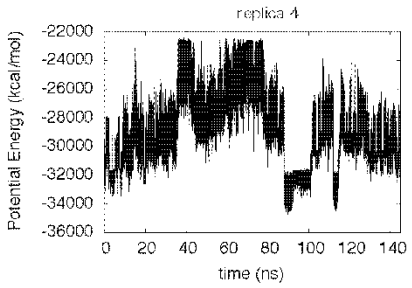


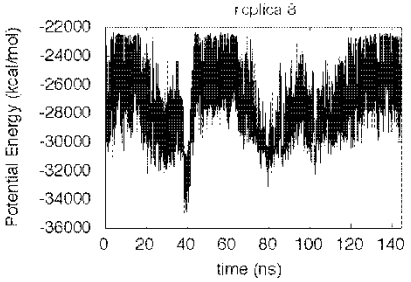




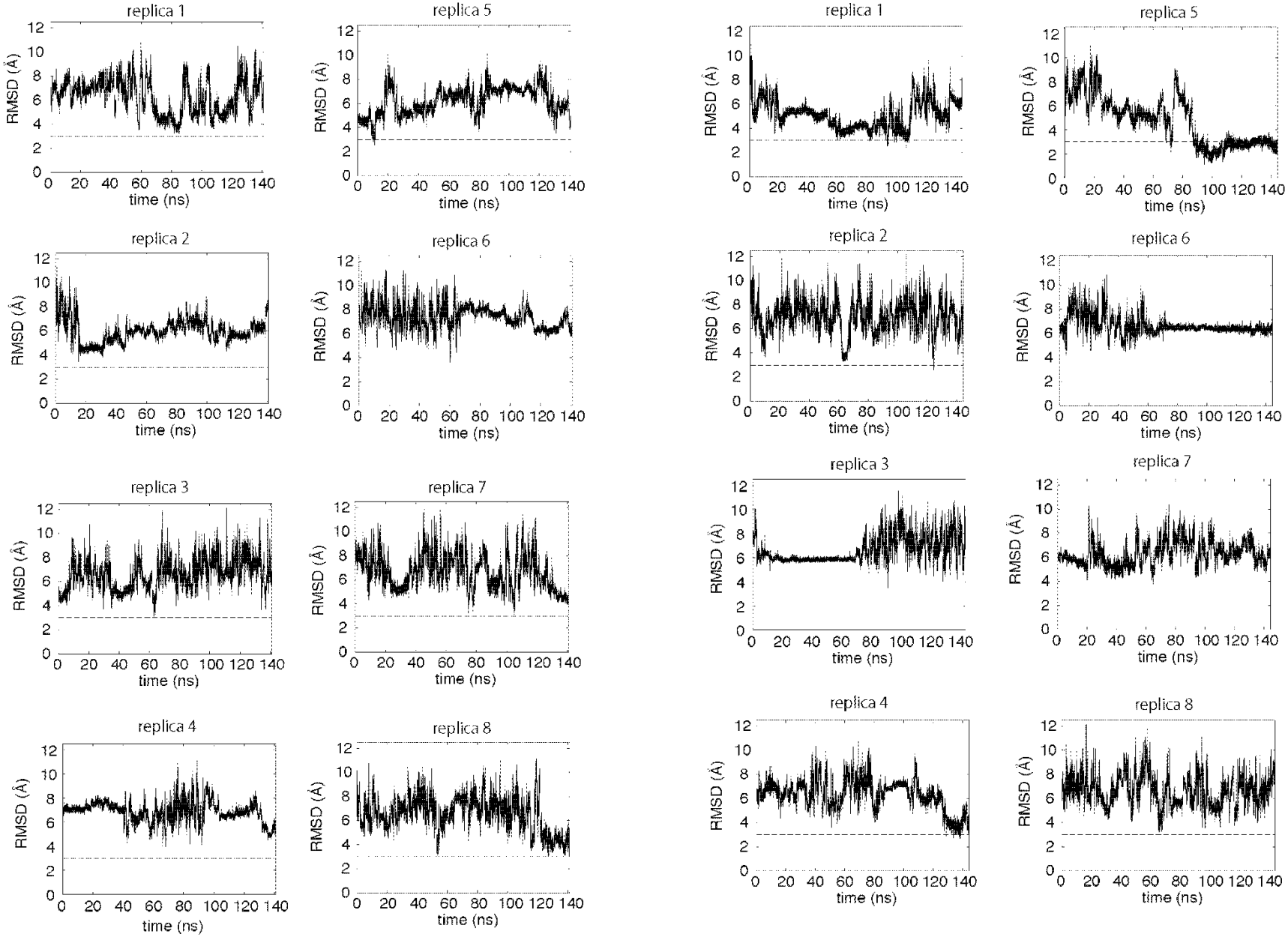




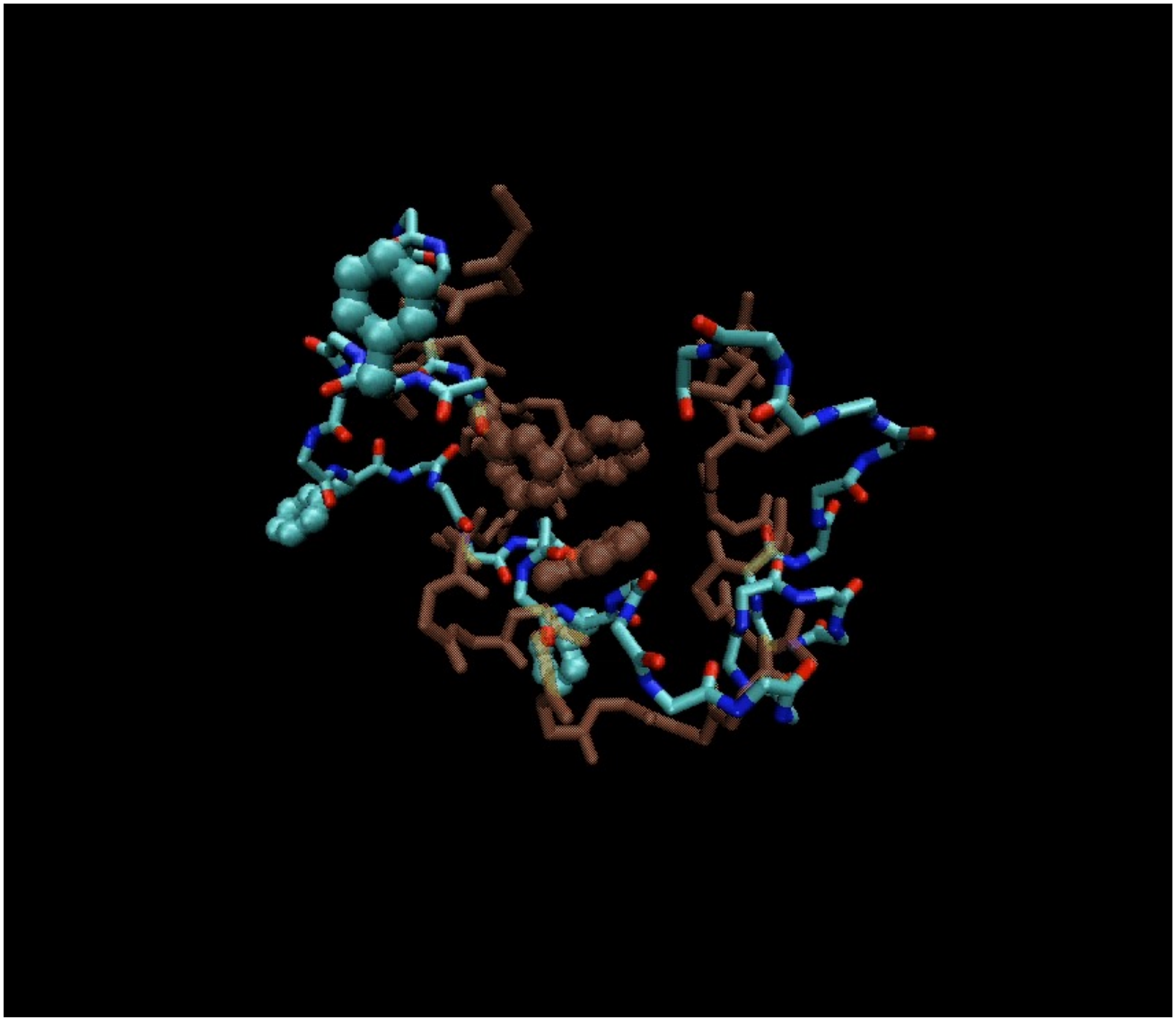




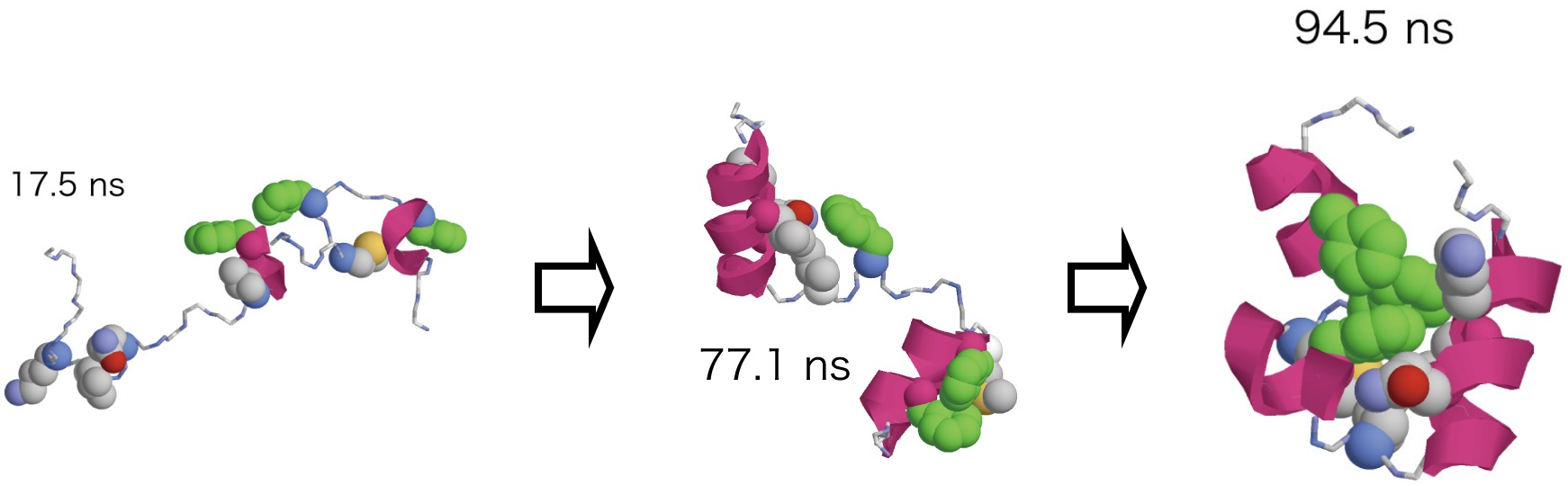
RMSD Time Series



MUCAREM2 (Replica 5)



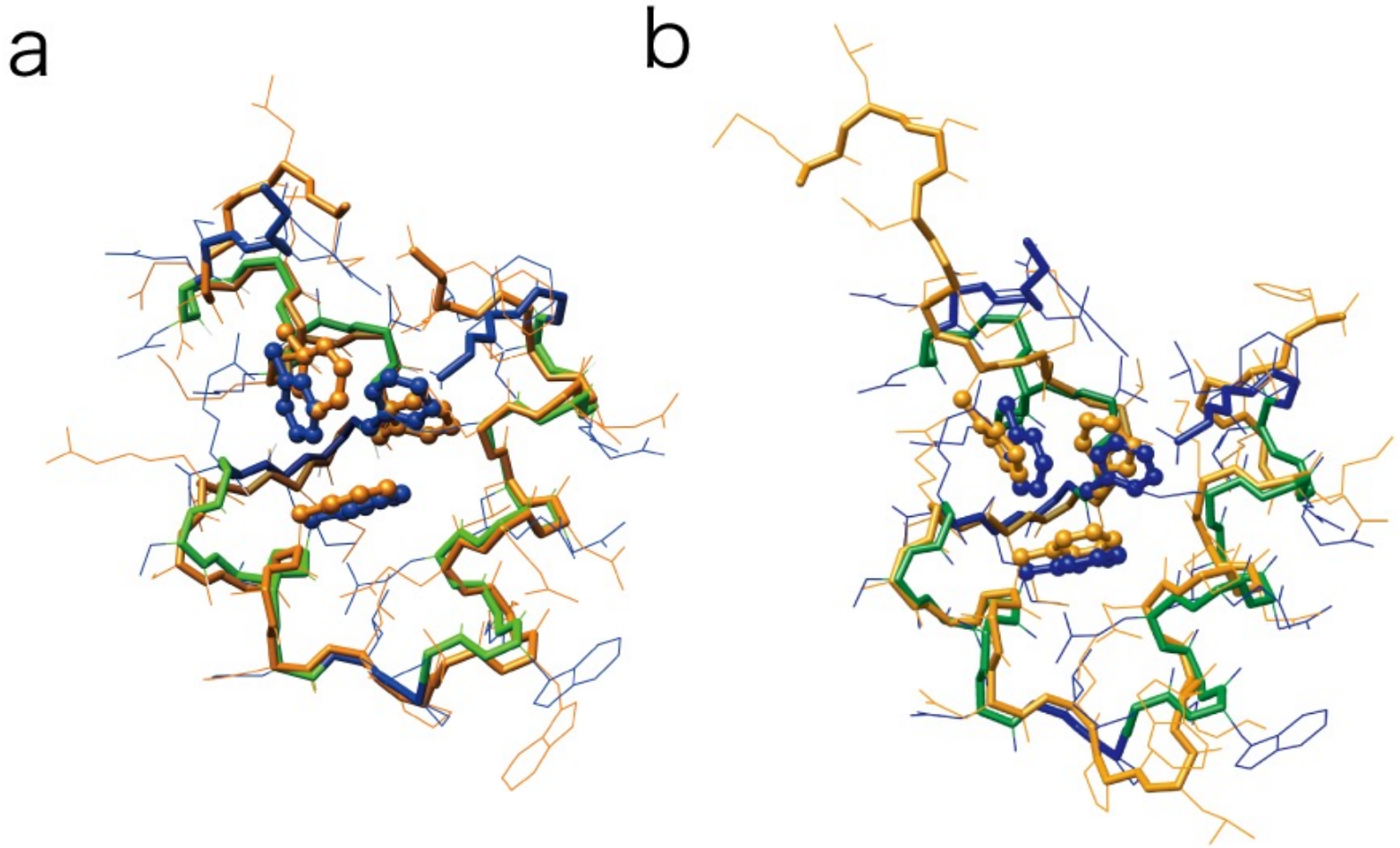
動画提供：**依田隆夫准教授**（長浜バイオ大学）



Snapshots from MUCAREM2

T. Yoda, Y. Sugita & Y.O., *Biophys. J.* **99**, 1637 (2010).

Native-Like Structures Obtained from MUCAREM2 (Left) and MUCAREM1 (Right)

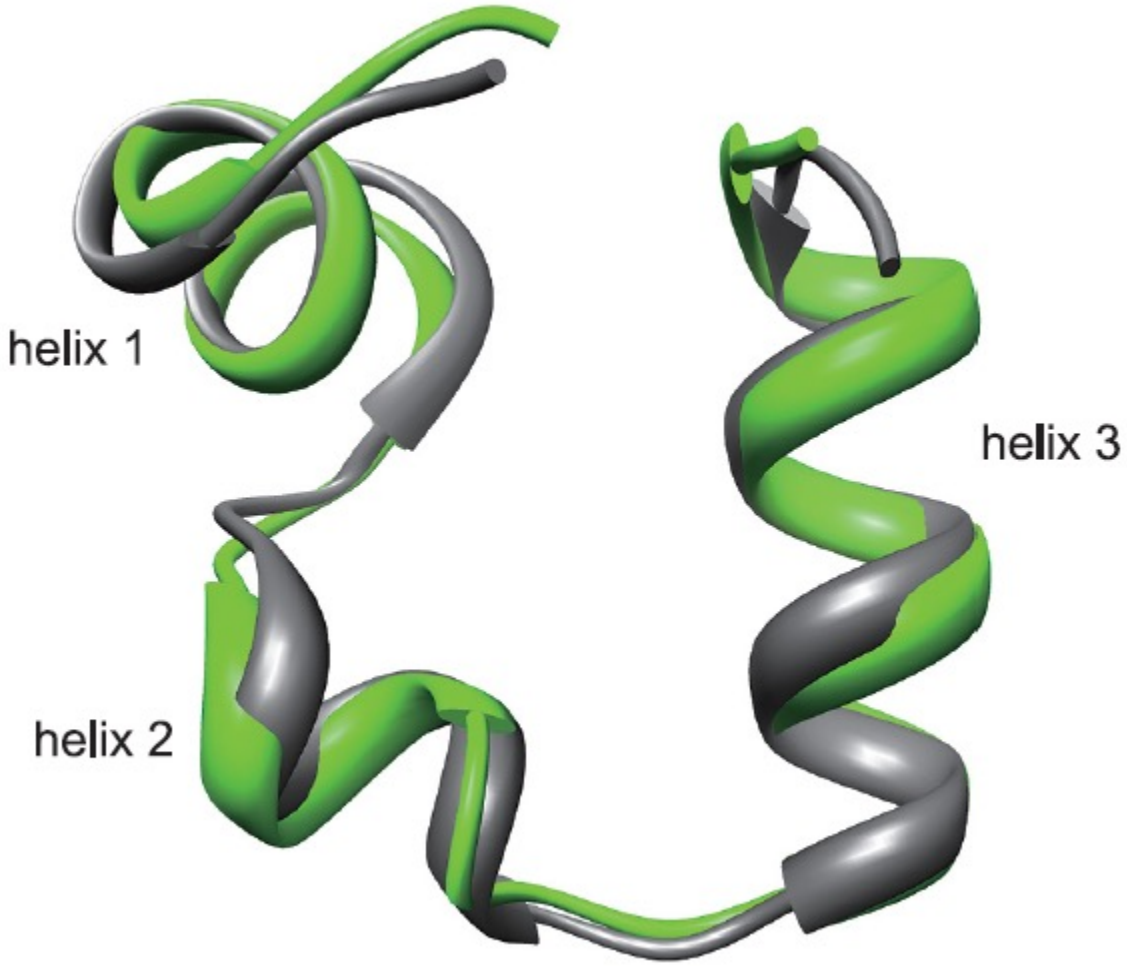


Main-Chain RMSD = 1.1 Å (residues 2 to 35) [Replica 5]

Main-Chain RMSD = 3.3 Å (residues 2 to 35) [Replica 8]
RMSD = 1.2 Å (residues 9 to 32)

Native-Like Structures Obtained from MUCAREM

T. Yoda, Y. Sugita & Y.O., *Biophys. J.* **99**, 1637 (2010).



Main-Chain RMSD = 1.1 Å　(residues 2 to 35) [Replica 5]
灰色：自然の構造（PDB ID: 1YRF）、緑色：シミュレーションの結果

α-Helix Formation

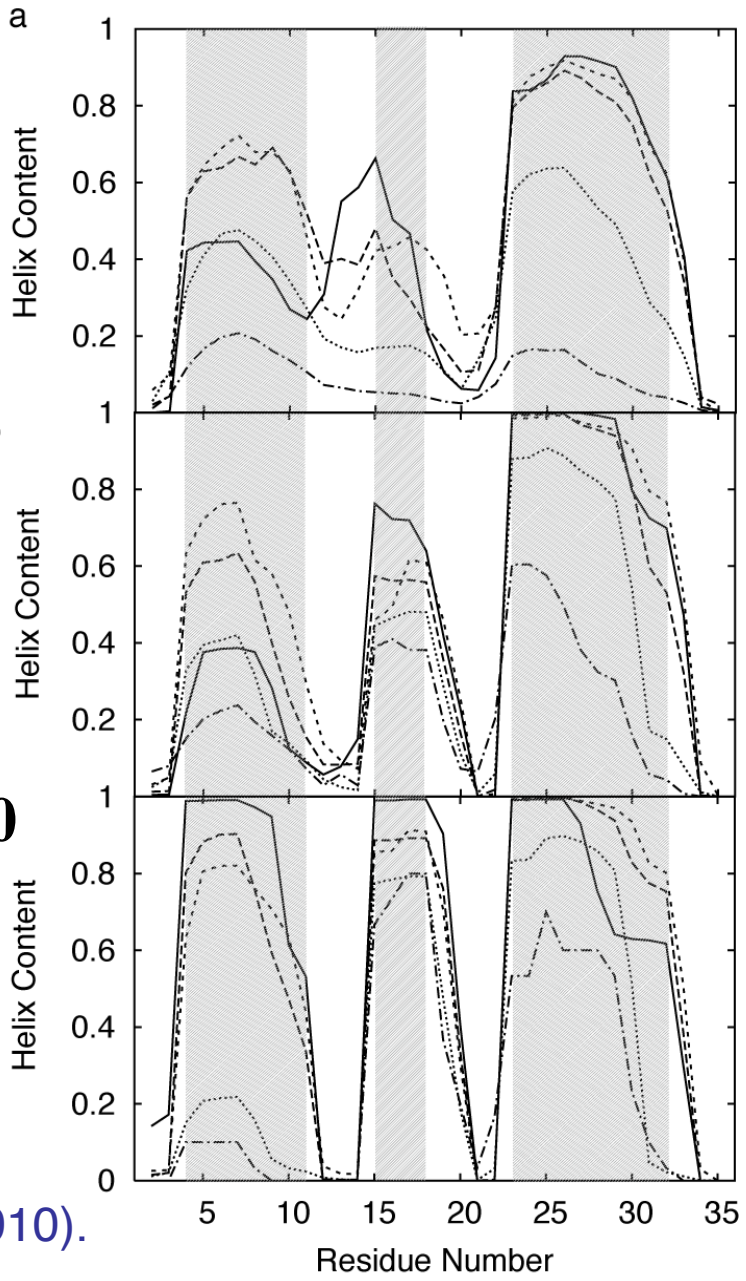
$N_n \leq 7$

N_n : no. of non-local side-chain native contacts (=18 for X-ray structure)

$N_n = 8, 9$

Solid: 300 K,
Dashed: 350 K,
Short Dashed: 400 K,
Dotted: 500 K,
Dotted Dash: 600 K

$N_n \geq 10$



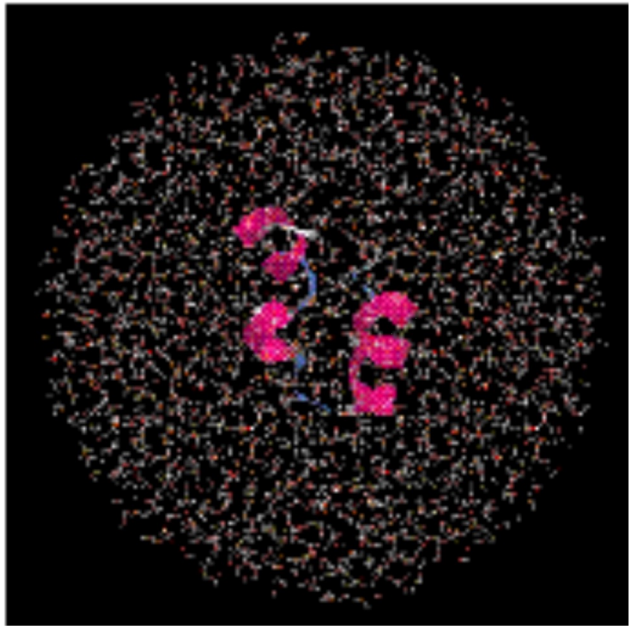
T. Yoda, Y. Sugita & Y.O., *Biophys. J.* **99**, 1637 (2010).

Salt Effects on Folding of a Small Globular Protein

T. Yoda, Y. Sugita & Y.O., *Proteins* **82**, 933-943 (2014).

Challenging the prediction of the 3-dimensional structure of a small protein by MUCAREM.

Villin headpiece subdomain
(36 amino acids; 596 atoms)
sphere of salted water with radius 30 Å
(3494 water molecules,
11 K^+, 13 Cl^- ≈ 0.2 M KCl);

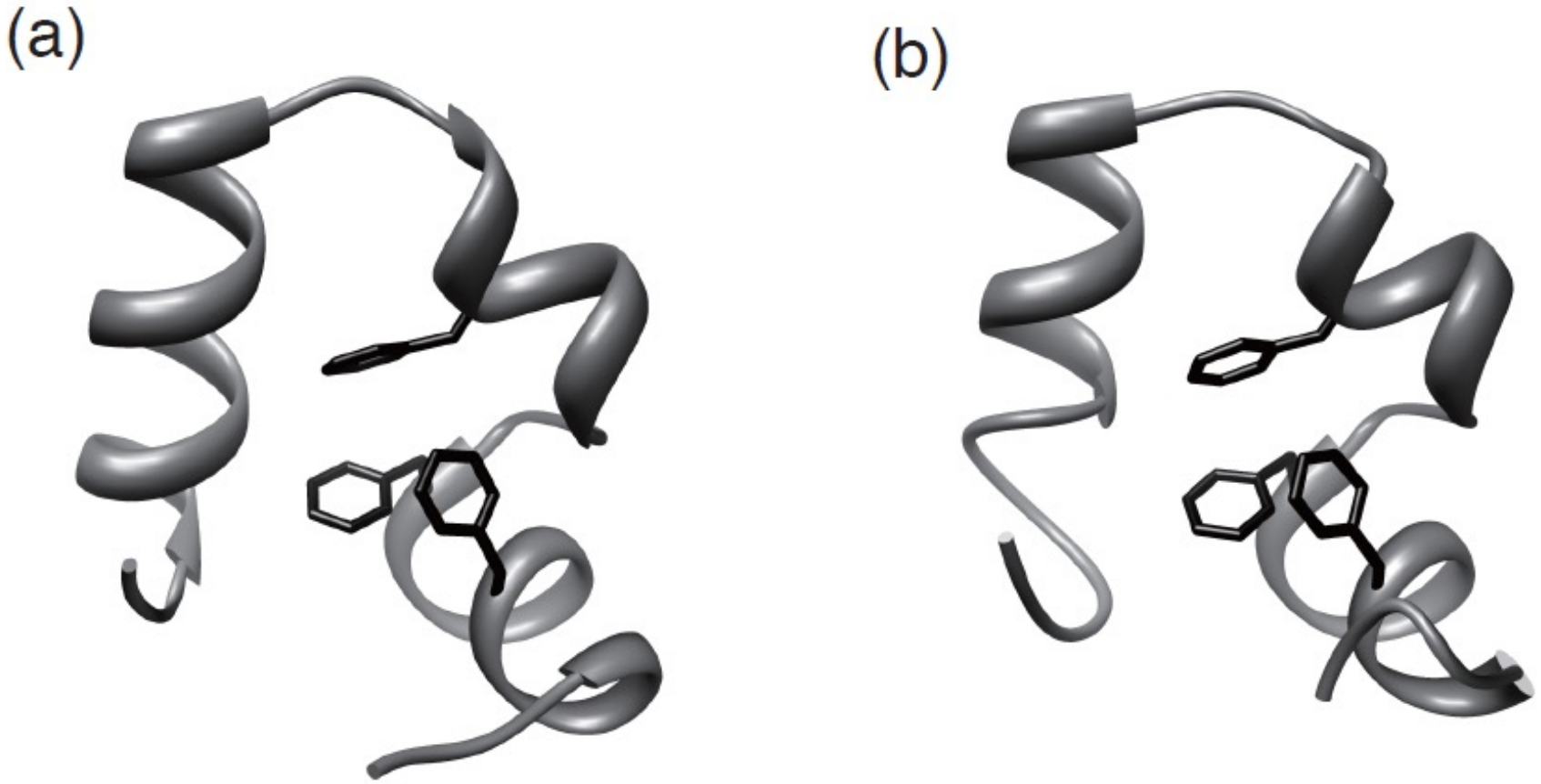


Primary Sequence of HP-36

```
1         10        20        30
MLSDEDFKAVFGMTRSAFANLPLWKQQNLKKEKGLF
 helix1    helix2    helix3
```

Native-Like Structure (Global Minimum in Free Energy) Obtained from MUCAREM Simulation (Left)

T. Yoda, Y. Sugita & Y.O., *Proteins* **82**, 933-943 (2014).

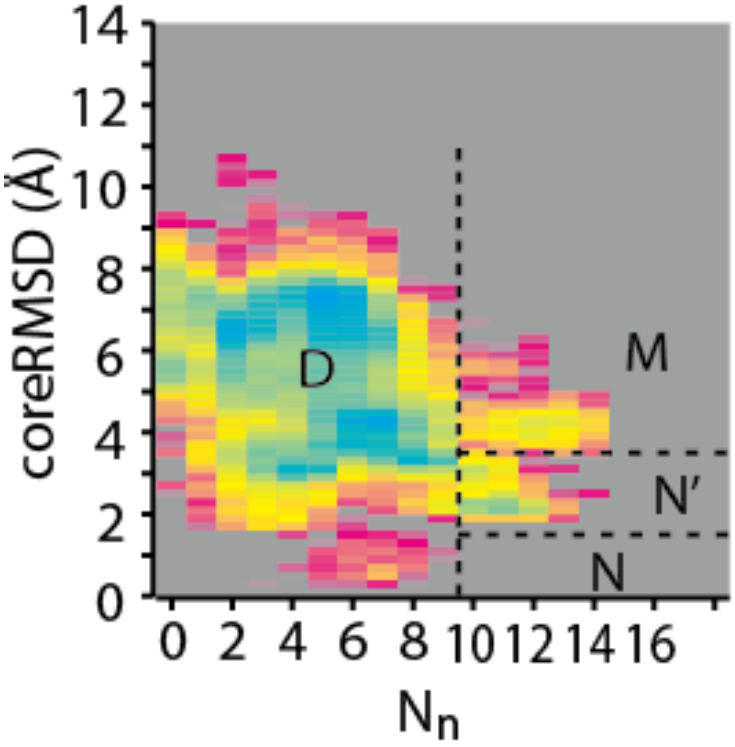


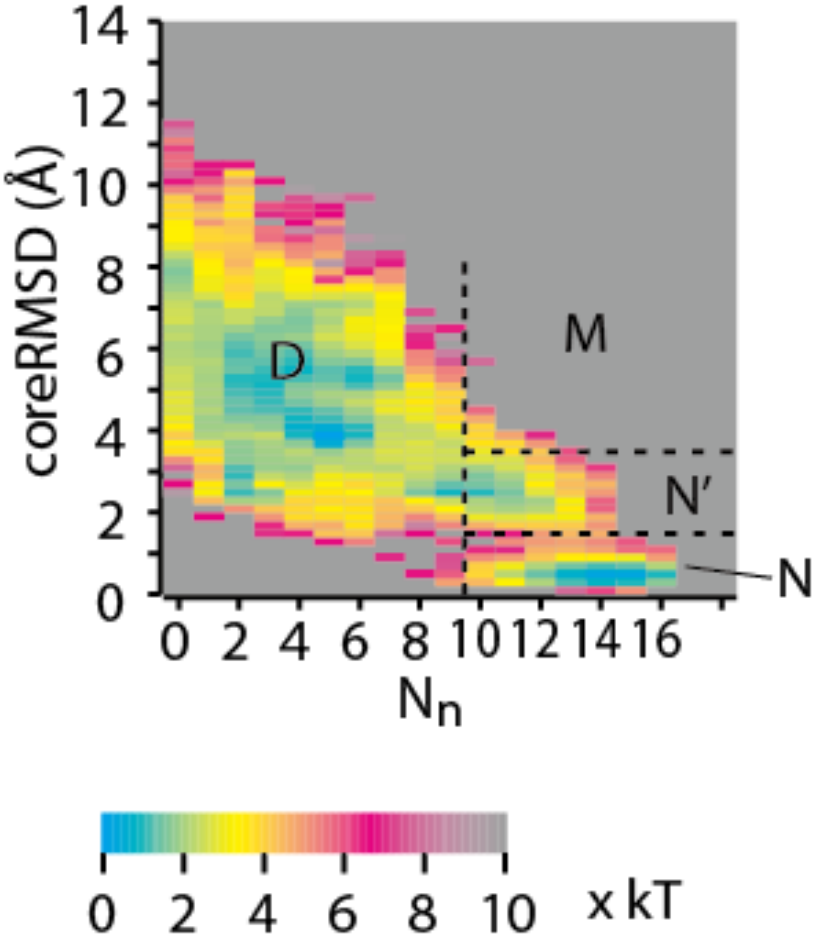
Main-Chain RMSD = 1.25 Å

Experimental Structure (PDB ID: 1YRF)

Free Energy Landscape

T. Yoda, Y. Sugita & Y.O., *Proteins* **82**, 933-943 (2014).





in Pure Water

in 0.2 M Salted Water

遺伝的アルゴリズムによるタンパク質の立体構造予測

Prediction of Protein Structures by Genetic Algorithms

T. Hiroyasu, M. Miki, M. Ogura & Y. O., *J. IPS Jpn.* **43**, 70-79 (2002).

Y. Sakae, T. Hiroyasu, M. Miki & Y. O., *Pacific Symposium on Biocomputing* **16**, 217-228 (2011).

Y. Sakae, T. Hiroyasu, M. Miki & Y. O., *J. Comput. Chem.* **32**, 1353-1360 (2011).

Y. Sakae, T. Hiroyasu, M. Miki, K. Ishii & Y.O., *J. Phys.: Conf. Ser.* **487**, 012003 (2014).

Y. Sakae, T. Hiroyasu, M. Miki, K. Ishii & Y.O., *Molecular Simulation* **41**, 1045-1049 (2015).

Y. Sakae, J. Straub, & Y.O., *J. Comput. Chem.* **40**, 475-481 (2018).

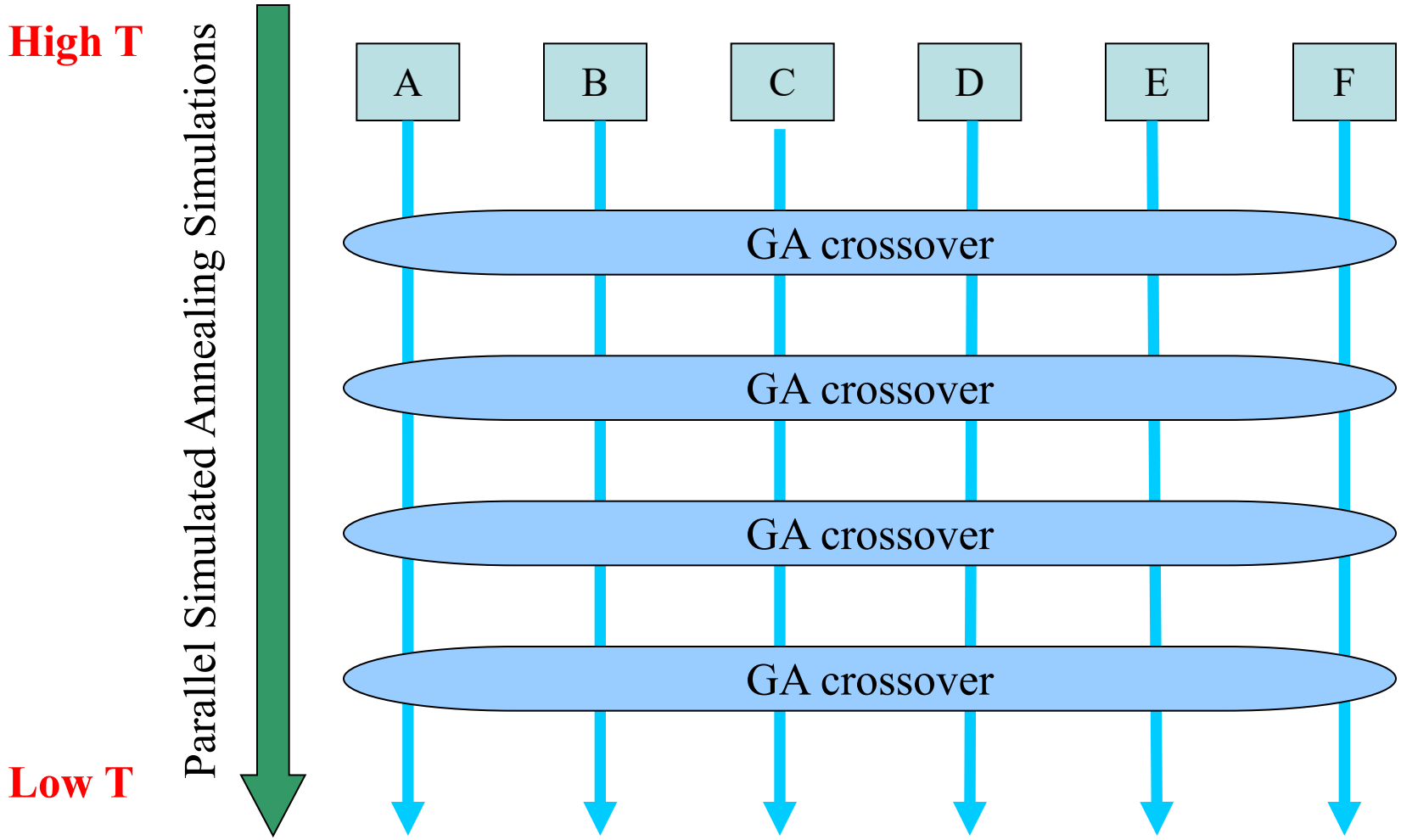
Parallel Simulated Annealing MC/MD with Genetic Crossover (PSAMC/GAc2 and PSAMD/GAc2)

T. Hiroyasu, M. Miki, M. Ogura & Y. O., *J. IPS Jpn.* **43**, 70-79 (2002).

Y. Sakae, T. Hiroyasu, M. Miki & Y. O., *J. Comput. Chem.* **32**, 1353-1360 (2011).

Step 1. Parallel simulated annealing simulations for a certain MC/MD steps

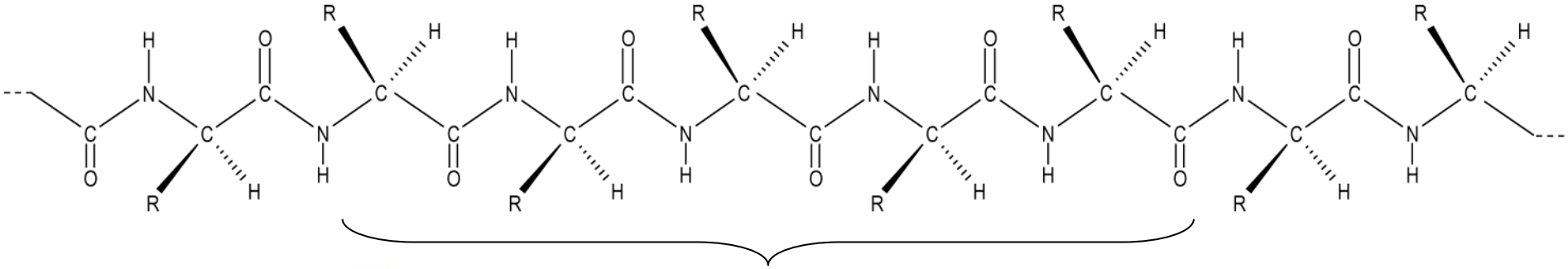
Step 2. Genetic crossover Repeat these two steps.

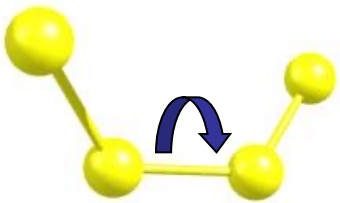


Genetic Crossover (2-Point Crossover)

Exchange a Randomly Chosen Pair of Dihedral Angle Sets

Structure A

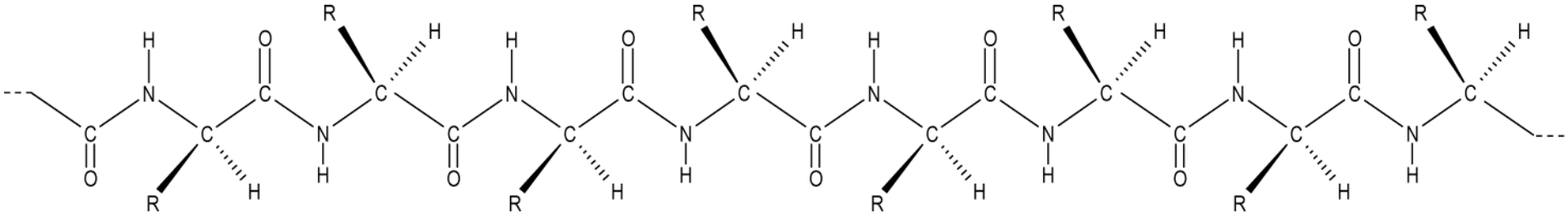




Dihedral Angle

All dihedral angles in *n* randomly selected consecutive amino-acid residues are exchanged. *n* = 2 – 10, motivated by the fragment assembly method (D. Baker).

Structure B



Instead of simulated annealing, we now repeat the following two steps:

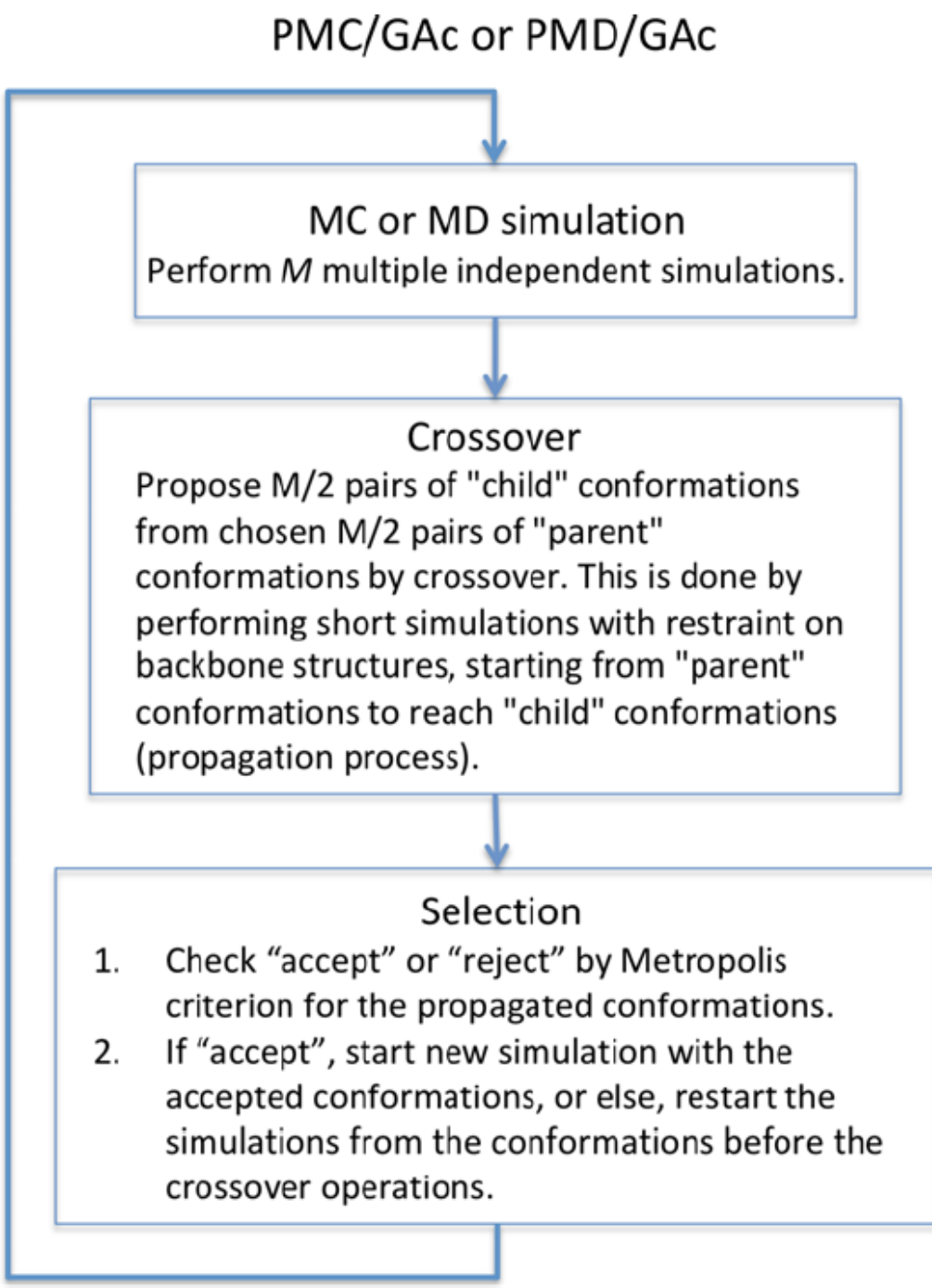
Step 1. [Local Update] Parallel MC/MD simulations at temperature *T*

Step 2. [Global Update] Genetic crossover

Hence, the detailed balance condition is assured and canonical ensemble at temperature *T* is obtained.

Cf. it is similar to hybrid MC.

By introducing M temperature values, we can easily generalize the method to REMC/REMD.



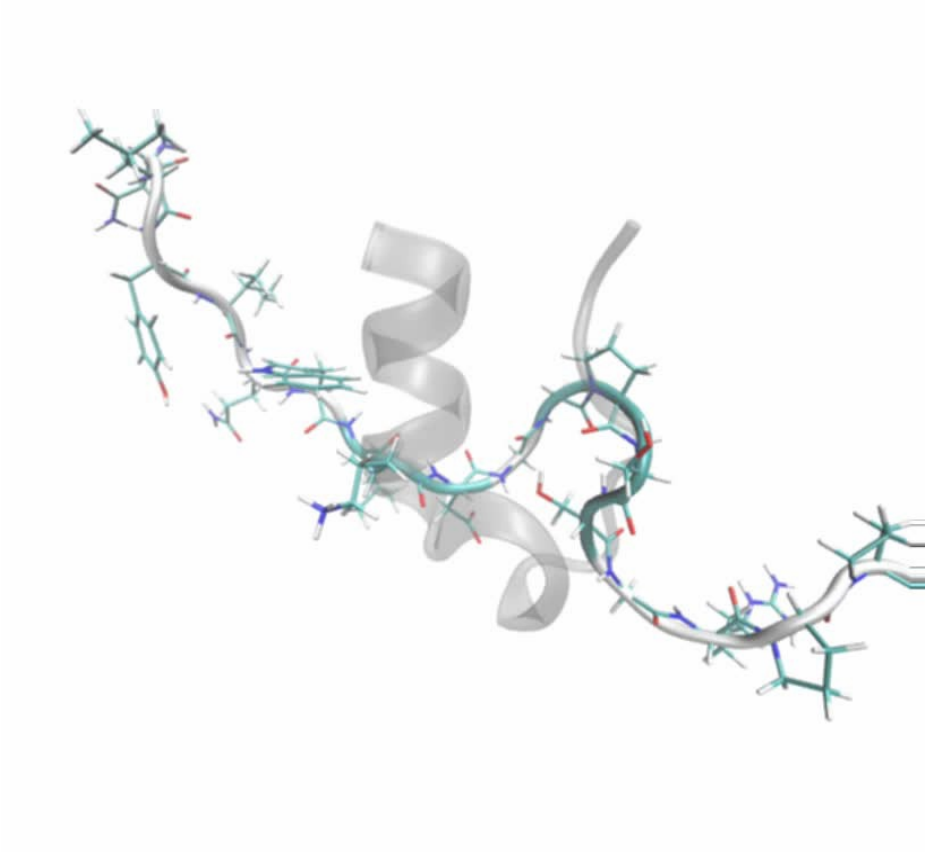
Y. Sakae, T. Hiroyasu, M. Miki, K. Ishii & Y.O., *Molecular Simulation* **41**, 1045-1049 (2015).

Y. Sakae, J. Straub, & Y.O., *J. Comput. Chem.* **40**, 475-481 (2018).

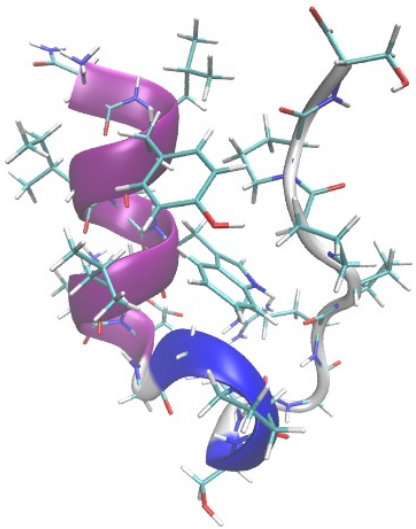
Example 1

Trp-Cage (PDB ID: 1L2Y)
20 residues

method: MD (Langevin dynamics)
temperature: 282 K
solvent: GB/SA
force field: AMBER ff03
no. of individuals: 16
simulation time per individual: 100ps × 100 （10ns）



Simulation （Individual No.2）



PDB Structure （NMR）

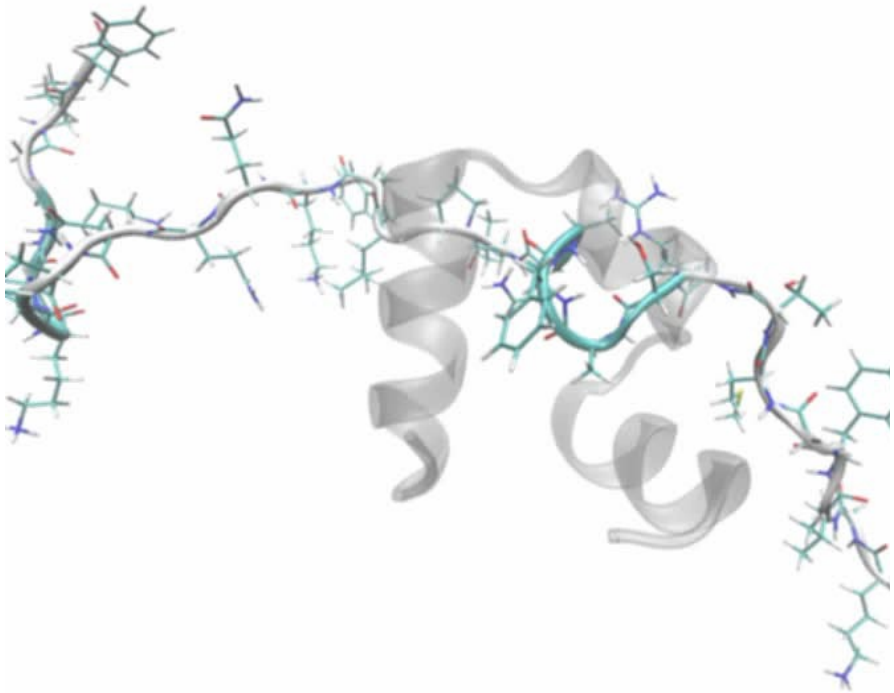
RMSD : 0.809 Å

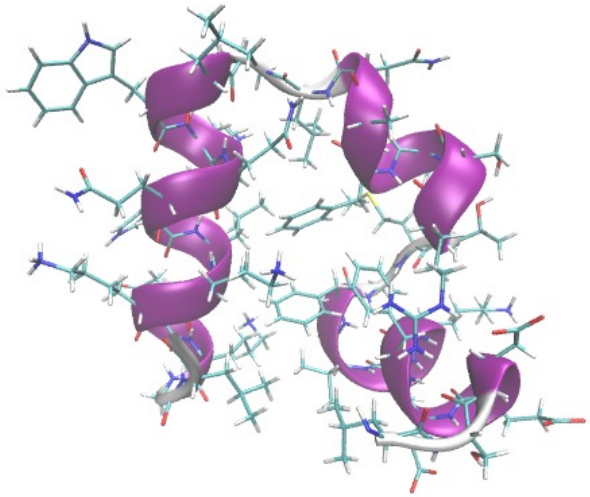
Y. Sakae, T. Hiroyasu, M. Miki, K. Ishii & Y.O., *Molecular Simulation* **41**, 1045-1049 (2015)

Example 2

Villin Headpiece (PDB ID: 1YRF)
36 residues

method: MD (Langevin dynamics)
temperature: 300 K
solvent: GB/SA
force field: AMBER ff03
no. of individuals: 32
simulation time per individual: 200ps × 100 (39.4ns)





Simulation (Individual No. 11)

PDB Structure (X-ray)

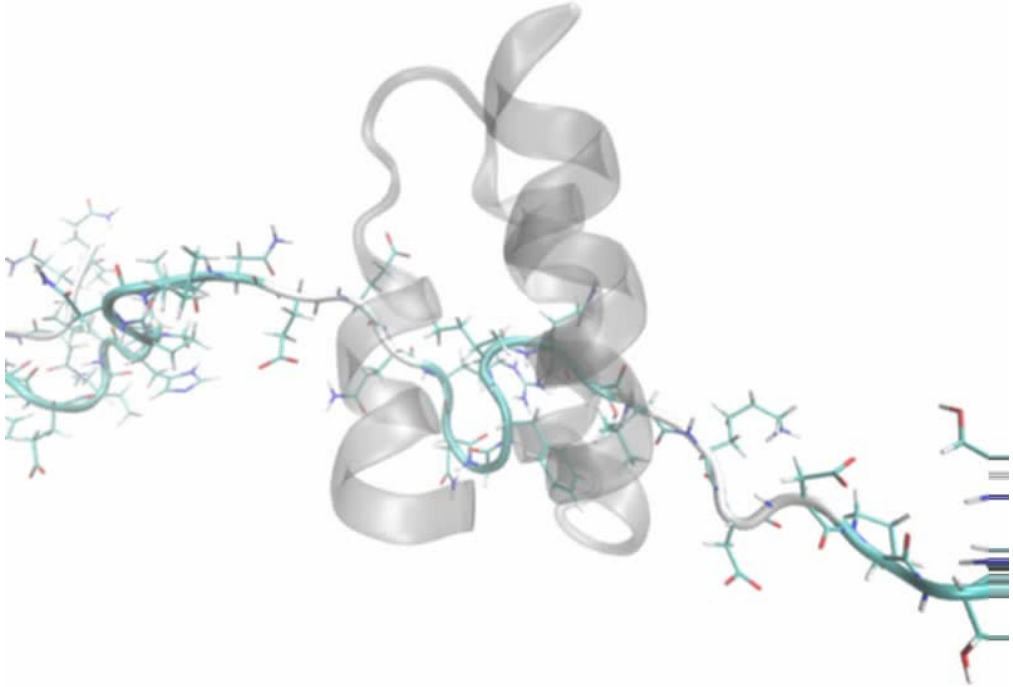
RMSD : 2.234 Å

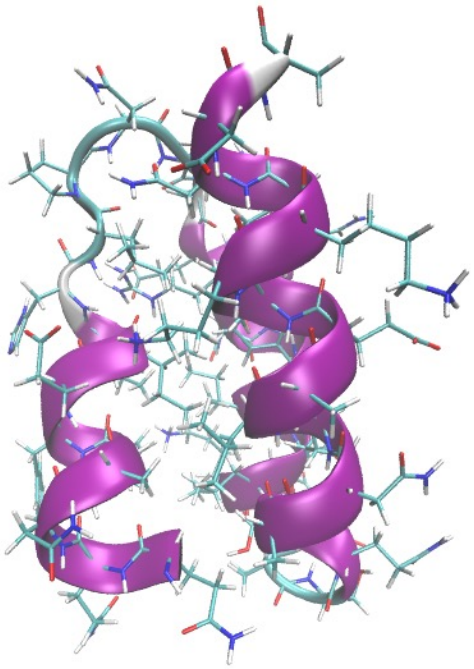
Y. Sakae, T. Hiroyasu, M. Miki, K. Ishii & Y.O., *Molecular Simulation* **41**, 1045-1049 (2015

Example 3

Protein A (PDB ID: 1BDD)
46 residues (10-55 out of 60)

method: MD (Langevin dynamics)
temperature: 300 K
solvent: GB/SA
force field: AMBER ff03
no. of individuals: 32
simulation time per individual: 1.0ns × 90 （90ns）





Simulation（Individual No. 5）

PDB Structure（NMR）

RMSD : 1.707 Å

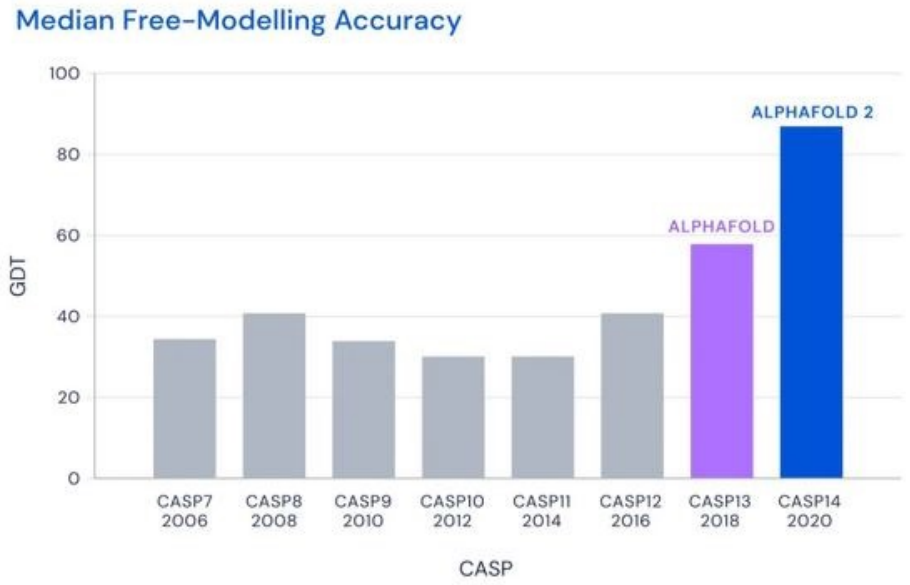
Y. Sakae, T. Hiroyasu, M. Miki, K. Ishii & Y.O., *J. Phys.: Conf. Ser.* **487**, 012003 (2014).

機械学習による蛋白質の立体構造予測（AlphaFold2）

with Ryuga Sugano

最近、AlphaFold,AlphaFold2というディープラーニングに基づく手法が、CASP(Critical Assessment of protein Structure Prediction)というタンパク質立体構造予測オンテストで圧倒的な正答率を挙げている。

CASP : タンパク質構造予測コンペ



https://doi.org/10.1126/science.370.6521.1144

https://www.itmedia.co.jp/news/articles/2012/01/news053.html

評価指標

評価指標としてRMSD(平均二乗偏差)を用いる

$$\mathrm{RMSD} = \min_{\text{平行移動、回転}} \sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta_i^2}$$

N : 比較したい全原子数
δ_i : i番目の原子の予測結果と対応する正解構造の距離(Å)

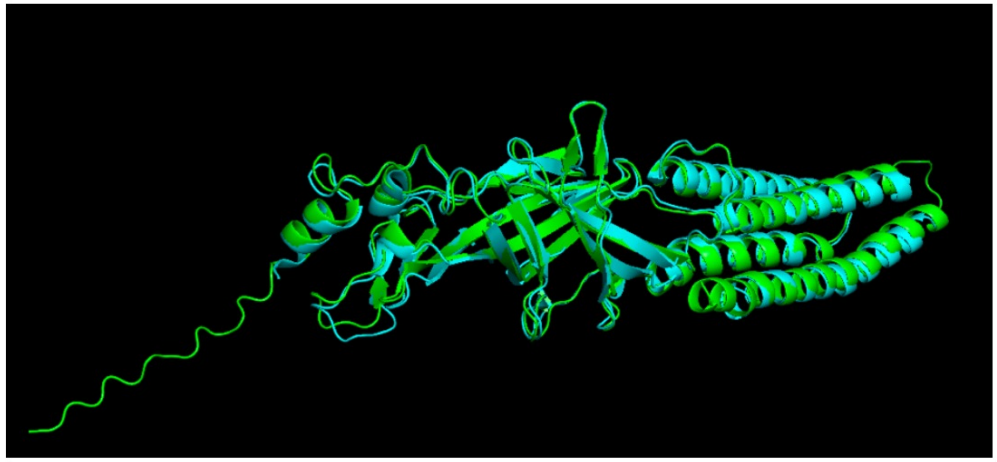
RMSD=0で完全一致、0に近いほど良い精度

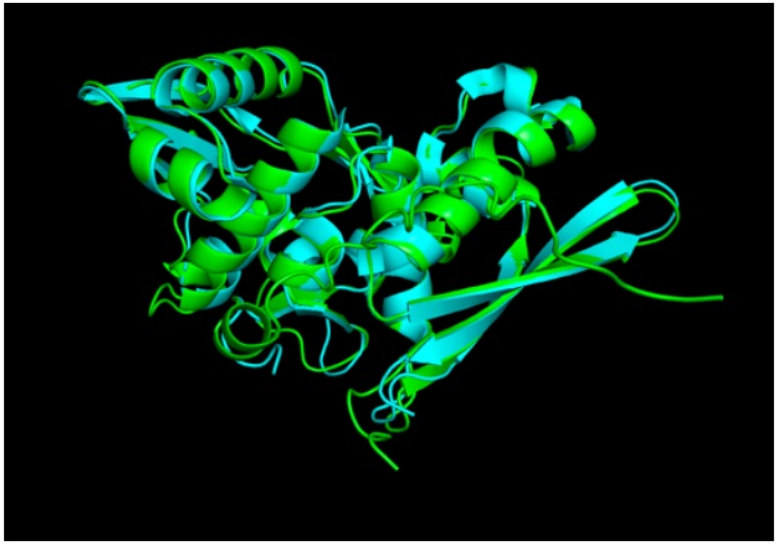
結果

②7RVA_chianC

成功(RMSD:0.677)

①5BKF(公開日：2021-09-08)　→　成功(RMSD:0.85)

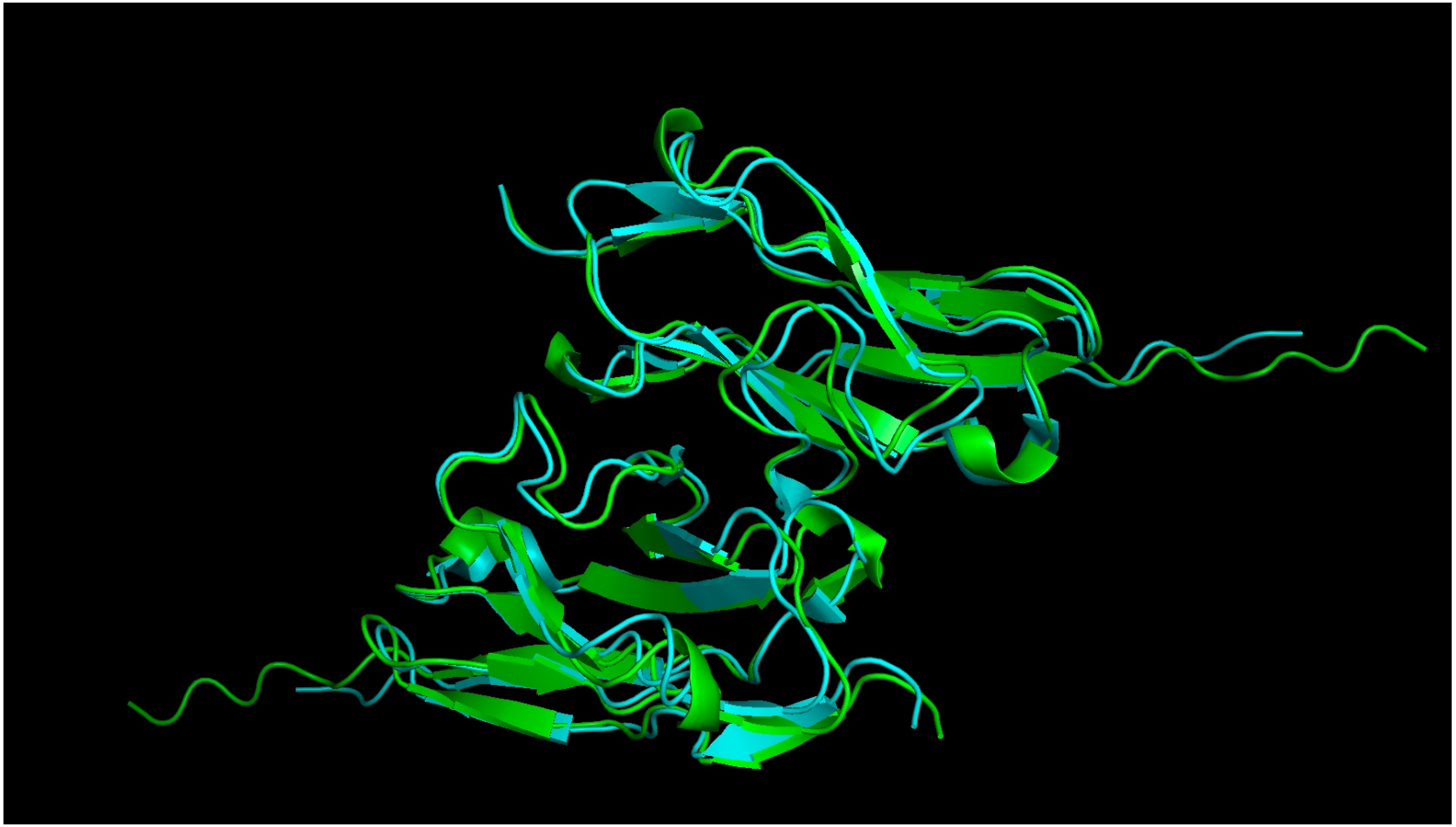




青：PDBの立体構造

緑：AlphaFold2の予測結果

結果(7LTW…二量体)

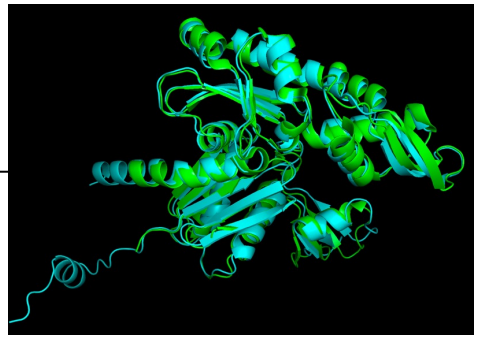


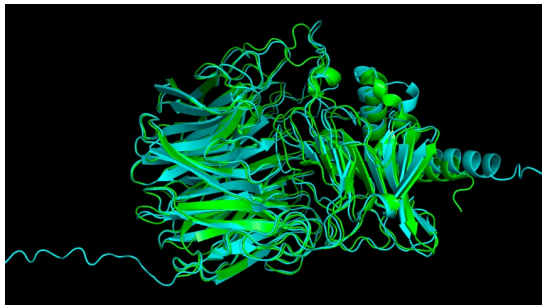
RMSD…1.485

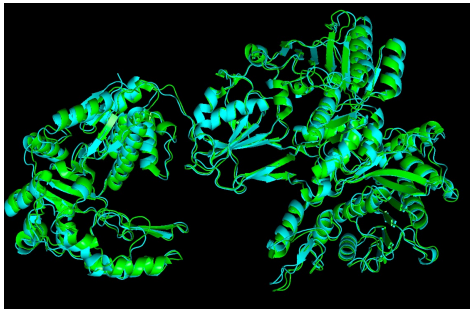
青：PDBの立体構造
緑：AlphaFold2の予測結果

AlphaFold2
ブラウザ版とインストール版(名大不老)

	ブラウザ版	インストール版(不老)
7E1C 355残基.	1時間55分 RMSD:1.406	54分 RMSD:1.591
7FH8 676残基	2時間32分 RMSD:0.914	1時間34分 RMSD:0.854
7WHP_A 1058残基	out of memory	1時間50分 RMSD:1.658







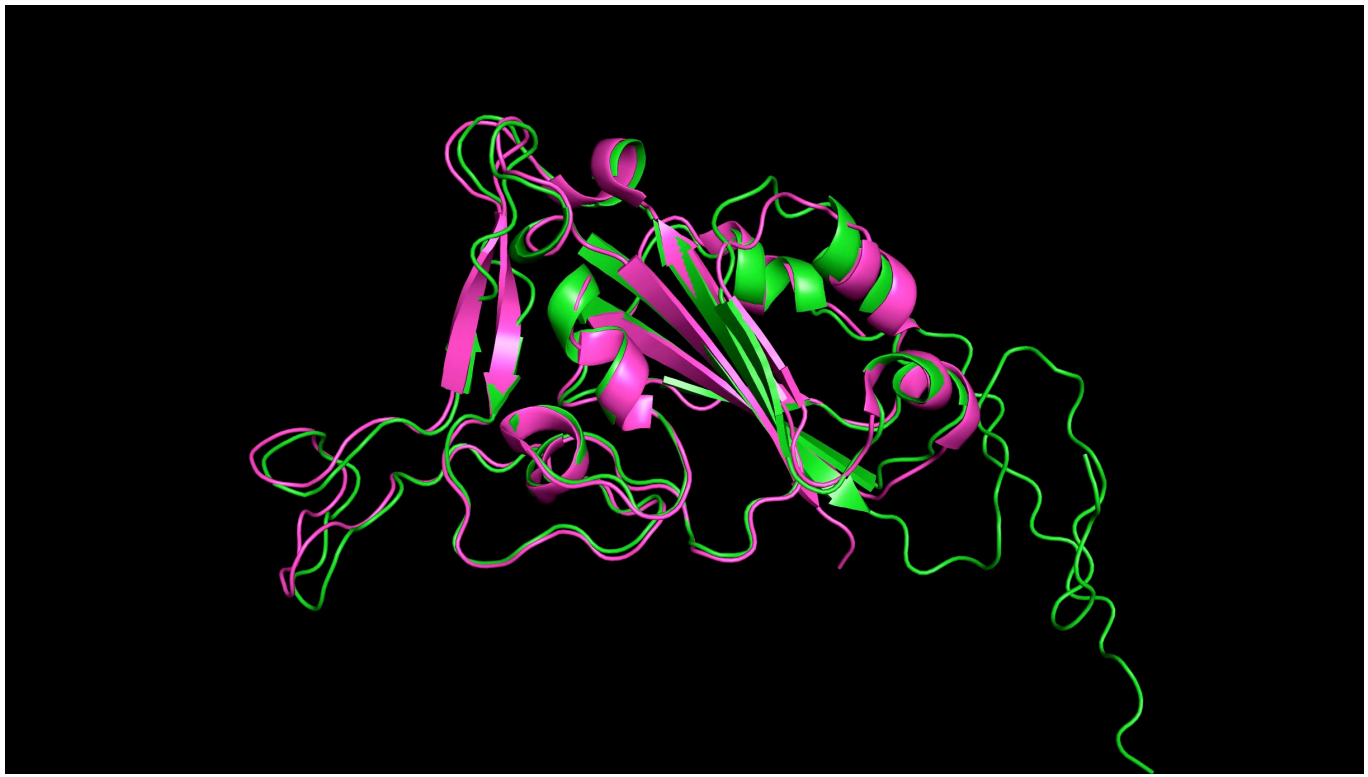
結果(7WHP_A,名大の不老を用いて)

7WHP_chainA
1058残基
RMSD…1.658

計算時間1時間
50分

青：AlphaFold2予測結果
緑：PDB構造

応用例：新型コロナウイルスのオミクロン株の Receptor Binding Domain (RBD)の構造)



223残基のうち
15残基に変異

RMSD:0.822

ピンク：オリジナル：新型コロナウイルスのRBDのPDB立体構造

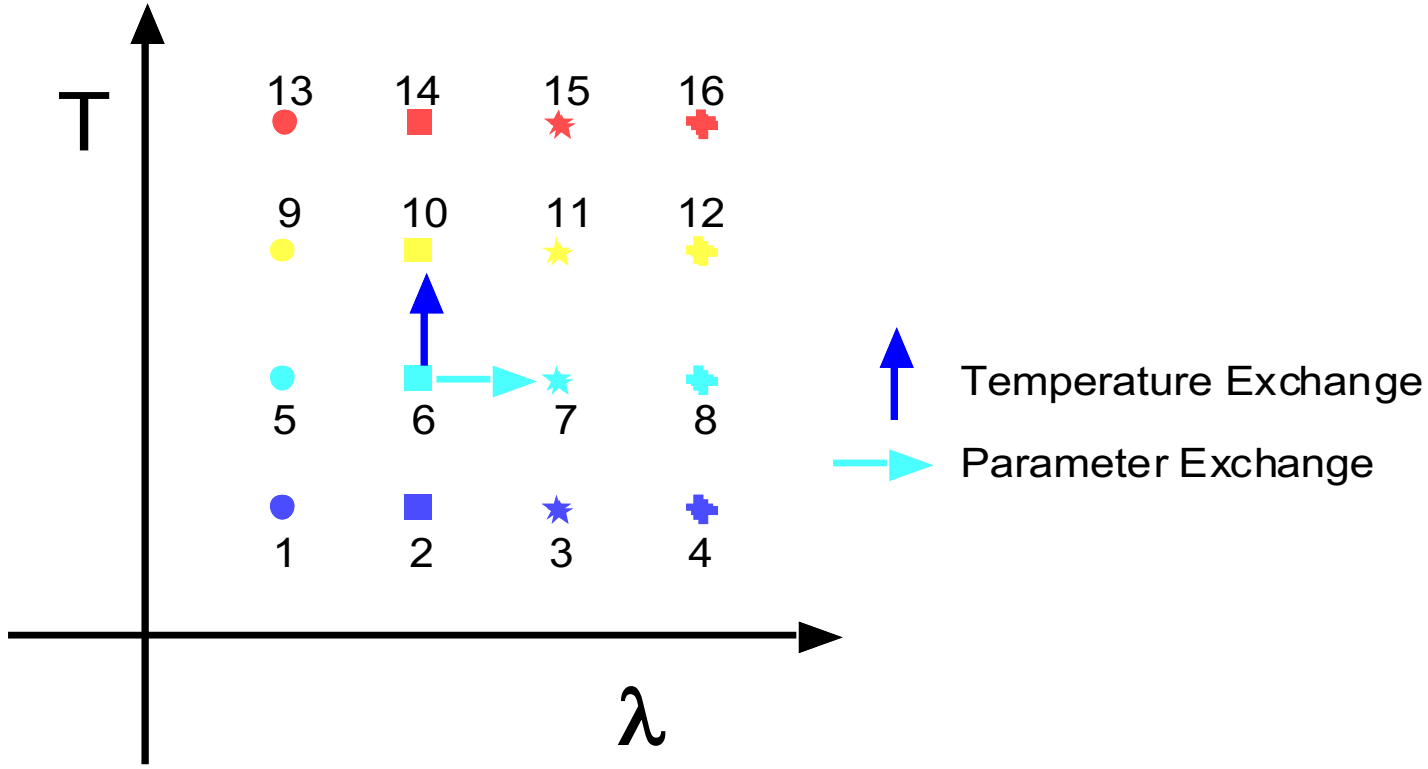
緑：オミクロン株のRBDのAlphaFold2の予測結果

Multidimensional Replica-Exchange Method (MREM)

Y. Sugita, A. Kitao & Y.O., *J. Chem. Phys.* **113**, 6042 (2000).

Hamiltonian Replica-Exchange Method (HREM)

H. Fukunishi, O. Watanabe & S. Takada, *J. Chem. Phys.* **116**, 9058 (2002).



Cf. Two-dimensional MUCA:

J. Lee, M.A. Novotny, & P.A. Rikvold, *Phys. Rev. E* **52**, 356 (1995).

J. Higo, N. Nakajima, H. Shirai, A. Kidera & H. Nakamura, *J. Comput. Chem.* **18**, 2086 (1997).

Multidimensional Replica-Exchange Method (MREM)（also referred to as Hamiltonian REM）

Y. Sugita, A. Kitao & Y.O., *J. Chem. Phys.* **113**, 6042 (2000).

1. System

M *Non-Interacting* Replicas of the Original System at *M* Different Sets of Temperatures and Parameters

2. Replica-Exchange

Step 1: Independent Canonical Simulations Performed for Each Replica

Step 2: A Pair of Replicas Corresponding to Neighboring Temperatures or Parameters are Exchanged

Repeat These 2 Steps

3. Canonical Distribution at Any Temperature

by Multiple Histogram Reweighting Techniques (WHAM)

Van der Waals Replica-Exchange Method (vWREM)

S.G. Itoh, H. Okumura & Y.O., *J. Chem. Phys.* **132**, 134105 (2010).

Hamiltonian

$$H_\lambda(q,p) = K(p) + E_\lambda(q)$$

Potential Energy

$$E_\lambda(q) = E_P(q_P) + E_{PS}(q_P, q_S) + E_S(q_S)$$

where P stands for protein and S solvent

$$E_P(q_P) = V_\lambda(q_P) + E_{rest}(q_P)$$

and

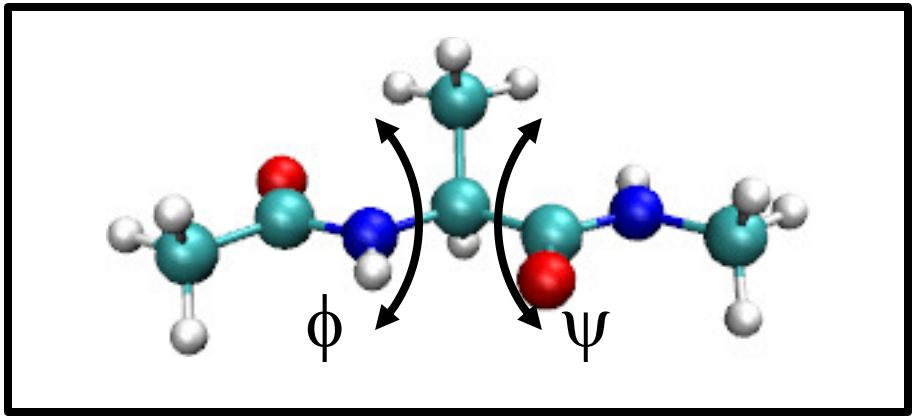
$$V_\lambda(q_P) = \sum_{k=1}^{N_P-1} \sum_{l=1}^{N_P} 4\varepsilon_{kl} \left[\left(\frac{\lambda\sigma_{kl}}{r_{kl}} \right)^{12} - \left(\frac{\lambda\sigma_{kl}}{r_{kl}} \right)^{6} \right]$$

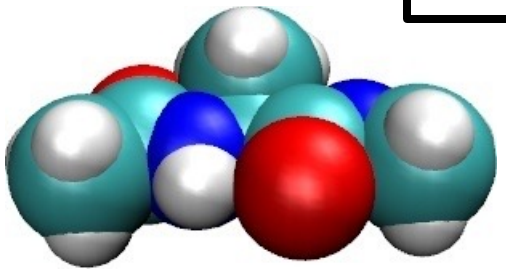
Namely, the van der Waals radii of the atoms only in the P-P interactions are scaled by λ.

Van der Waals Replica-Exchange Method (vWREM)

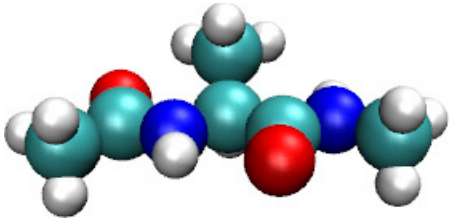
S.G. Itoh, H. Okumura & Y.O., *J. Chem. Phys.* **132**, 134105 (2010).

Alanine dipeptide





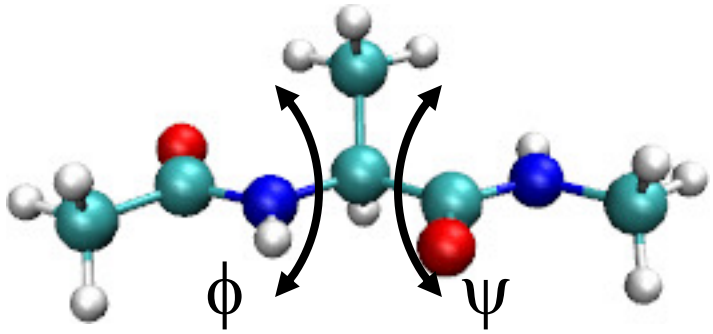
Decrease atom size

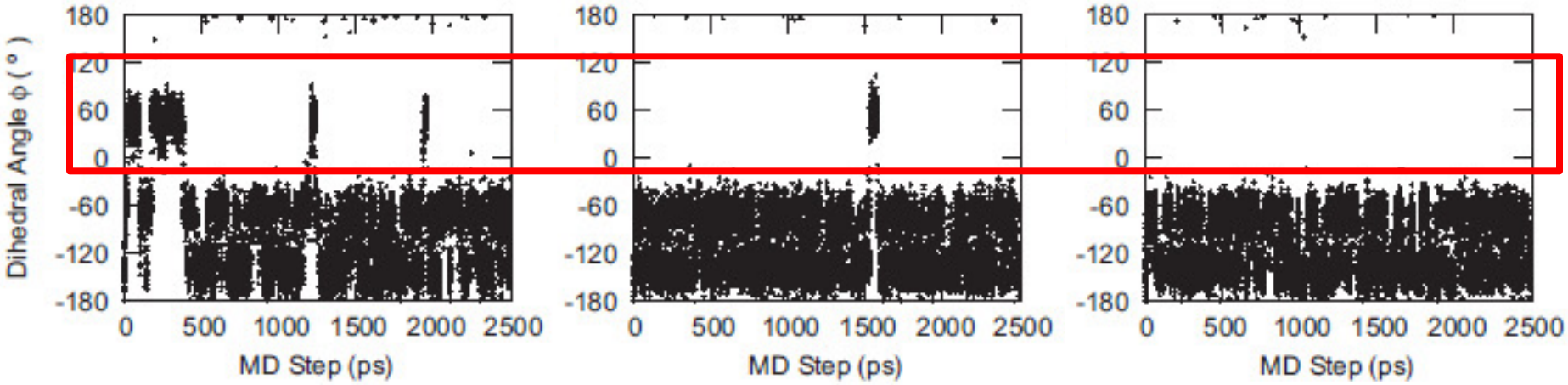


Original atom size

Move easily!

Time series of dihedral angles ϕ





Van der Waals REMD > Conventional REMD > Canonical MD

Effectiveness of samplings

Multi-Overlap Algorithm

MC version: B. Berg, H. Noguchi & Y.O., *Phys. Rev. E* **68**, 036126 (2003).
MD version: S.G. Itoh & Y.O., *Chem. Phys. Lett.* **400**, 308 (2004);
S.G. Itoh & Y.O., *J. Chem. Phys.* **124**, 104103 (2006).

Multi-overlap algorithm realizes a random walk in the configurational space between two appropriate reference configurations.

▪ dihedral-angle distance

A dihedral-angle distance is defined by

$$d = \frac{1}{\pi n} \sum_{i=1}^{n} d_a(v_i, v_i^0) .$$

where v_i is the dihedral angle i, v_i^0 is the dihedral angle i of the reference configuration, and the distance $d_a(v_i, v_i^0)$ is given by

$$d_a(v_i, v_i^0) = \min(|\, v_i - v_i^0 \,|, 2\pi - |\, v_i - v_i^0 \,|)$$

$$0 \le d_a(v_i, v_i^0) \le \pi.$$

Therefore,

$$0 \le d \le 1.$$

If $d = 0$, all dihedral angles are coincident with those of the reference configuration.

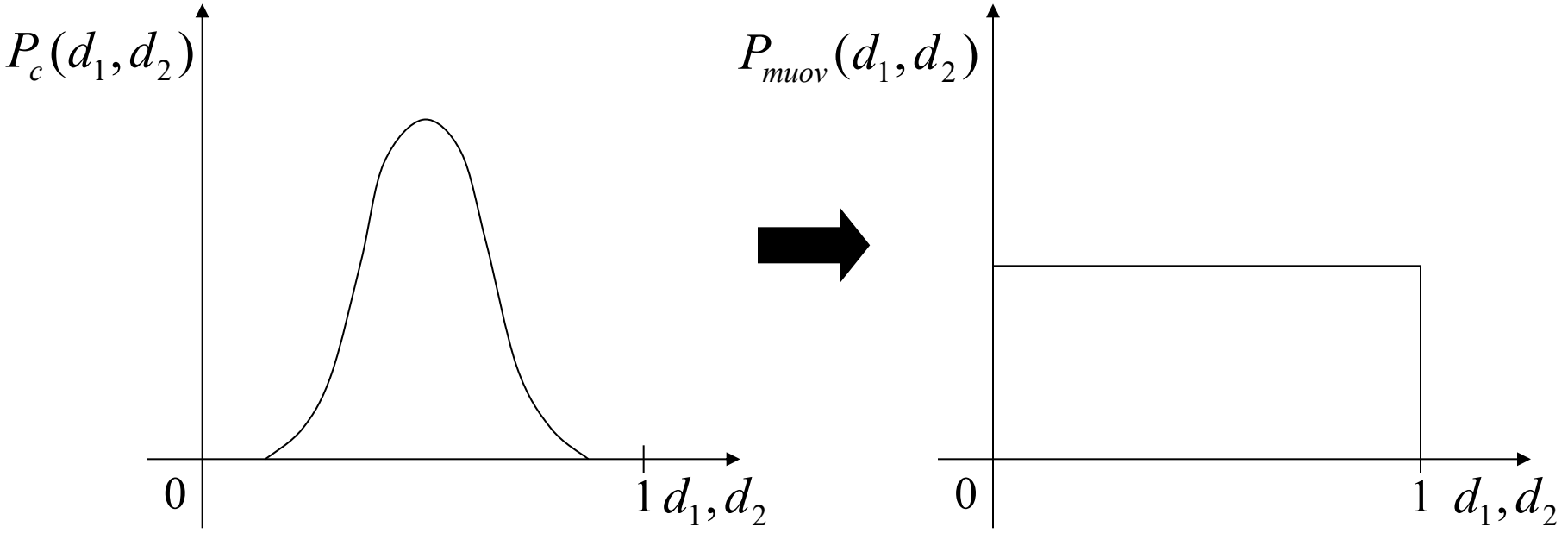
⟶ dihedral-angle distance is an indicator of how similar the configuration is to the reference configuration.

In the canonical ensemble at a constant temperature, the probability distribution with respect to two dihedral-angle distances d_1 and d_2 is given by

$$P_c(d_1, d_2) = \int dEn(d_1, d_2; E) \exp(-\beta E)$$

In the multi-overlap ensemble at a constant temperature, utilizing a function $f(d_1, d_2)$, the probability distribution with respect to two dihedral-angle distances is constant:

$$P_{muov}(d_1, d_2) = \int dEn(d_1, d_2; E) \exp(-\beta E + f(d_1, d_2)) = const$$



The multi-overlap weight factor is defined by

$$W_{muov} = \exp(-\beta E_{muov}) = \exp[-\beta(E - k_B T f(d_1, d_2)].$$

The multi-overlap MD simulation is performed by solving the following modified Newton equation:

$$\frac{dp_i}{dt} = -\frac{\partial E_{muov}}{\partial q_i} = -\frac{\partial E}{\partial q_i} + k_B T \frac{\partial f(d_1, d_2)}{\partial q_i}.$$

The function $f(d_1, d_2)$ is determined by the following process. At first, we set $f(d_1, d_2) = 0$. Then, if $d_1 = c_1, d_2 = c_2$ after ith MD step, the function $f(d_1, d_2)$ is updated by

$$f_i(d_1 = c_1, d_2 = c_2) = f_{i-1}(d_1 = c_1, d_2 = c_2) - a.$$

Here a is an appropriately chosen positive constant. This procedure is iterated until the probability distribution P_{muov} is flat.

(S. G. Itoh and Y. Okamoto, Chem. Phys. Lett. 400 (2004) 308)

The density of states is defined by

$$n(E) = \int dd_1 dd_2 N(d_1, d_2; E) \exp(\beta E - f(d_1, d_2)),$$

where $N(d_1, d_2; E)$ is the histogram of the probability distribution.

The expectation value of a physical quantity A at any temperature T' is give by

$$\langle A \rangle_{T'} = \frac{\int dE A(E) n(E) e^{-\beta' E}}{\int dE n(E) e^{-\beta' E}}$$

simulation : canonical MD simulation

multi-overlap MD simulation

system : Met-enkephalin (Tyr-Gly-Gly-Phe-Met) in vacuum.
the N-terminus and the C-terminus were blocked with the acetyl group and the N-methyl group, respectively.

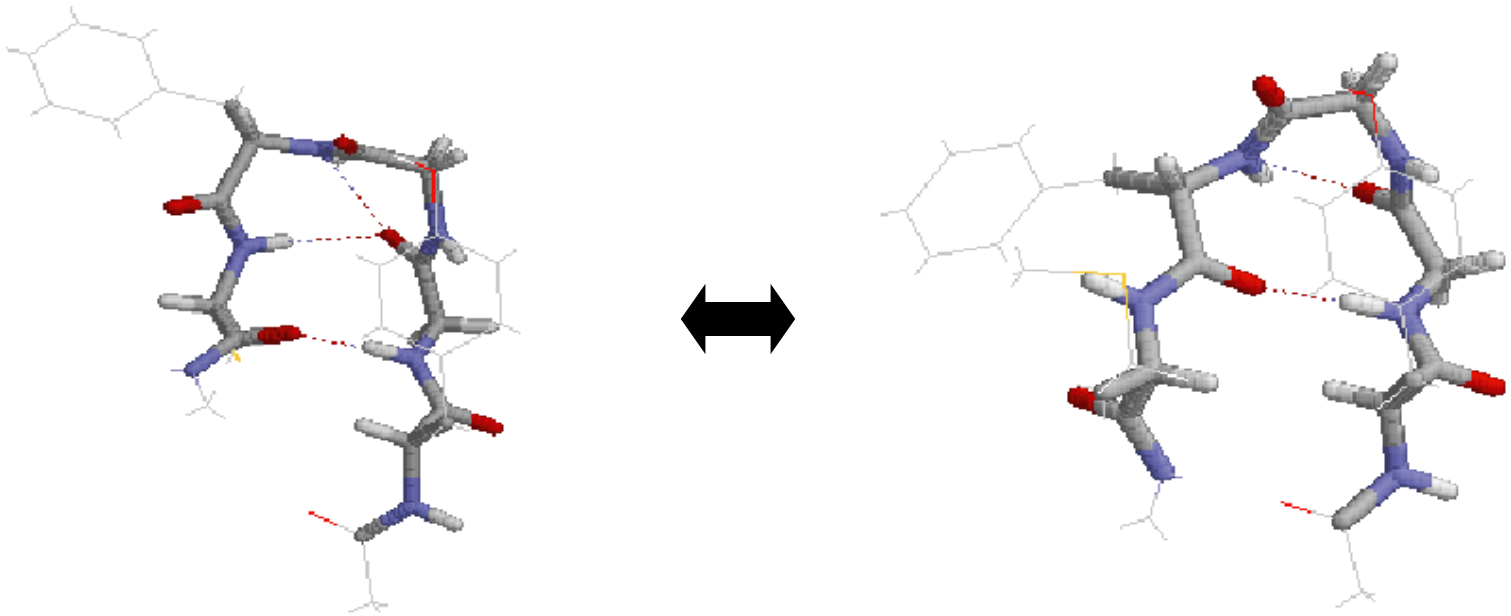
temperature : 300K (Gaussian constraint method)

force field : CHARMM param 22 (all-atom model)

time step : 0.5fs (leap-frog method)

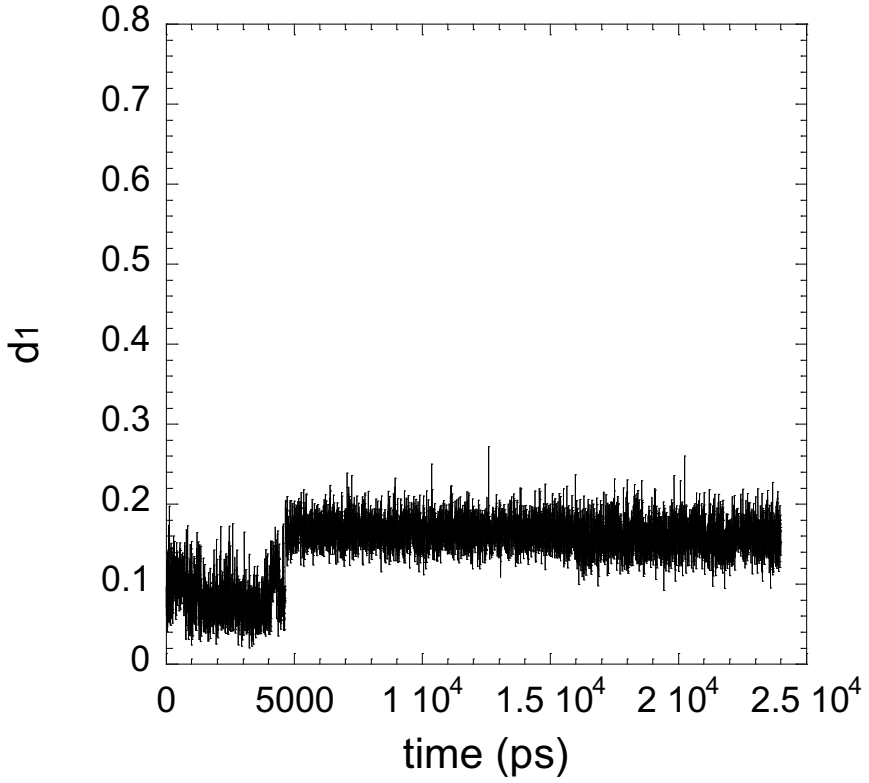
total simulation time : 24ns (+1ns equilibration)

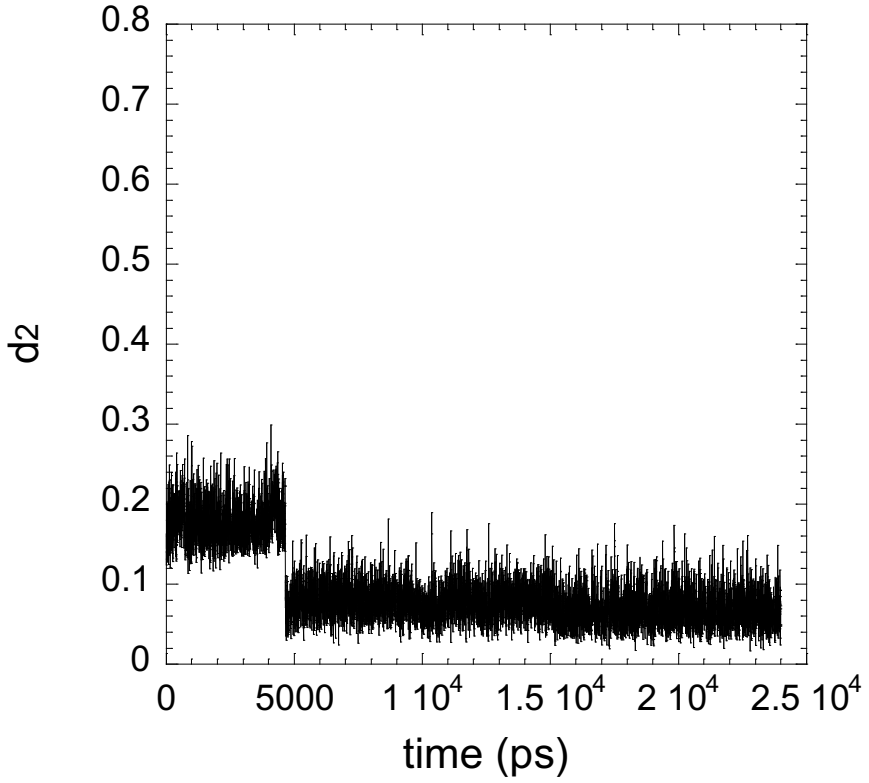
We performed a multi-overlap MD simulation for Met-enkephalin with two reference configurations.



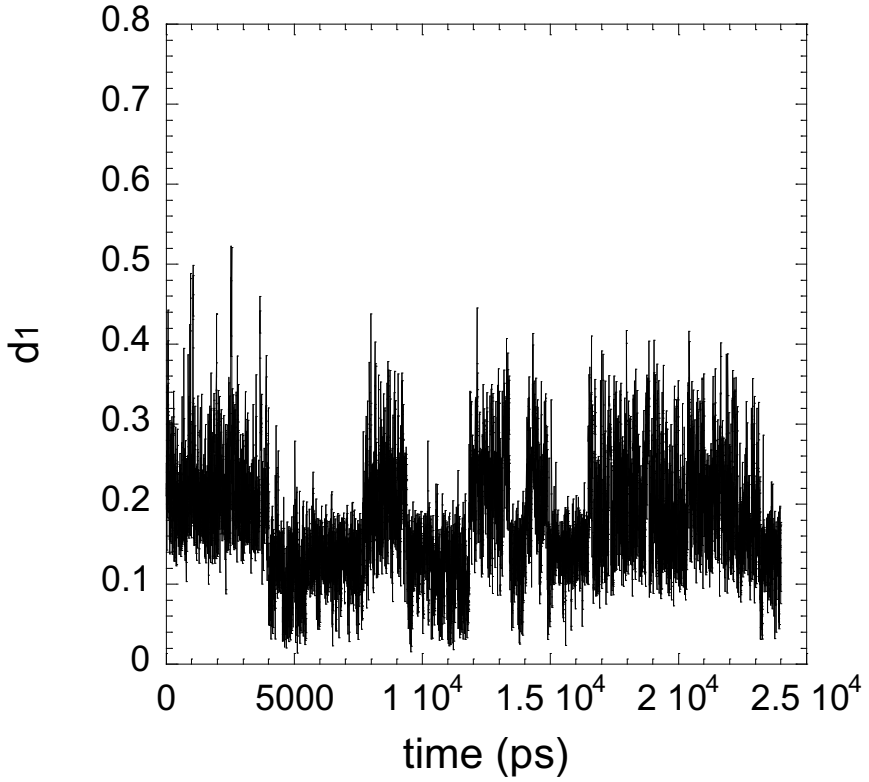
Reference configuration 1

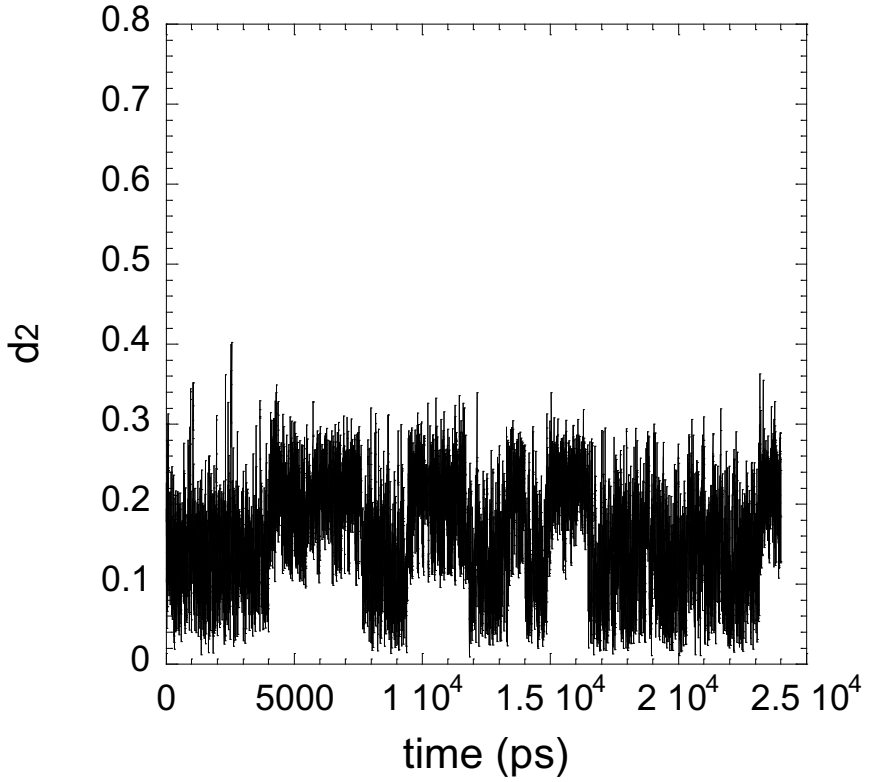
Reference configuration 2



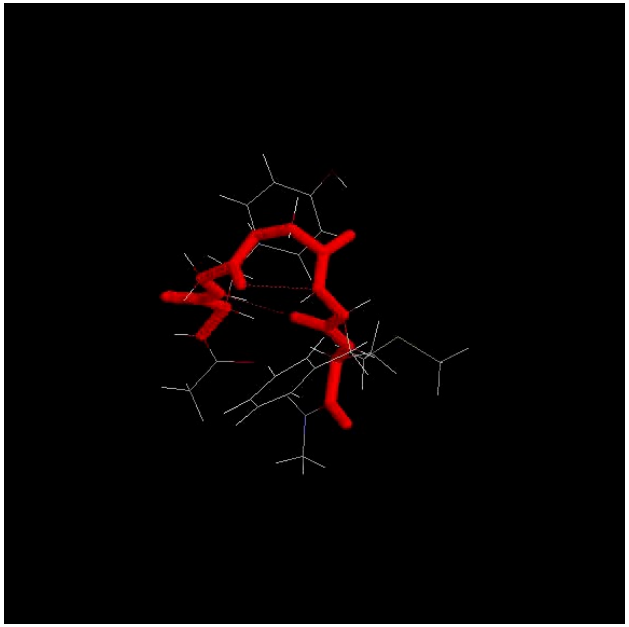


The time series of the dihedral-angle distances in a conventional canonical MD simulation.

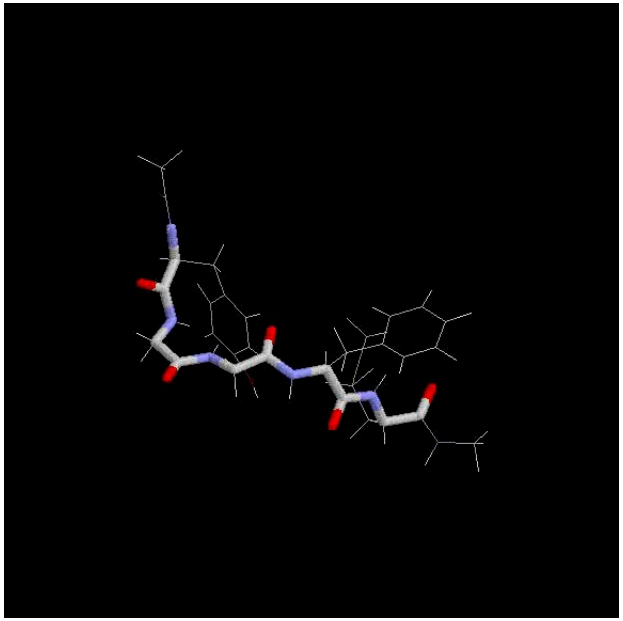




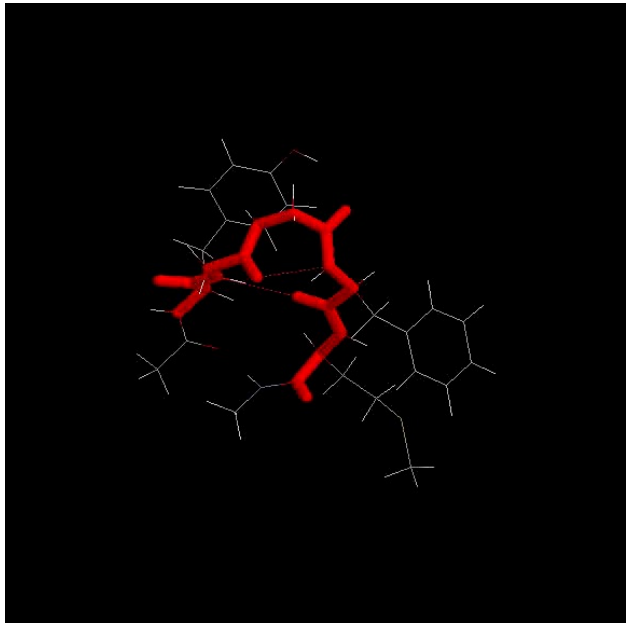
The time series of the dihedral-angle distances in the multi-overlap MD simulation.



Canonical MD



Multicanonical MD



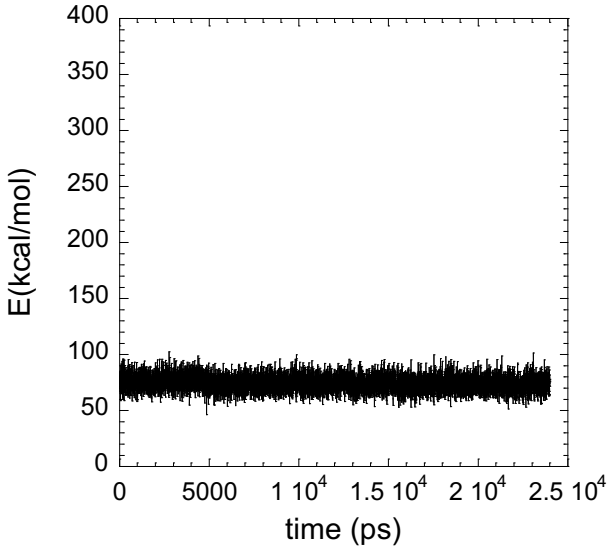
Multioverlap MD

$r_1 < 0.5 \Rightarrow$ 

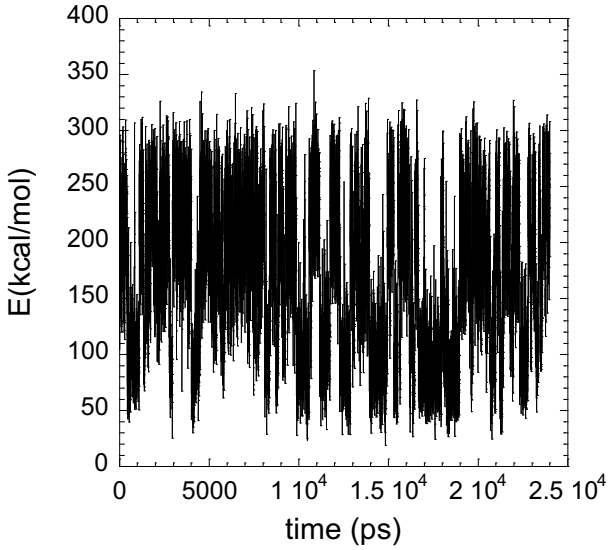
$r_2 < 0.5 \Rightarrow$ 

Simulation and movie by
S.G. Itoh

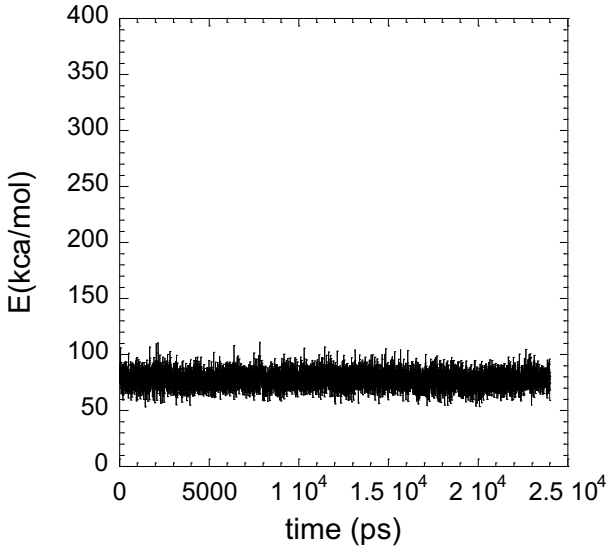
Canonical MD



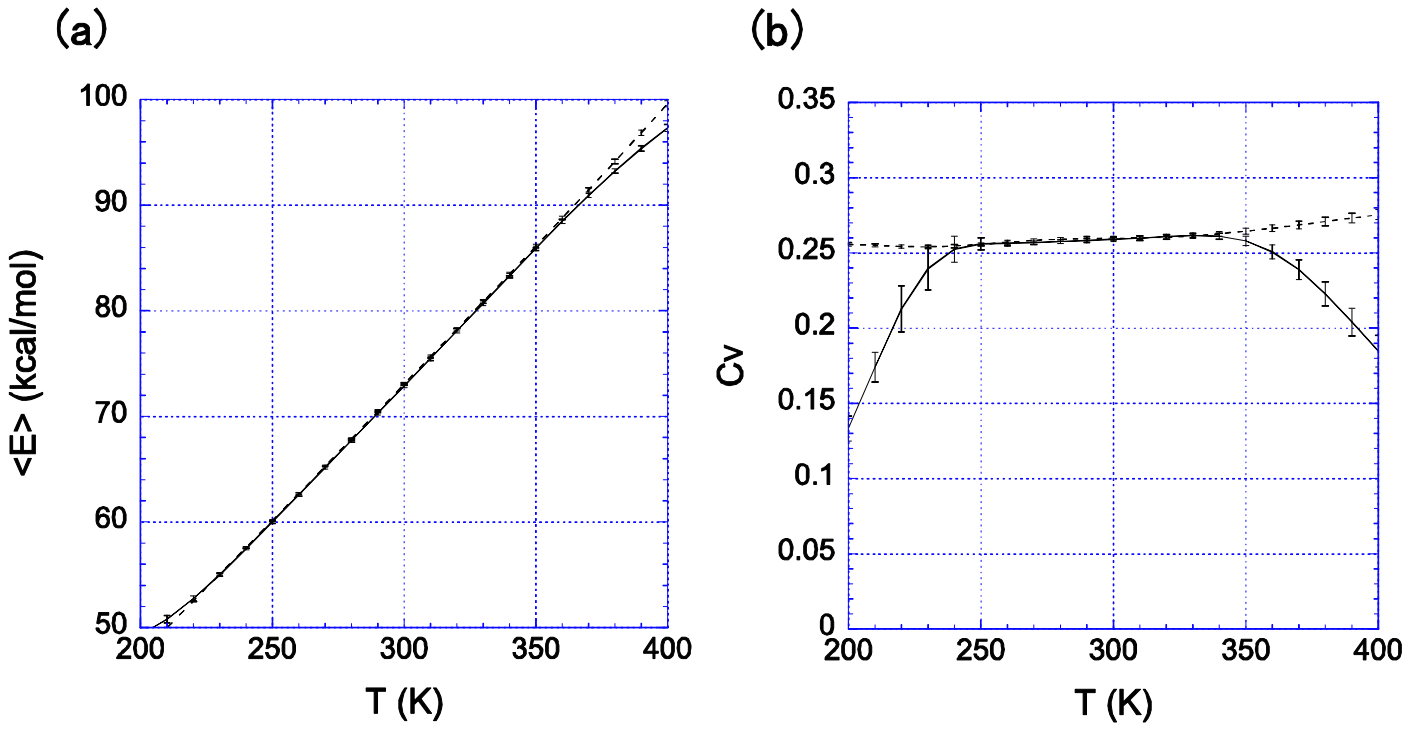
Multicanonical MD



Multioverlap MD



Time series of potential energy

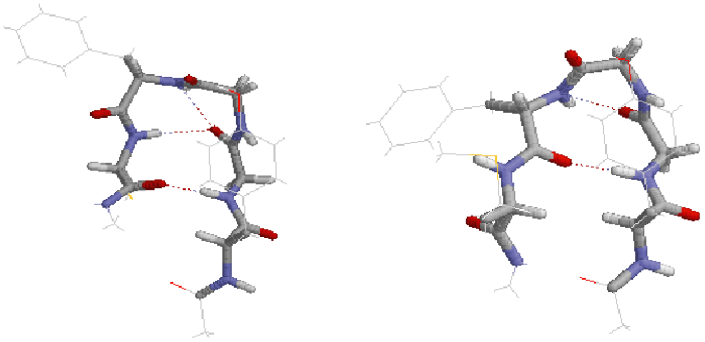


Average potential energy and specific heat as functions of temperature:

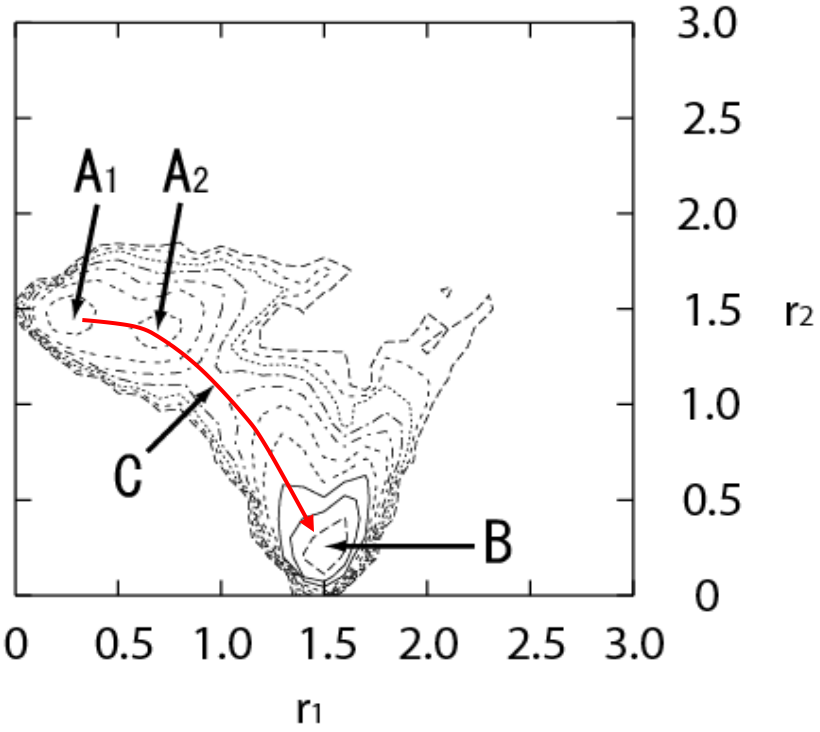
Solid curve: MUOV, Dashed curve: MUCA

MUOV results are reliable only in the range

$\sim 250\ \mathrm{K} < T < \sim 350\ \mathrm{K}$

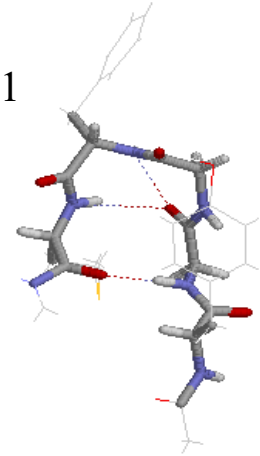


ref. 1 ref. 2

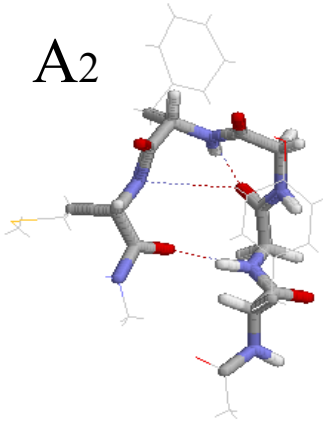


The free-energy landscape obtained from the multi-overlap MD simulation at T=300K with respect to the RMSD axes r_1, r_2.

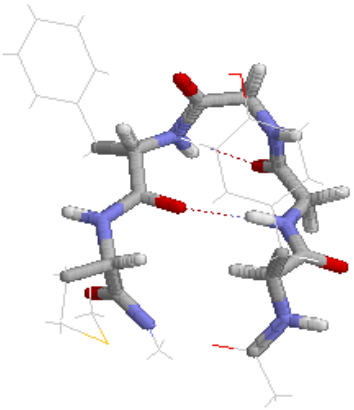
A1



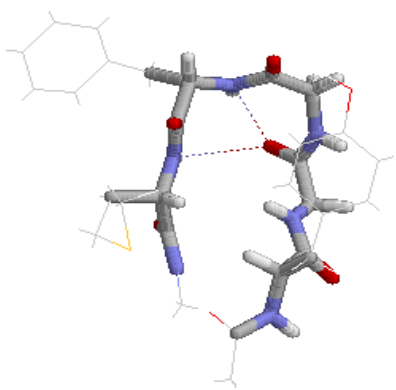
A2



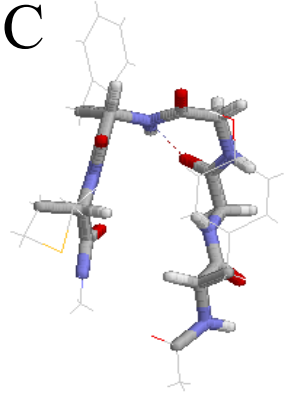
B



C



C



Multicanonical-Multioverlap Algorithm

S.G. Itoh & Y.O., *Phys. Rev. E* **76**, 026705 (2007).

In canonical ensemble at a constant temperature, the probability distribution with variables E, d is given by

$$P_c(E,d) \propto n(E;d) W_c(E;T_0) \equiv n(E;d) \exp(-\beta_0 E),$$

where $n(E;d)$ is the density of states

We employ $E_{mco}(E,d)$ so as to obtain a constant probability distribution with E, d

$$\begin{aligned} P_{mco}(E,d) &= n(E;d) W_{mco}(E;d) \\ &\equiv n(E;d) \exp(-\beta_0 E_{mco}(E,d)) \\ &= const \end{aligned}$$

・再重法

状態密度は次のように表される

$$n(E) = \sum_d N(E;d) \exp(\beta_0 E_{mco}(E,d))$$

ここで $N(E;d)$ はシミュレーションにより得られたヒストグラム.

この状態密度を用いて,物理量 A の温度T での統計平均を次のように得ることができる

$$\langle A \rangle_T = \frac{\sum_E A(E) n(E) e^{-\beta E}}{\sum_E n(E) e^{-\beta E}}$$

・シミュレーション条件

シミュレーション：マルチカノニカル-マルチオーバーラップ
MDシミュレーション
マルチカノニカルMDシミュレーション
マルチオーバーラップMDシミュレーション

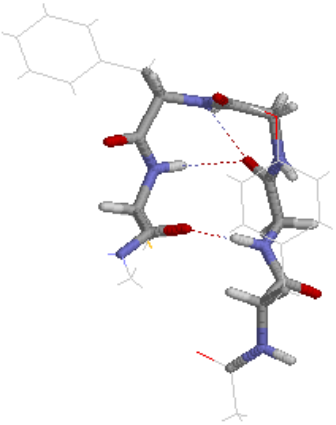
系：真空中のMet-enkephalin (Tyr-Gly-Gly-Phe-Met)
N末はアセチル基,C末はN-メチルアミド基でそれぞれ修飾.

系の温度：300K (ガウス束縛法)

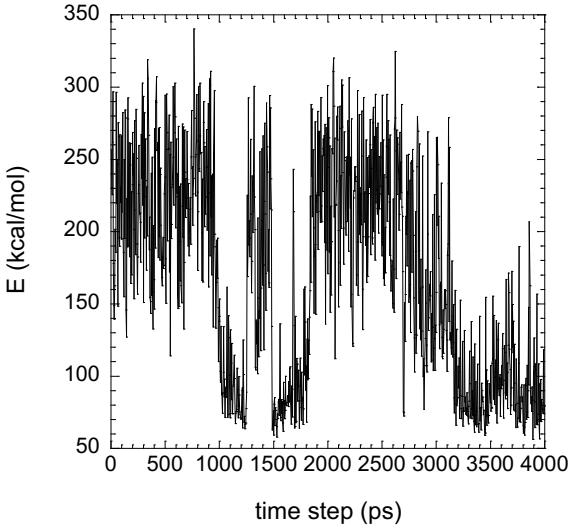
力場：CHARMM param 22 (all-atom model)

時間刻み：0.2fs (leap-frog法)

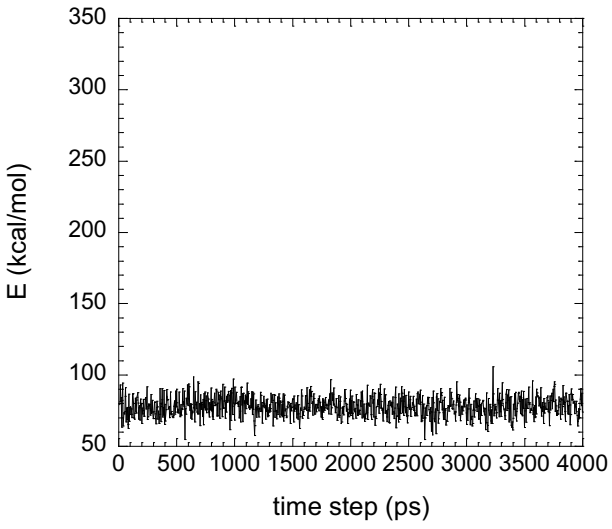
全シミュレーション時間：4ns (+1nsの平衡化)



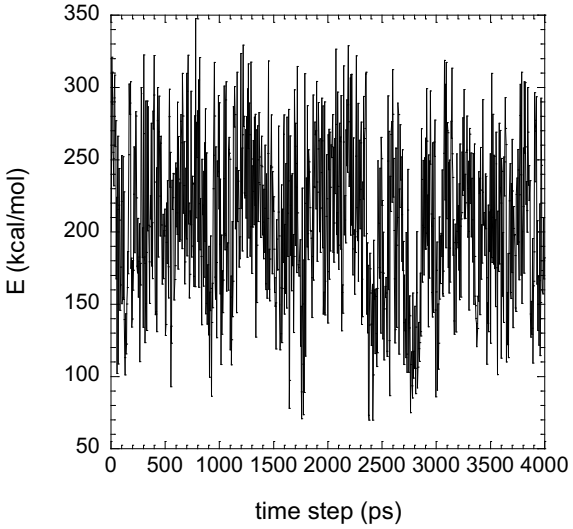
Reference conformation



Multicanonical MD



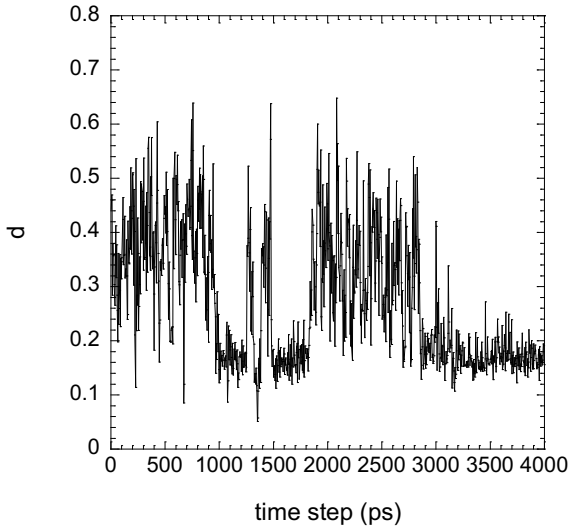
Multioverlap MD



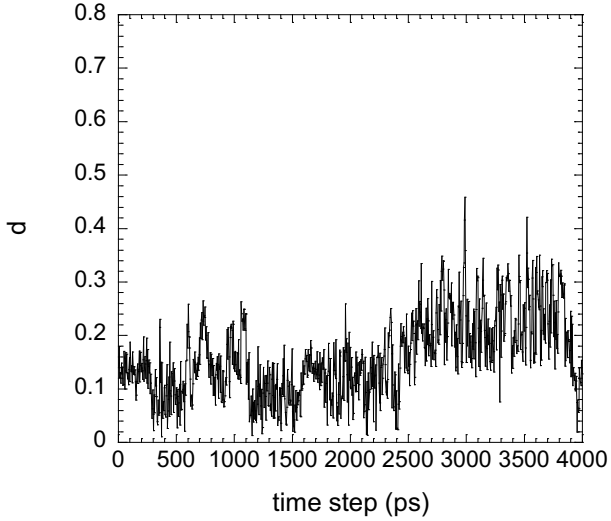
Multicanonical-multioverlap MD

Time series of the potential energy for each simulations

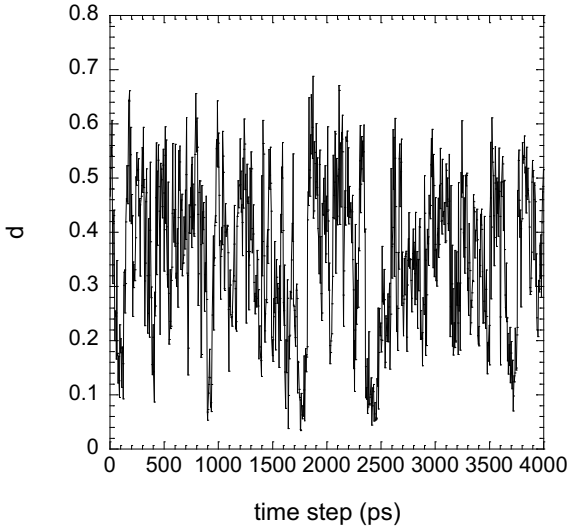
MUCA-MUOV now allows accurate reweighting in a wide temperature range.



Multicanonical MD

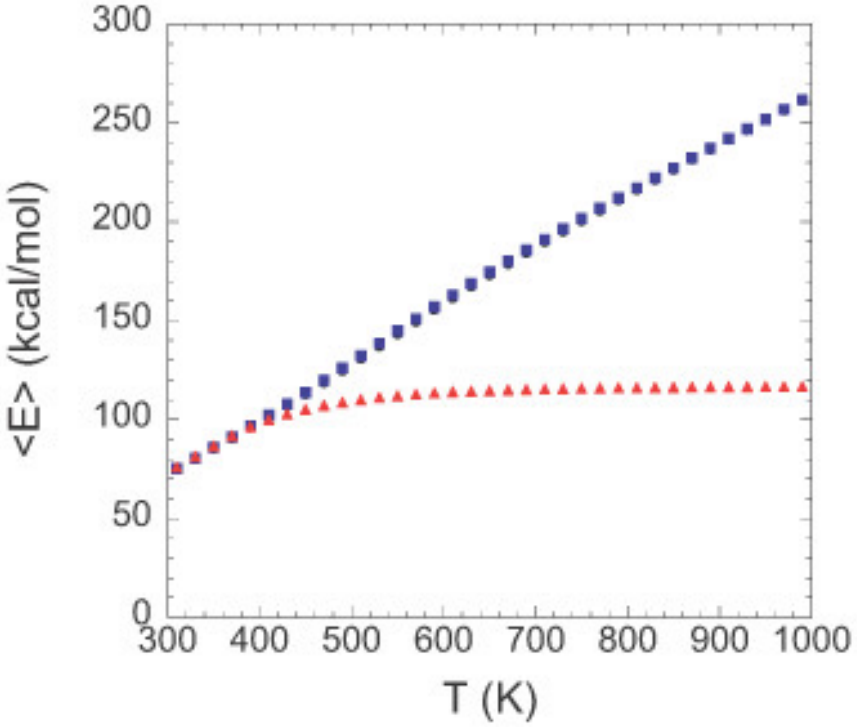


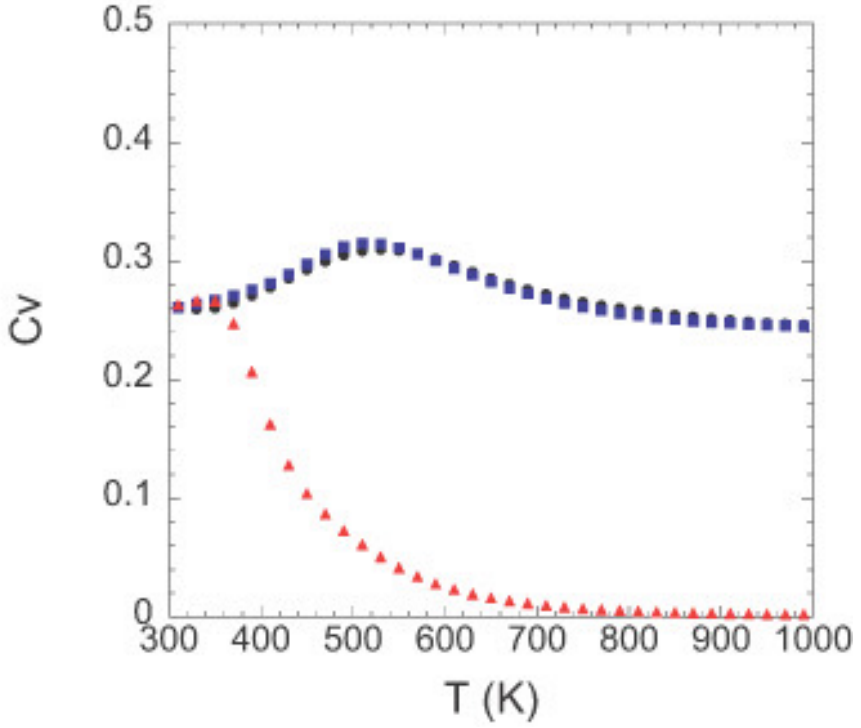
Multioverlap MD



Multicanonical-multioverlap MD

Time series of the dihedral-angle distance for each simulations





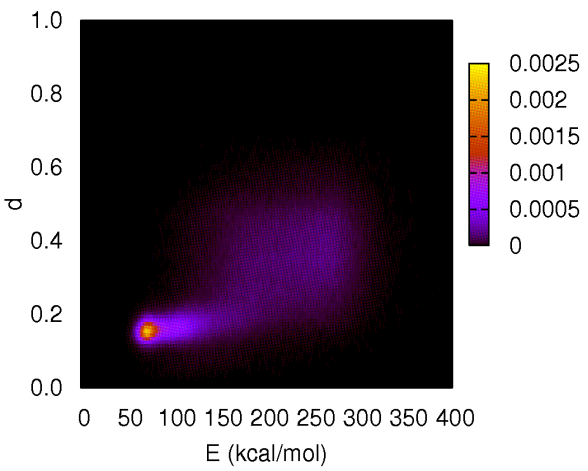
Average potential energy and specific heat as functions of temperature:

Blue and **black** curves: MUCA & **MUCA-MUOV**

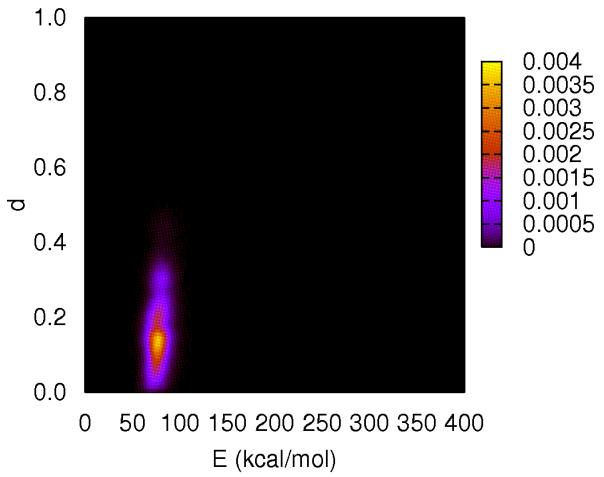
Red curve: MUOV

MUCA and MUCA-MUOV results are reliable in a wide temperature range.

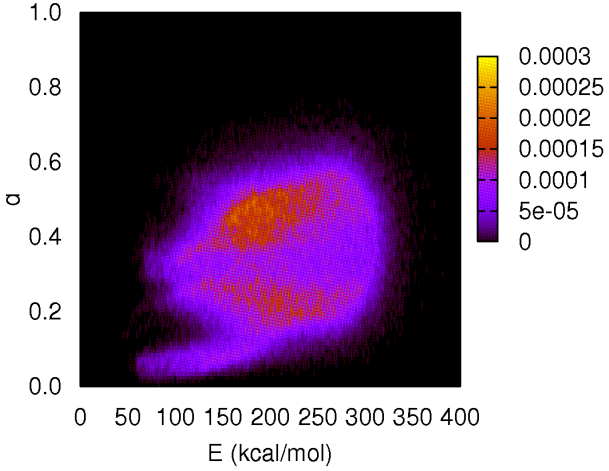
S.G. Itoh & Y.O., *Phys. Rev. E* **76**, 026705 (2007).



Multicanonical MD



Multioverlap MD



Multicanonical-multioverlap MD

Histogram with E and d axes for each simulations

From Multidimensional REM to Multidimensional MUCA and ST

A. Mitsutake & Y.O., *Phys. Rev. E* **79**, 047701 (2009);
J.Chem. Phys. **130**, 214105 (2009);
A. Mitsutake, *J.Chem. Phys.* **131**, 094105 (2009).

MMUCA: random walk in multidimensional energy

MST: random walk in multidimensional parameter

MREM: random walk in multidimensional parameter

e.g., $E_\lambda = E + \lambda V$ $\qquad W_{mu}(E,V) = \dfrac{1}{n(E,V)}$

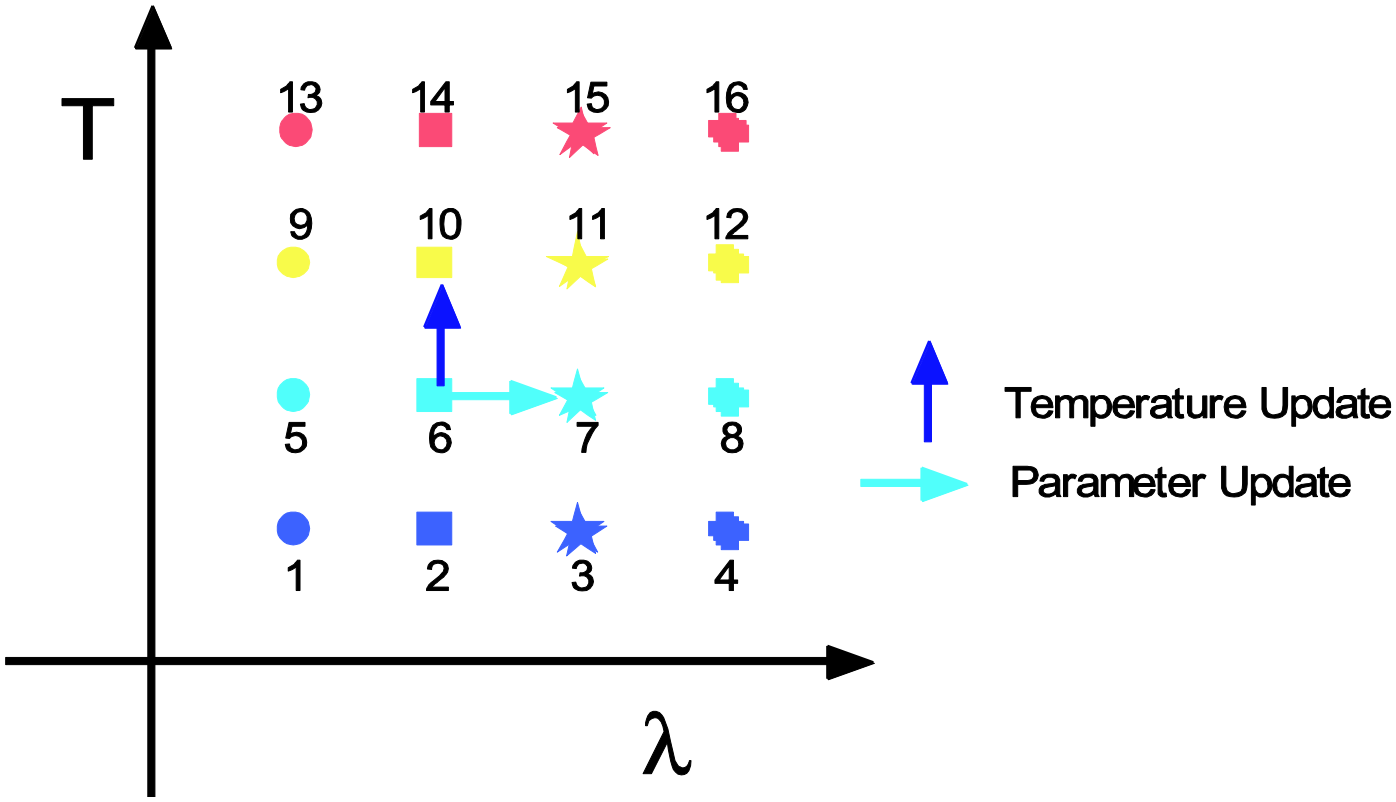
$$W_{ST}(E,V;T_m,\lambda_n) = \exp(-\beta_m(E+\lambda_n V) + f_{m,n})$$

WHAM eqns.

$$n(E,V) = \frac{\sum_{m,n=1}^{M} N_{m,n}(E,V)}{\sum_{m,n=1}^{M} n_{m,n} e^{f_{m,n} - \beta_m(E+\lambda_n V)}}, \text{ where } e^{-f_{m,n}} = \sum_{E,V} n(E,V) e^{-\beta_m(E+\lambda_n V)}.$$

Multidimensional ST

A. Mitsutake & Y.O., *Phys. Rev. E* **79**, 047701 (2009);
J.Chem. Phys. **130**, 214105 (2009);
A. Mitsutake, *J.Chem. Phys.* **131**, 094105 (2009).



Examples of Multidimensional REM, MUCA, and ST

A. Mitsutake & Y.O., *Phys. Rev. E* **79**, 047701 (2009);
J. Chem. Phys. **130**, 214105 (2009);
A. Mitsutake, *J. Chem. Phys.* **131**, 094105 (2009).

$$E_\lambda = E + \lambda V$$

$\lambda = h =$ external field
$V = M =$ magnetization

1. Simulated Tempering and Magnetizing

random walk in temperature *T* and external field *h*

∗ Ising Model

T. Nagai & Y.O., *Phys. Rev. E* **86**, 056705 (2012).

$$H = E - hM$$

$$E = -\sum_{\langle i,j \rangle} \sigma_i \sigma_j$$

$$M = \sum_{i}^{N} \sigma_i \,,$$

∗ 3-state Potts Model

T. Nagai, Y.O., & W. Janke,
J. Stat. Mech. (2013) P02039.

$$H = E - hM \,,$$

$$E = -\sum_{\langle i,j \rangle} \delta_{\sigma_i,\sigma_j} \,,$$

$$M = \sum_{i=1}^{N} \delta_{0,\sigma_i} \,,$$

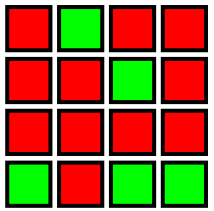
T. Nagai & Y.O., *Physics Procedia* **34**, 100 (2012).
T. Nagai & Y.O., *Phys. Rev. E* **86**, 056705 (2012).

Application of Two-dimensional ST method to Two-dimensional Ising model and Its Crossover

System and method

Hamiltonian: $H = E - hM$,

where $E = -\sum_{<i,j>} \sigma_i \sigma_j$, $M = \sum_i \sigma_i$

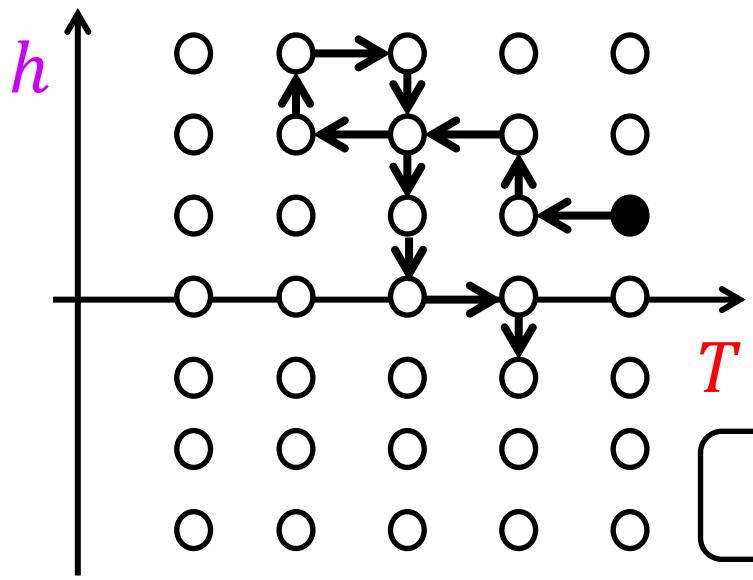


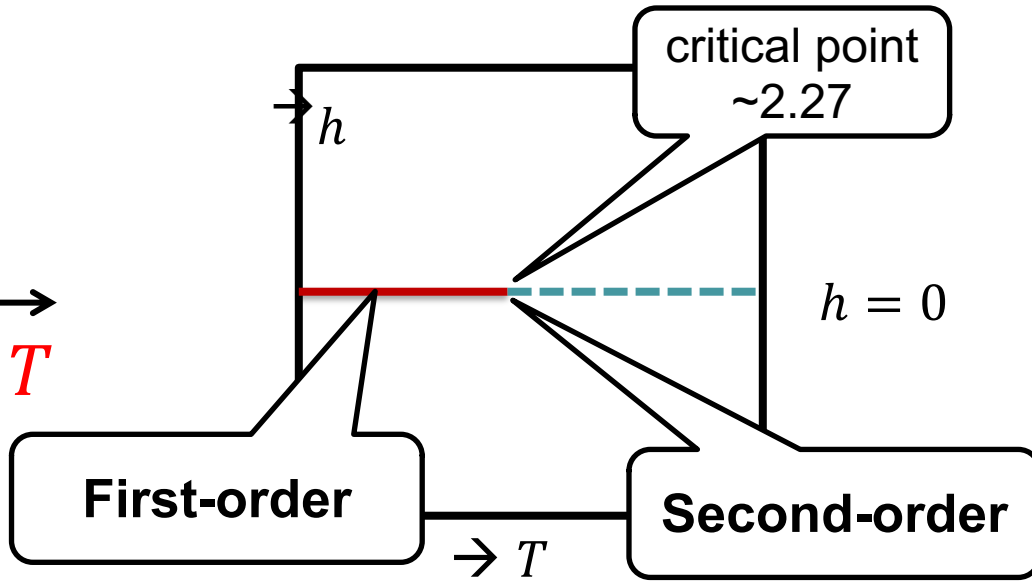
Two-dimensional Ising model (4X4)

Weight: $w(\{\sigma_i\}, T, h) = e^{-\beta(E - hM) + a(T,h)}$

“Simulated Tempering and Magnetizing” (STM).

（Both temperature *T* and external field *h* are updated）





Two-dimensional simulated tempering

- Consider the Boltzmann factor $e^{-\beta(E(x)-hM)+a(T,h)}$ as the disjoint probability of (x, T, h), where $T \in \{T_1, T_2, \ldots, T_{N_T}\}$ and $h \in \{h_1, h_2, \ldots, h_{N_h}\}$, respectively. Here, $a(T, h)$ is a parameter.
- The parameters are updated by the Metropolis criteria: $P_{accept} = \min(1, \Delta)$.

$$\Delta = e^{-\left(\beta^{trial}-\beta^{old}\right)E(x)+\left(\beta^{trial}h^{trial}-\beta^{old}h^{old}\right)M(x)+a(T^{trial},h^{trial})-a(T^{old},h^{old})}$$

Ref) A. Mitsutake and Y.O.: *J. Chem. Phys.* **130**, 214105 (2009).

Reweighting methods

The WHAM or MBAR method enable one to obtain the density function and free energy of systems.

- This gives free energy necessary for the ST methods.
- Knowing the density function let one calculate thermal average at any point.

WHAM

$$n(E,M) = \frac{\sum_{T_i,h_j} n_{T_i,h_j}(E,M)}{\sum_{T_i,h_j} N_{T_i,h_j} \exp(f(T_i,h_j) - (E - h_j M)/T_i)}$$

$$f(T_i,h_j) = -\log \sum_{E,M} n(E,M) \exp(-(E - h_j M)/T_i)\,,$$

S. Kumar, et al.: *J. Comput. Chem.* **13**, 1011 (1992).
A. Mitsutake and Y.O.: *J. Chem. Phys.* **130**, 214105 (2009).

Reweighting methods

MBAR

$$f(T_i, h_j) = -\log \sum_{i=1}^{N_T} \sum_{j=1}^{N_h} \sum_{n=1}^{N} \frac{\exp(-(E_n - h_j M_n)/T_i)}{\sum_{k=1}^{N_T} \sum_{l=1}^{N_h} N_{kl} \exp(f(T_k, h_l) - (E_n - h_l M_n)/T_k)}$$

$$\langle A \rangle_{T,h} = \sum_{n=1}^{N} W_{na} A(x_n)$$

$$W_{na} = \frac{1}{\langle c_a \rangle} \frac{\exp(-(E_n - h M_n)/T)}{\sum_{k=1}^{N_T} \sum_{l=1}^{N_h} N_{kl} \exp(f(T_k, h_l) - (E_n - h_l M_n)/T_k)}$$

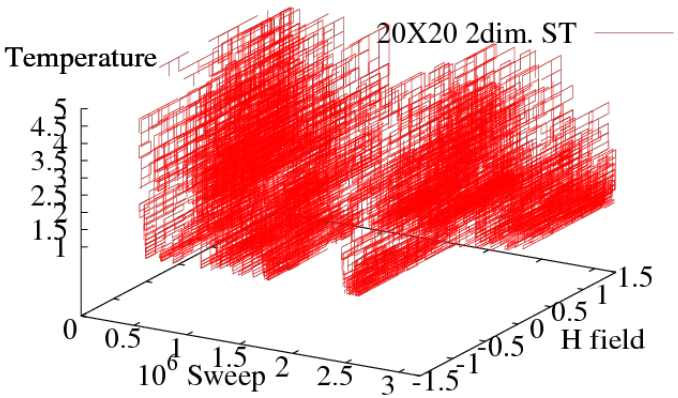
$$\langle c_a \rangle = \sum_{n=1}^{N} \frac{\exp(-(E_n - h M_n)/T)}{\sum_{k=1}^{N_T} \sum_{l=1}^{N_h} N_{kl} \exp(f(T_k, h_l) - (E_n - h_l M_n)/T_k)}$$

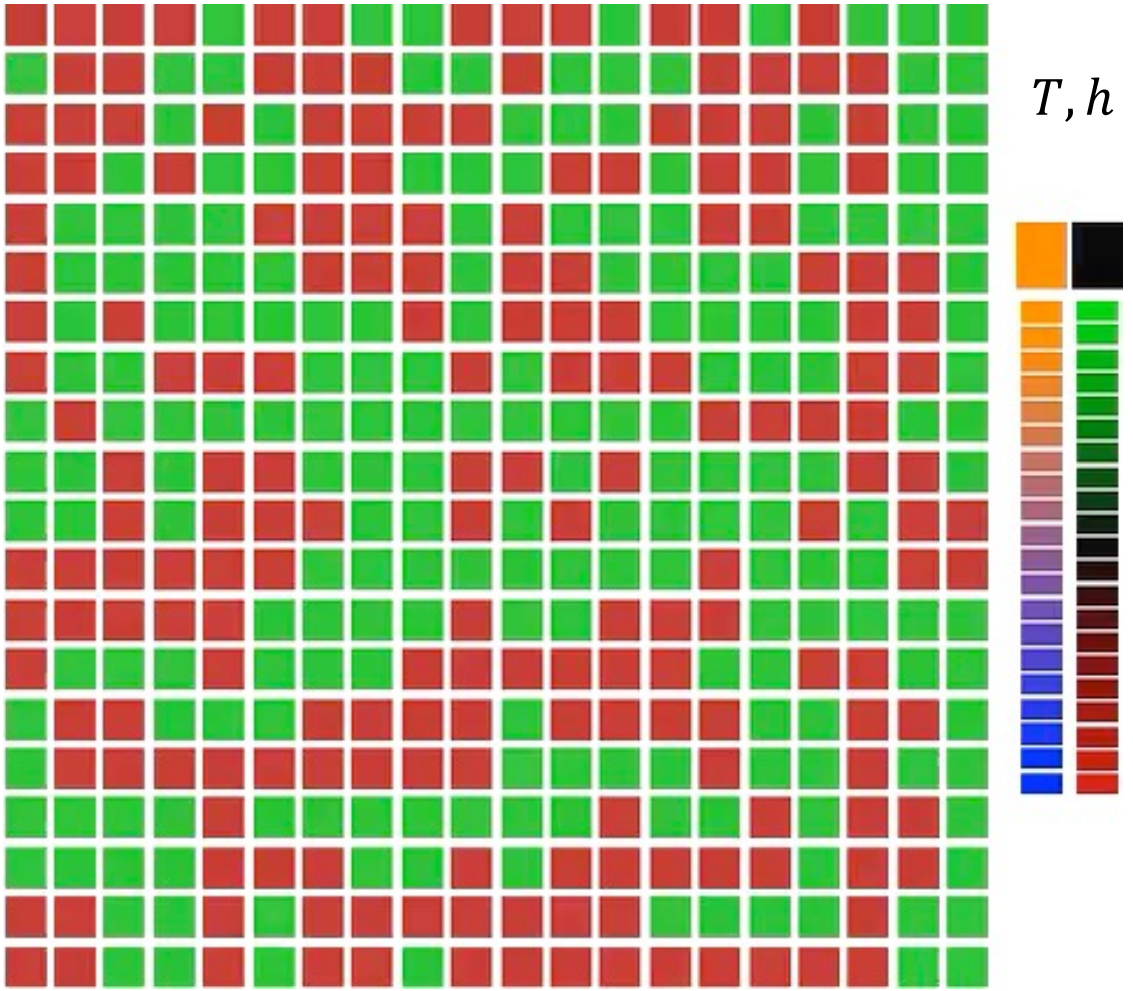
M. Shirts and J. Chodera: *J. Chem. Phys.* **129**, 124105 (2008).

Simulation and movie by T. Nagai

Results and Discussion

- In this simulation, the temperature and external field realized the random walks. Therefore the simulations are performed properly.





緑：σ=+1
赤：σ=-1

Temperature: 20 points in [1:5]
External field: 21 points in [-1.5, 1.5]
System: 20X20

T. Nagai & Y.O., *Phys. Rev. E* **86**, 056705 (2012).

Results and Discussion

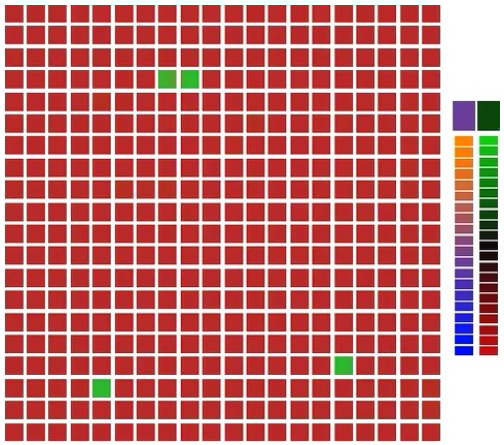
- To investigate the effect of multi-dimensional generalization,
  - We performed “Simulated Magnetizing (SM),” where T is fixed and h is allowed to change.

T. Nagai & Y.O., *Phys. Rev. E* **86**, 056705 (2012).

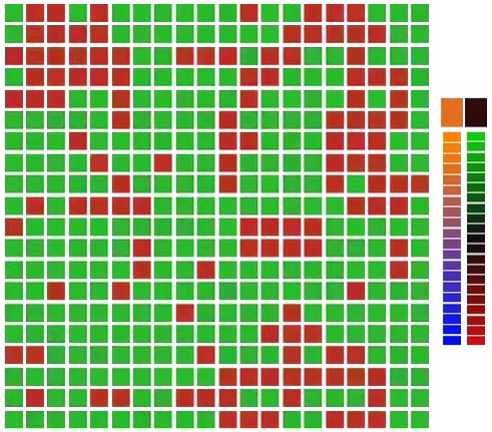
Simulation and movie by T. Nagai

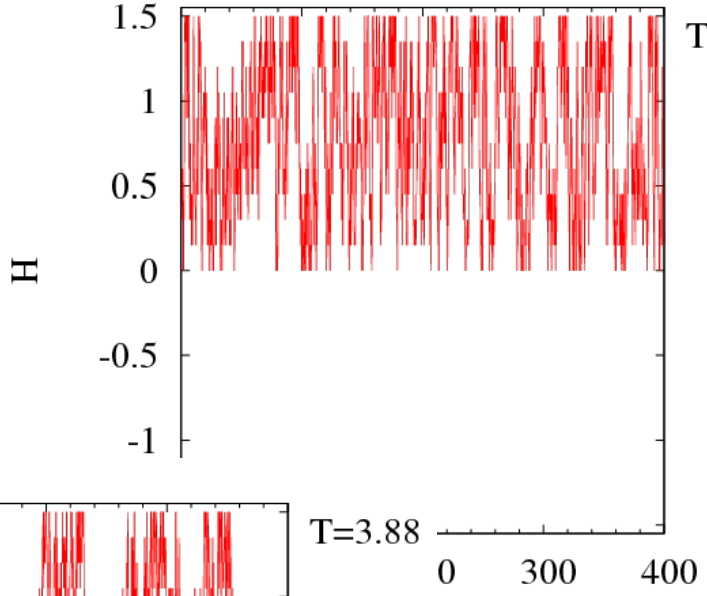
Results and Discussion

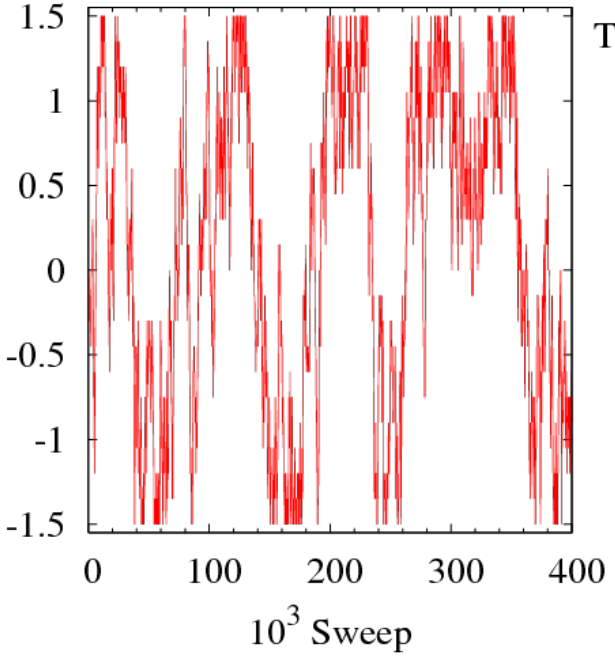
T is fixed at 1.97 (below T_c)



T is fixed at 3.88 (above T_c)





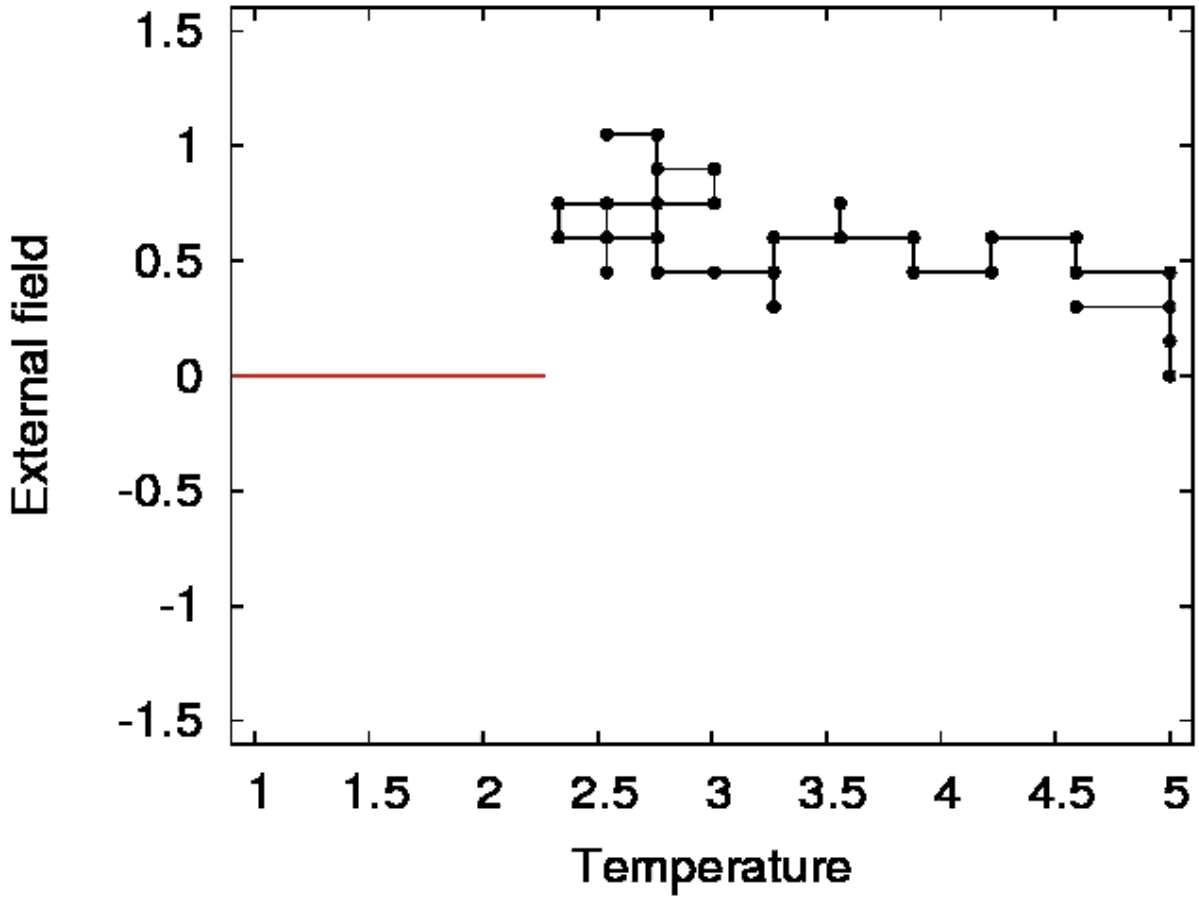


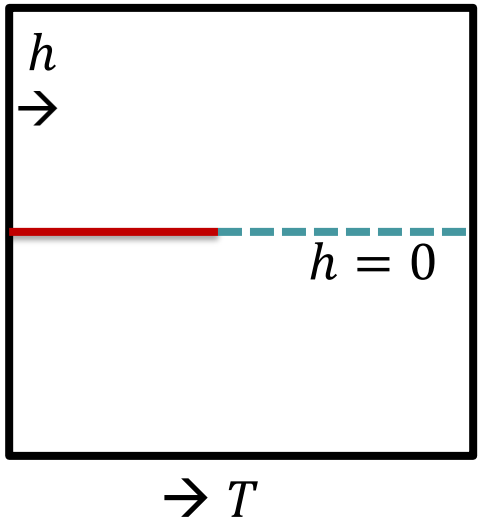
Below the T_c, H changes in only one side. Do not go over 0.

Simulation and movie by T. Nagai

Results and Discussion

This animation illustrates how the temperature and external field moved during the STM simulation.





At every frame,
5000 MC sweeps.
Change by 1000
MC sweeps.

ST-Freq: 50
TvsH=0.5
L=20

Results and Discussion

T. Nagai & Y.O., *Phys. Rev. E* **86**, 056705 (2012).

M as a function of *Lt* when *h*=0.

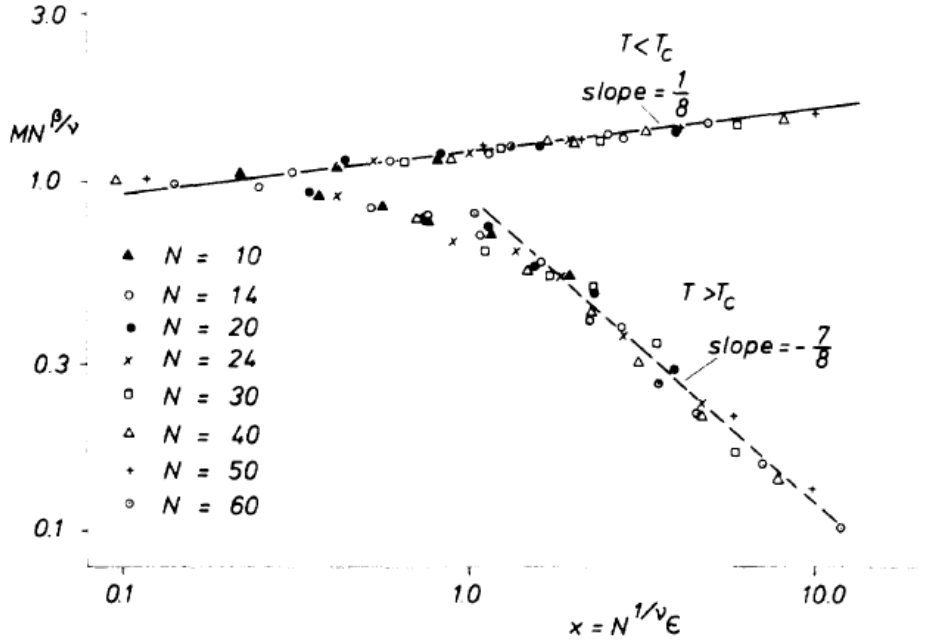
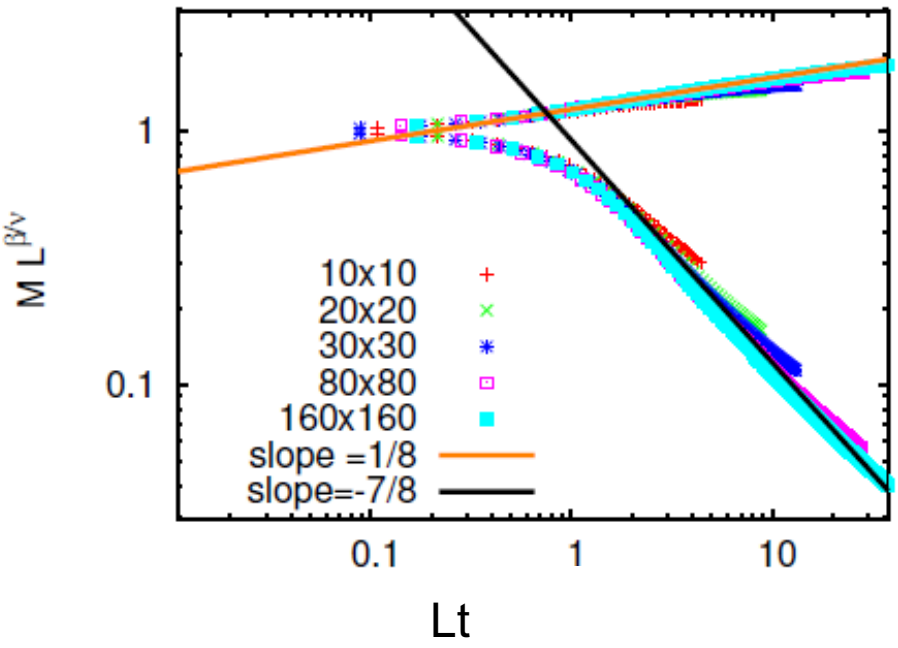


FIG. 14. Finite-size scaling plot for the order parameter for lattices with p.b.c. Data for $N=10$, ▲; $N=14$, ○; $N=20$, ●; $N=24$, ×; $N=30$, □; $N=40$, △; $N=50$, +; $N=60$, ⊙. For $T<T_c$ the solid line is $1.22\,x^{1/8}$; for $T>T_c$ the dashed line is $0.92\,x^{-7/8}$.



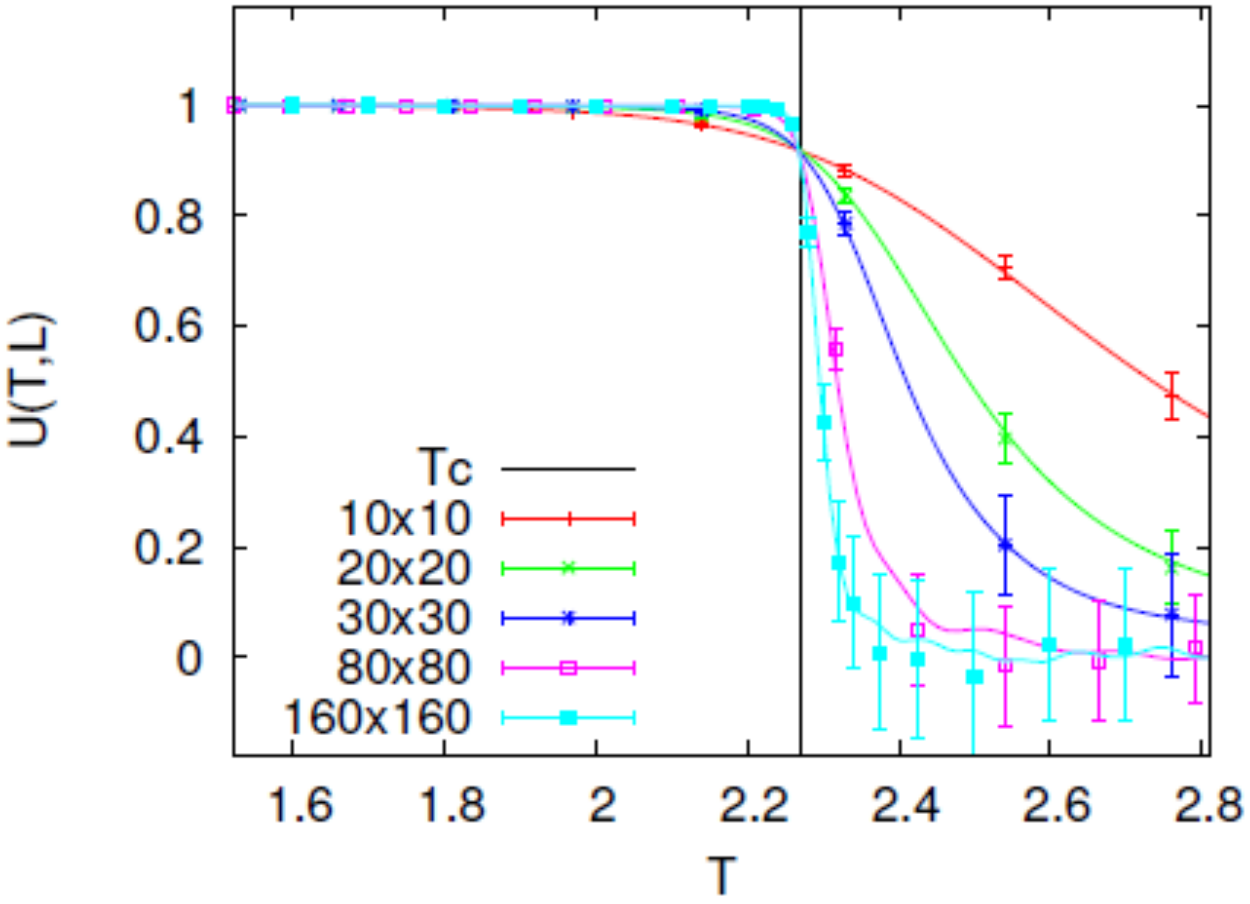
We applied the finite-size scaling form.
Here, $M \sim t^{1/8}$

Ref) D.P. Landau: Phys. Rev. B **13**, 2997 (1976).

Results and Discussion

T. Nagai & Y.O., *Phys. Rev. E* **86**, 056705 (2012).

The Binder cumulant for $h = 0$.
The curves were obtained with the reweighting method.



$$U(T,h,L) \equiv \frac{1}{2}\left(3 - \frac{\langle m^4 \rangle}{\langle m^2 \rangle^2}\right)$$

K. Binder, *Z. f. Physik B* **43**, 119 (1981).

Examples of Multidimensional REM, MUCA, and ST

A. Mitsutake & Y.O., *Phys. Rev. E* **79**, 047701 (2009);
J. Chem. Phys. **130**, 214105 (2009);
A. Mitsutake, *J. Chem. Phys.* **131**, 094105 (2009).

$$E_\lambda = E + \lambda V$$

2. Grandcanonical Ensemble

$\lambda = -\mu =$ - chemical potential $\quad V = N =$ no. of particles

* **REM: random walk in**
temperature *T* and chemical potential *μ*

D. Matsubara & Y.O., *J. Chem. Phys.* **152**, 194108 (2020).

Lennard-Jones Fluid System:

$$E(s) = \sum_{i=2}^{N} \sum_{j<i}^{N-1} 4\varepsilon \left(\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right) \longrightarrow E^*(s) = \sum_{i=2}^{N} \sum_{j<i}^{N-1} 4 \left(\left(\frac{1}{r_{ij}^*} \right)^{12} - \left(\frac{1}{r_{ij}^*} \right)^{6} \right)$$

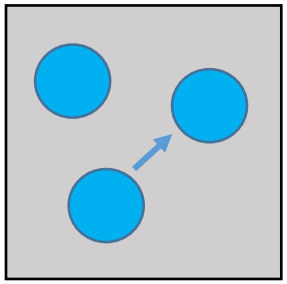
Reduced unit: $r^* = r/\sigma$, $E^* = E/\varepsilon$, $\mu^* = \mu/\varepsilon$, $T^* = k_B T/\varepsilon$, $V^* = L^{*3}$, $L^* = L/\sigma$, $\rho^* = \rho\sigma^3 = \langle N \rangle / V^*$, $V^* = L^{*3}$

Grandcanonical MC (GCMC)

メトロポリス判定 Metropolis Criterion

$$\omega\left(x^{(\nu)} \rightarrow x^{(\nu+1)}\right) = \min\left(1, \frac{P_{\mu VT}\left(x^{(\nu+1)}\right)}{P_{\mu VT}\left(x^{(\nu)}\right)}\right)$$

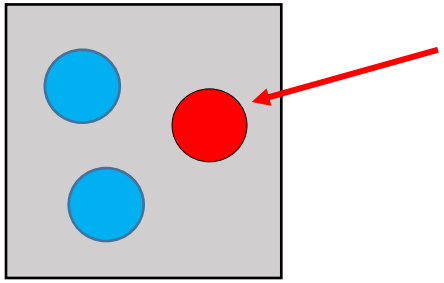
Translation:



$$x^{(\nu)} = \{s_N, N\} \qquad x^{(\nu+1)} = \{s_N + \Delta s_N, N\}$$

$$\frac{P_{\mu VT}\left(x^{(\nu+1)}\right)}{P_{\mu VT}\left(x^{(\nu)}\right)} = \exp\{-\beta[E(s_N + \Delta s_N) - E(s_N)]\}$$

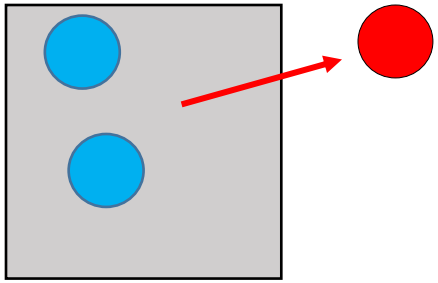
Insert:



$$x^{(\nu)} = \{s_N, N\} \qquad x^{(\nu+1)} = \{s_{N+1}, N+1\}$$

$$\frac{P_{\mu VT}\left(x^{(\nu+1)}\right)}{P_{\mu VT}\left(x^{(\nu)}\right)} = \frac{V}{(N+1)} \exp\{-\beta[E(s_{N+1}) - E(s_N)] + \beta\mu\}$$

Delete:



$$x^{(\nu)} = \{s_N, N\} \qquad x^{(\nu+1)} = \{s_{N-1}, N-1\}$$

$$\frac{P_{\mu VT}\left(x^{(\nu+1)}\right)}{P_{\mu VT}\left(x^{(\nu)}\right)} = \frac{N}{V} \exp\{-\beta[E(s_N) - E(s_N)] - \beta\mu\}$$

Grandcanonical REM (GCREM)

D. Matsubara & Y.O., *J. Chem. Phys.* **152**, 194108 (2020).

M Replicas with

$$\lambda_m = \lambda_{m_0,m_1} \equiv (T_{m_0}, \mu_{m_1}),$$
$$m = (m_0, m_1),$$
$$m_0 = 1, 2, \ldots, M_0,$$
$$m_1 = 1, 2, \ldots, M_1,$$
$$M = M_0 \times M_1,$$

$$T_{m_0} : T_1, T_2, \cdots, T_{M_0}$$

$$\mu_{m_1} : \mu_1, \mu_2, \cdots, \mu_{M_1}$$

Step 1: Each replica i performs an independent grand canonical MC or MD simulation ($i = 1, \ldots, M$) for a certain MC or MD steps.

Step 2: A pair of replicas, $x_m^{[i]}$ and $x_n^{[j]}$, are exchanged with the probability $w\left(x_m^{[i]} \mid x_n^{[j]}\right)$ in Eq. (20).

$$w(X \to X') \equiv w\left(x_m^{[i]} \mid x_n^{[j]}\right) = \min\left(1, \frac{W_{GCREM}(X')}{W_{GCREM}(X)}\right)$$
$$= \min(1, \exp(-\Delta)). \qquad (20)$$

$$\Delta = \Delta_T = -(\beta_{n_0} - \beta_{m_0})\left\{\left(E^{[j]} - E^{[i]}\right) - \mu\left(N^{[j]} - N^{[i]}\right)\right\} \quad \text{for } T\text{-exchange}$$

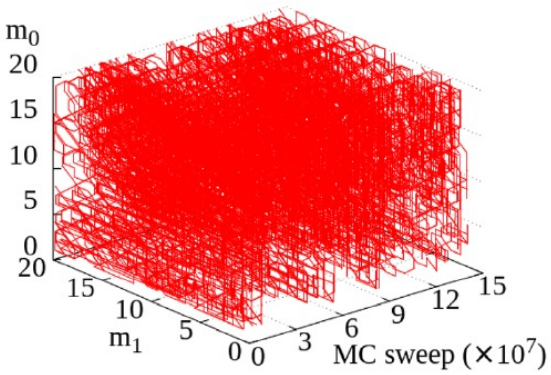
$$\Delta = \Delta_\mu = \beta(\mu_{n_1} - \mu_{m_1})\left(N^{[j]} - N^{[i]}\right) \quad \text{for } \mu\text{-exchange}$$

Repeat these two steps.

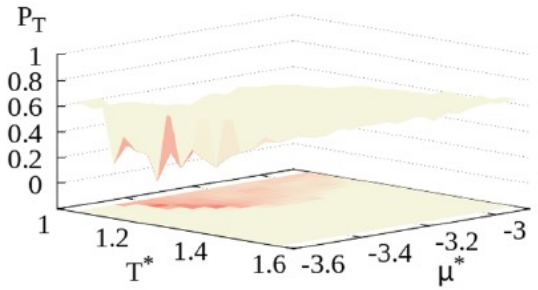
RESULTS

Grandcanonical REM (GCREM)

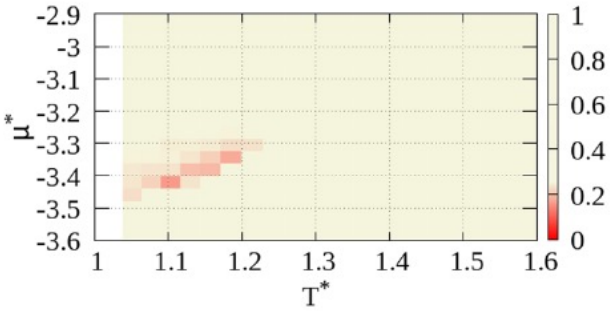
D. Matsubara & Y.O., *J. Chem. Phys.* **152**, 194108 (2020).



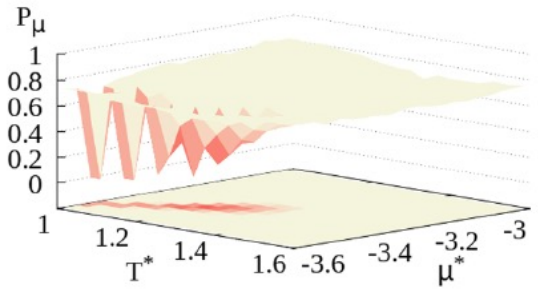
Random walk in T-μ space



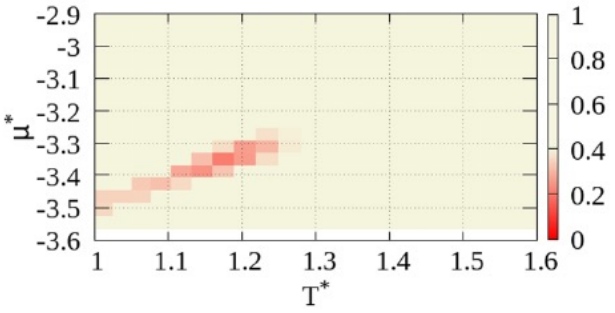
(a) 3D plot of T-exchange



(b) 2D projection color map of T-exchange



(c) 3D plot of μ-exchange



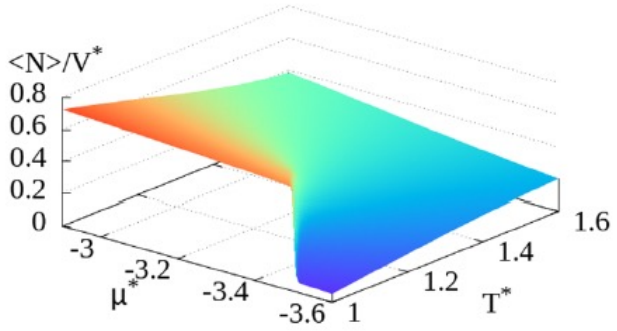
(d) 2D projection color map of μ-exchange

Acceptance ratios of T-exchange (a,b) and for μ-exchange (c,d)

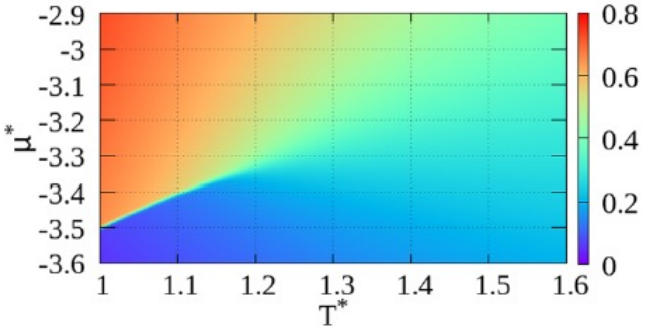
Grandcanonical REM (GCREM)

RESULTS

D. Matsubara & Y.O., *J. Chem. Phys.* **152**, 194108 (2020).

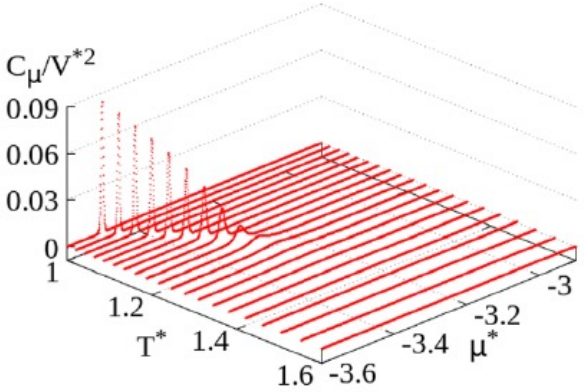


(a) 3D plot of $\langle N \rangle$

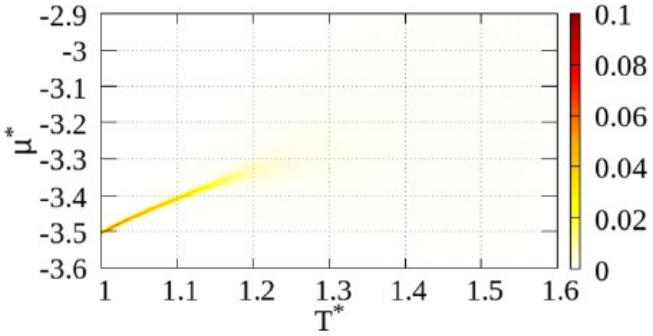


(b) 2D projection color map of $\langle N \rangle$

Average number of particles <*N*>



(a) 3D plot of C_μ



(b) 2D projection color map of C_μ

Particle number fluctuations $C\mu$

$$C_\mu \equiv \langle N^2 \rangle - \langle N \rangle^2$$

Grandcanonical REM (GCREM)

RESULTS

D. Matsubara & Y.O., *J. Chem. Phys.* **152**, 194108 (2020).

Binder Cumulant

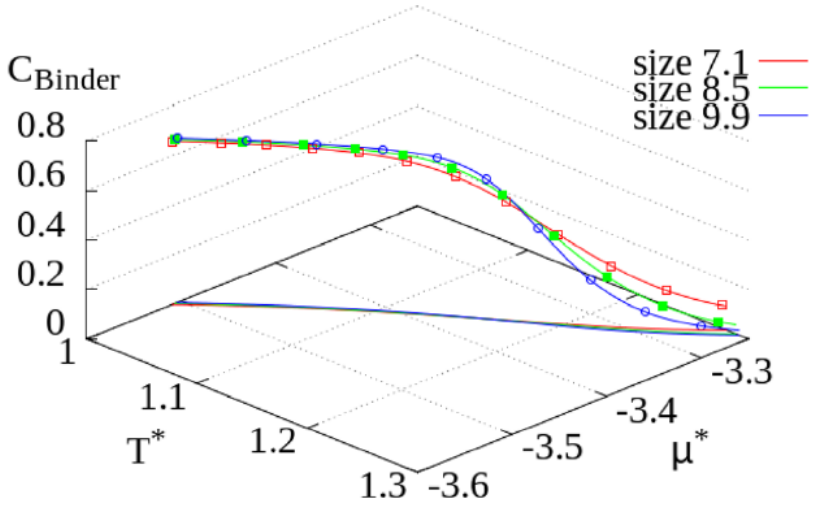
$$C_{Binder} = 1 - \frac{\langle m^4 \rangle}{3\langle m^2 \rangle^2}$$

where the number density is used for m:

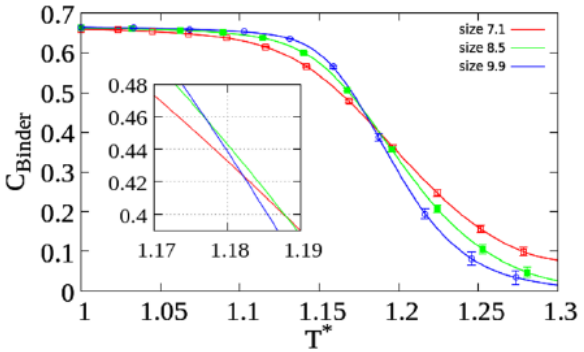
$$m = \rho^* = \frac{< N >}{V^*}$$

These results give:

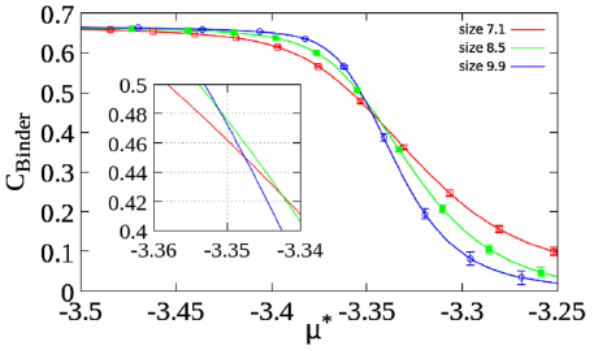
$$T_c^* = 1.183 \pm 0.005$$
$$\mu_c^* = -3.346 \pm 0.004$$
$$\rho_c^* = 0.336 \pm 0.006$$



(a) C_{Binder} in 3D space and its projection onto the (T^*, μ^*) plane.



(b) C_{Binder} projected on T^* axis



(c) C_{Binder} projected on μ^* axis

Examples of Multidimensional REM, MUCA, and ST

A. Mitsutake & Y.O., *Phys. Rev. E* **79**, 047701 (2009);
J. Chem. Phys. **130**, 214105 (2009);
A. Mitsutake, *J. Chem. Phys.* **131**, 094105 (2009).

$$E_\lambda = E + \lambda V$$

3. Isobaric-Isothermal Ensemble（定圧定温アンサンブル）

$$\lambda = P = \text{pressure} \qquad V = \text{volume}$$

* **MUCA: Multibaric-Multithermal Algorithm (MUBATH)**
  random walk in potential energy *E* and volume *V*

H. Okumura & Y.O., *Chem. Phys. Lett.* **383**, 391 (2004). (MC version)
H. Okumura & Y.O., *Chem. Phys. Lett.* **391**, 248 (2004). (MD version)

* **REM: random walk in temperature *T* and pressure *P***

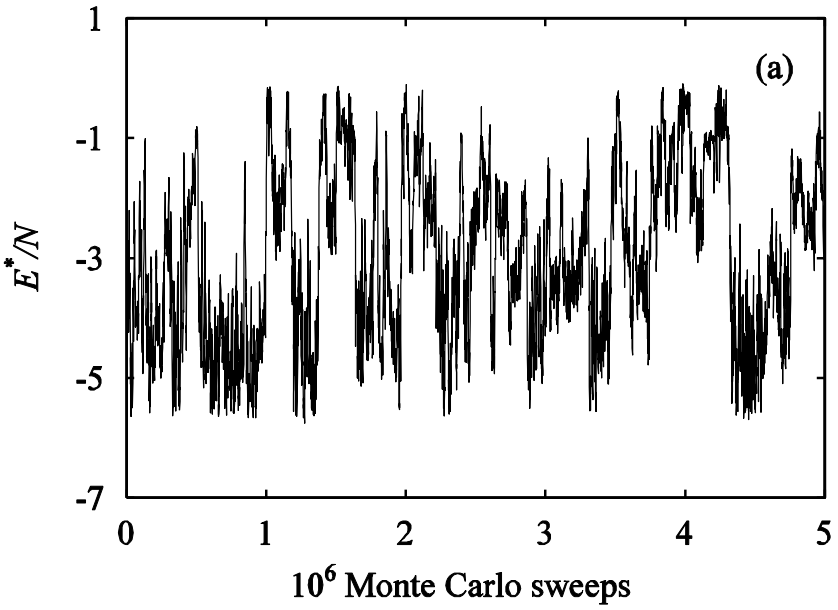
Y. Sugita & Y.O., in *Lect. Notes in Computational Science & Engineering*, ed. by T. Schlick and H.Gun (2002) pp. 304-332; cond-mat/0102296.
T. Okabe, M. Kawata, Y.O. & M. Mikami, *Chem. Phys. Lett.* **335**, 435 (2001).
Also, see D. Paschek & A. Garcia, *Phys. Rev. Lett.* **93**, 238105 (2004).

* **ST: random walk in temperature *T* and pressure *P***

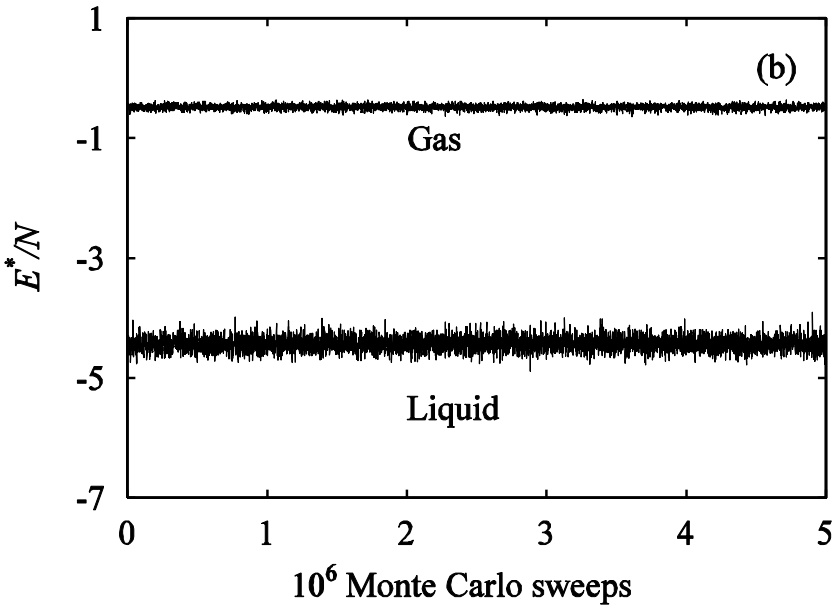
Y. Mori & Y.O., *J. Phys. Soc. Jpn.* **79**, 074003 (2010).

Liquid-Gas Phase Transitions Studied by Multibaric-Multithermal Simulations

MUBATH



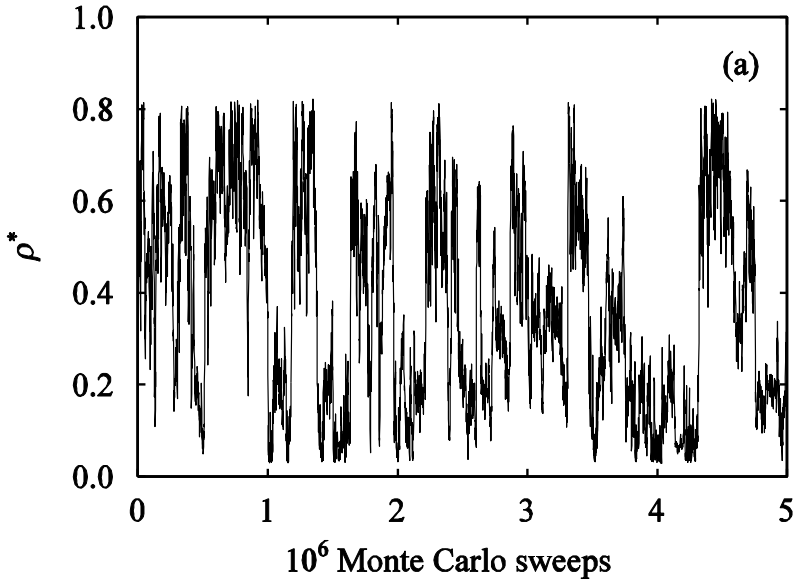
Conventional NPT



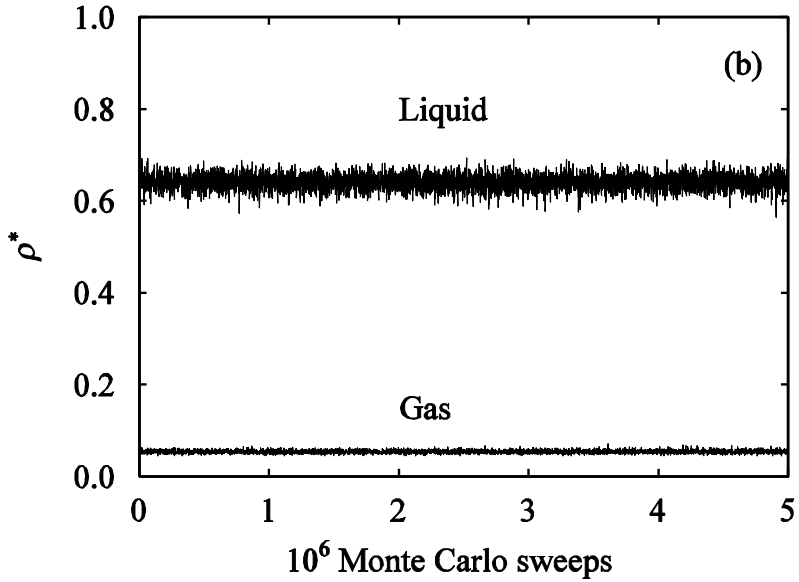
H. Okumura & Y.O., *J. Phys. Soc. Jpn.* **73**, 3304 (2004).

Liquid-Gas Phase Transitions Studied by Multibaric-Multithermal Simulations

MUBATH

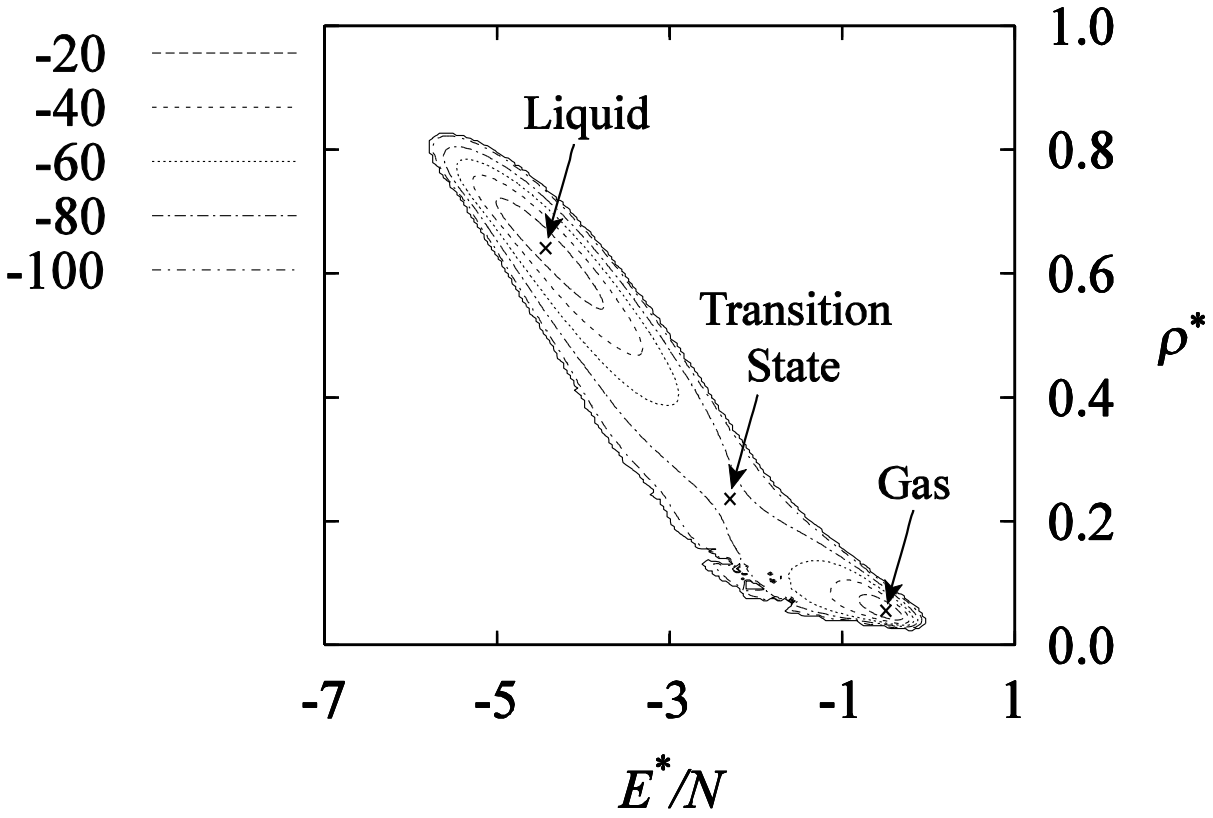


Conventional NPT



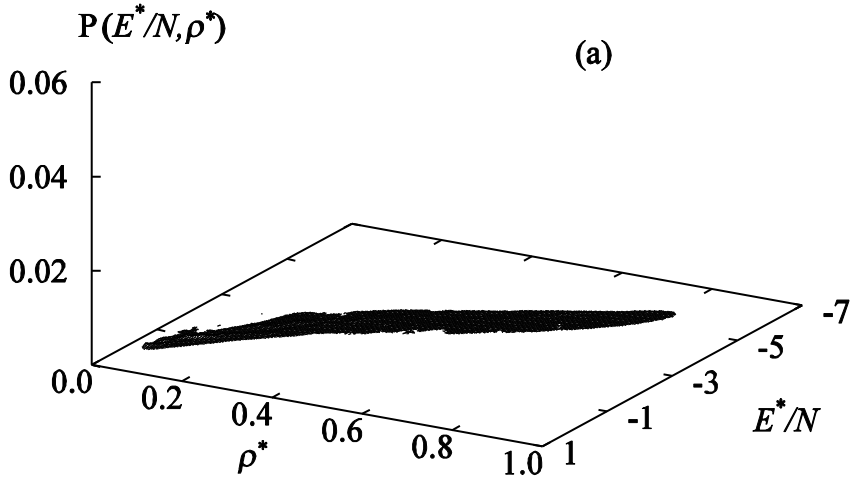
H. Okumura & Y.O., *J. Phys. Soc. Jpn.* **73**, 3304 (2004).

Liquid-Gas Phase Transitions Studied by Multibaric-Multithermal Simulations



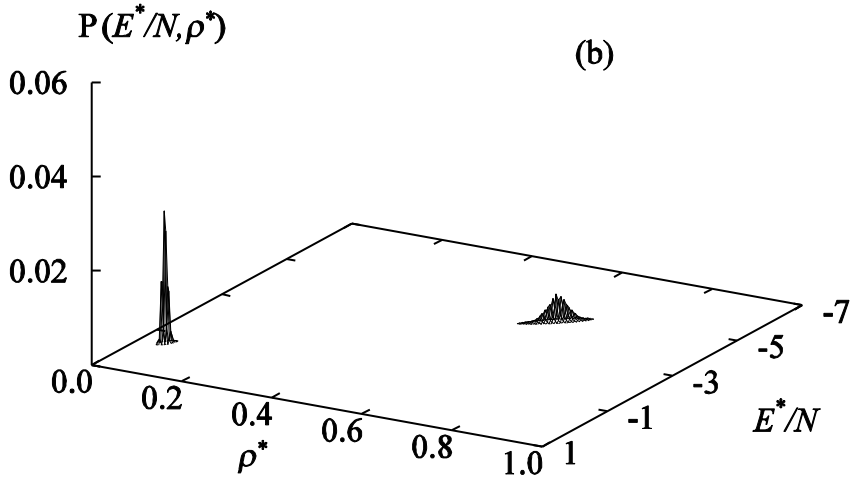
H. Okumura & Y.O., *J. Phys. Soc. Jpn.* **73**, 3304 (2004).

Liquid-Gas Phase Transitions Studied by Multibaric-Multithermal Simulations



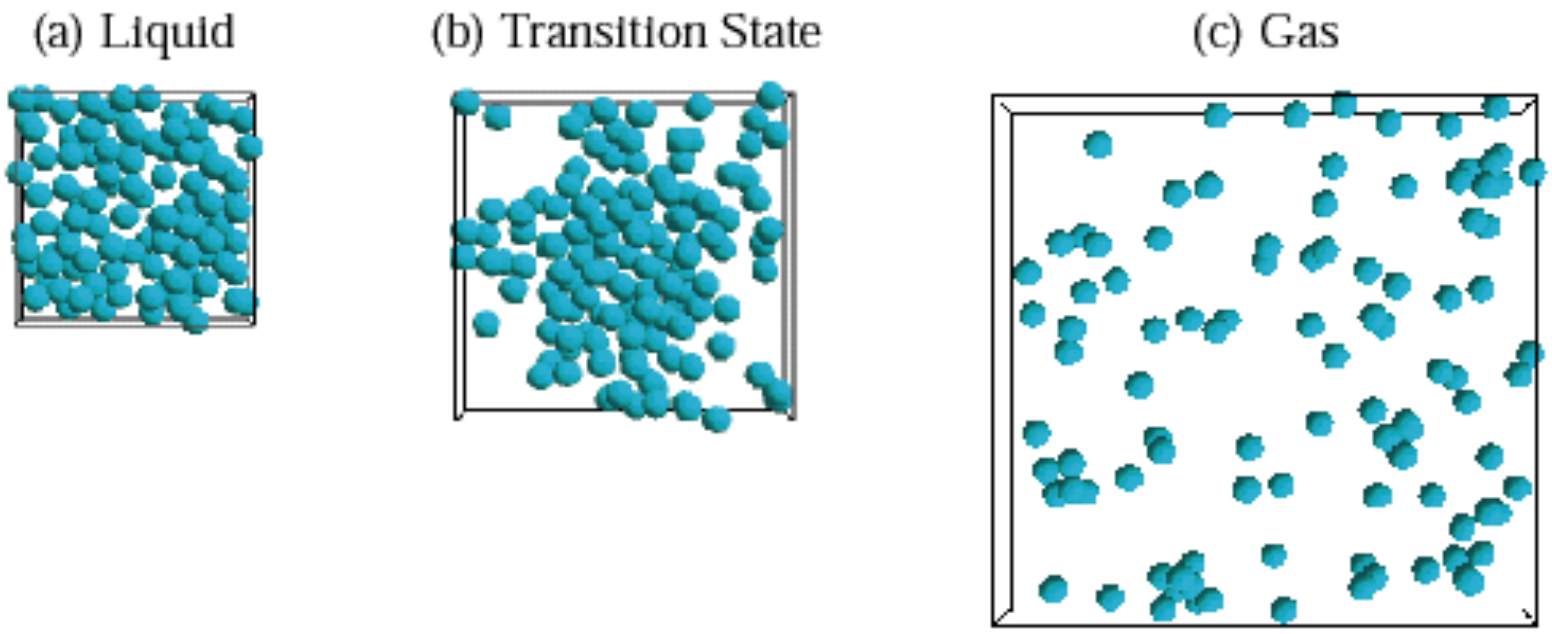
H. Okumura & Y.O.,
J. Phys. Soc. Jpn. **73**, 3304 (2004).

MUBATH



NPT (reweighted)

Liquid-Gas Phase Transitions Studied by Multibaric-Multithermal Simulations



H. Okumura & Y.O., *J. Phys. Soc. Jpn.* **73**, 3304 (2004).

ST Simulation in Isobaric-Isothermal Ensemble

Y. Mori & Y.O., *J. Phys. Soc. Jpn.* **79**, 074003 (2010);
Y. Mori & Y.O., *J. Comput. Chem.* **38**, 1167-1173 (2017);
G. La Penna, Y. Mori, R. Kitahara, K. Akasaka, & Y.O., *J. Chem. Phys.* **145**, 085104 (2016);
Y. Mori et al., in preparation.

$$E_\lambda = E + \lambda V$$

temperature and pressure become dynamical variables.

$$\lambda = P = \text{pressure} \qquad V = \text{volume}$$

random walk in temperature and pressure

CF. Multibaric-Multithermal Algorithm (MUBATH)

random walk in potential energy *E* and volume *V*

H. Okumura & Y.O., *Chem. Phys. Lett.* **383**, 391 (2004). (MC version)
H. Okumura & Y.O., *Chem. Phys. Lett.* **391**, 248 (2004). (MD version)

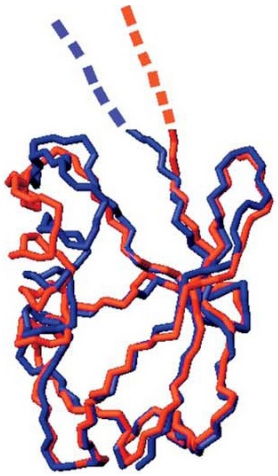
CF. Y. Mori & H. Okumura, *J. Phys. Chem. Lett.* **4**, 2079–2083 (2013).

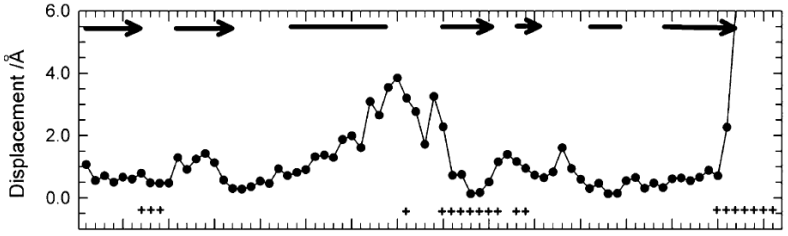
Cf. previous works:
D. Paschek & A. Garcia, *Phys. Rev. Lett.* **93**, 238105 (2004);
T. Imai, S. Ohyama, A. Kovalenko, & F. Hirata, *Protein Sci.* **16**, 1927 (2007);
T. Imai & Y. Sugita, *J. Phys. Chem. B* **114**, 2281 (2010).

Pressure-Induced Unfolding of Ubiquitin

NMR Experiments

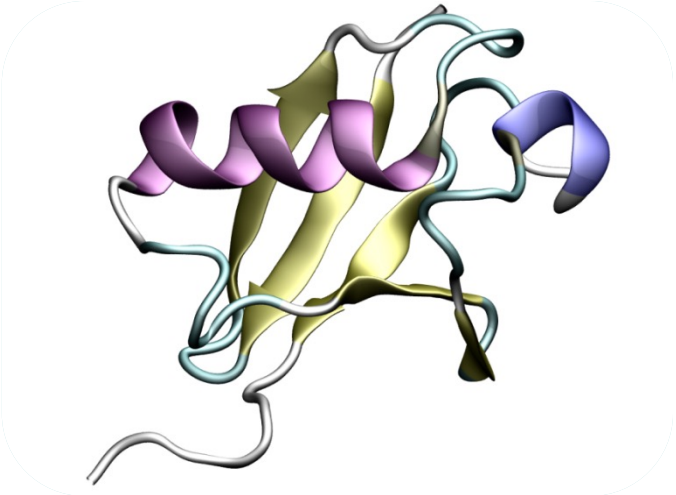
30 bar – 3000 bar





Simulated System

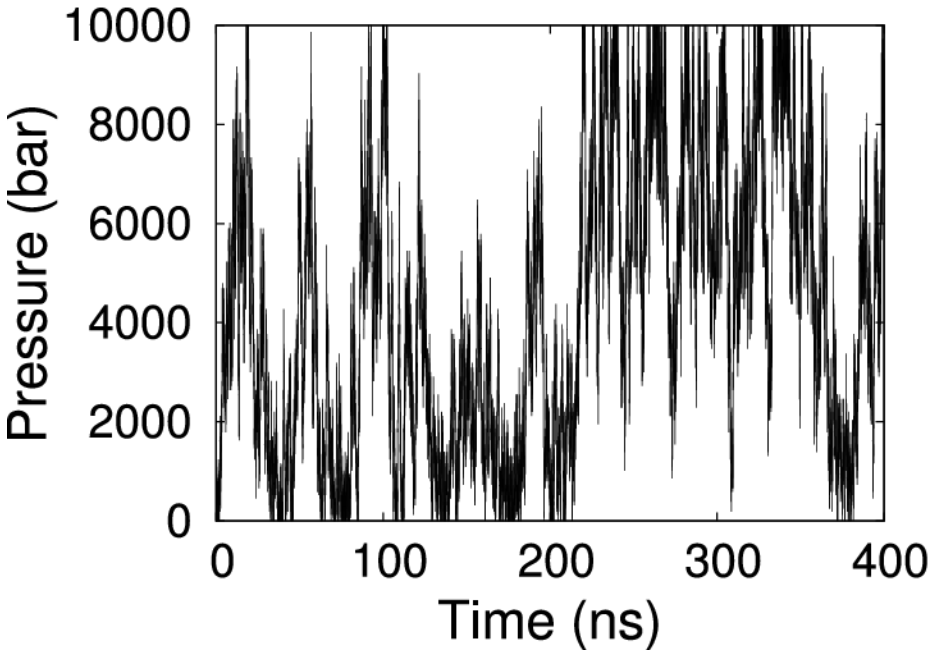
- 76 amino acids
- 6232 water molecules
- 19985 atoms

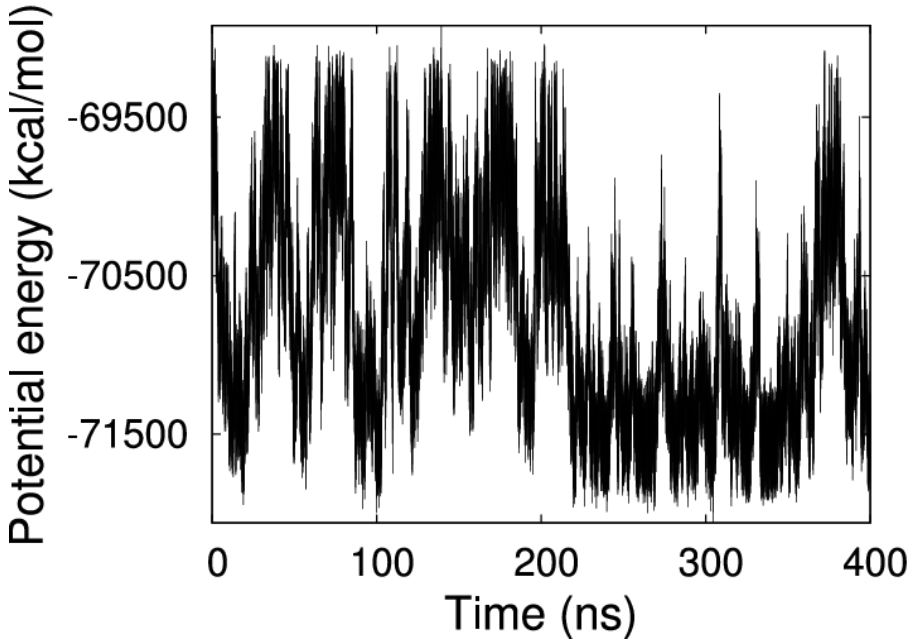


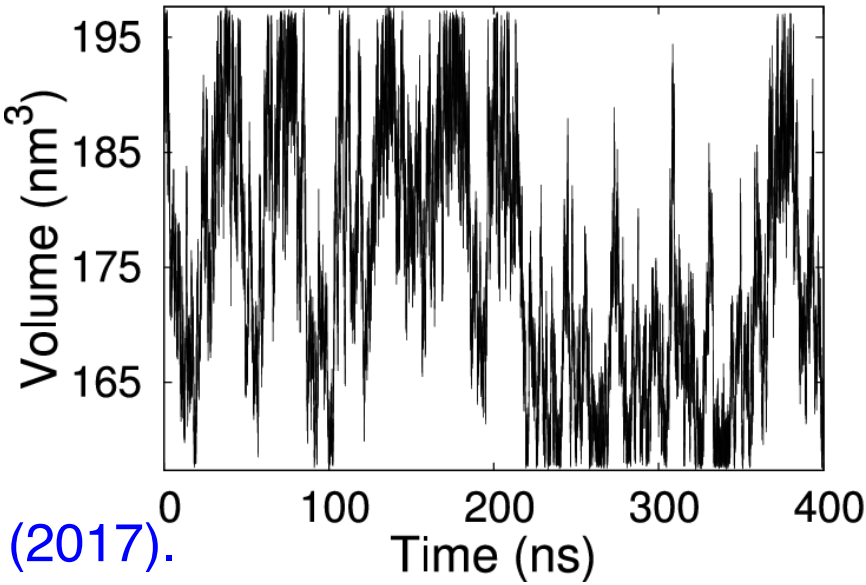
PDB: 1V80

R. Kitahara, S. Yokoyama & K. Akasaka, *J. Mol. Biol.* **347**, 277 (2005).

Time series of pressure *P*, potential energy *E*, and volume *V* for ubiquitin





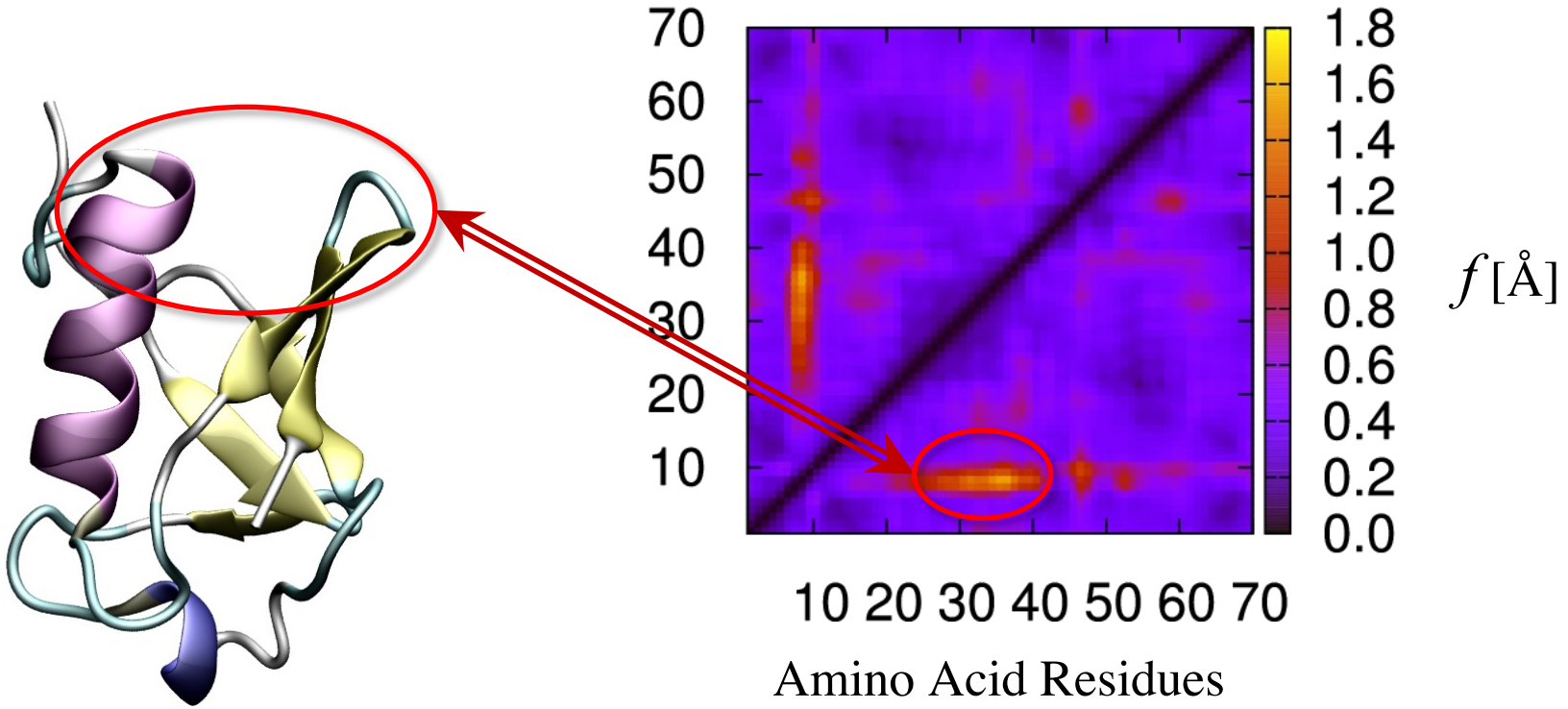


Y. Mori & Y.O., *J. Comput. Chem.* **38**, 1167 (2017).

Large structural fluctuations

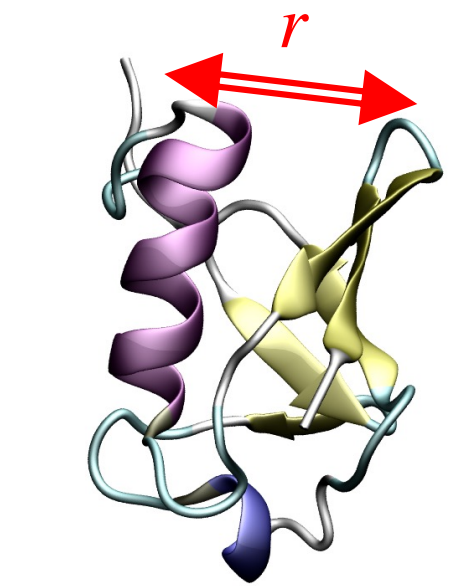
- Fluctuations of distance *d* between pairs of Cα atoms.

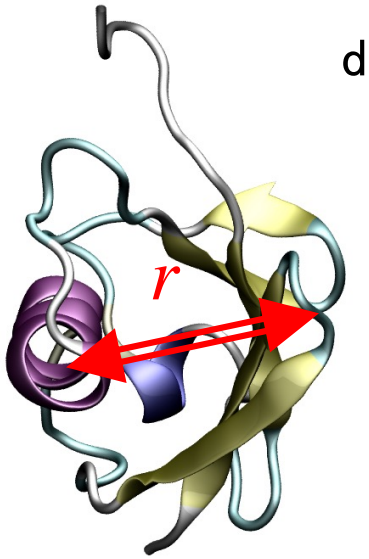
$$f \equiv \sqrt{\langle d^2 \rangle - \langle d \rangle^2}$$



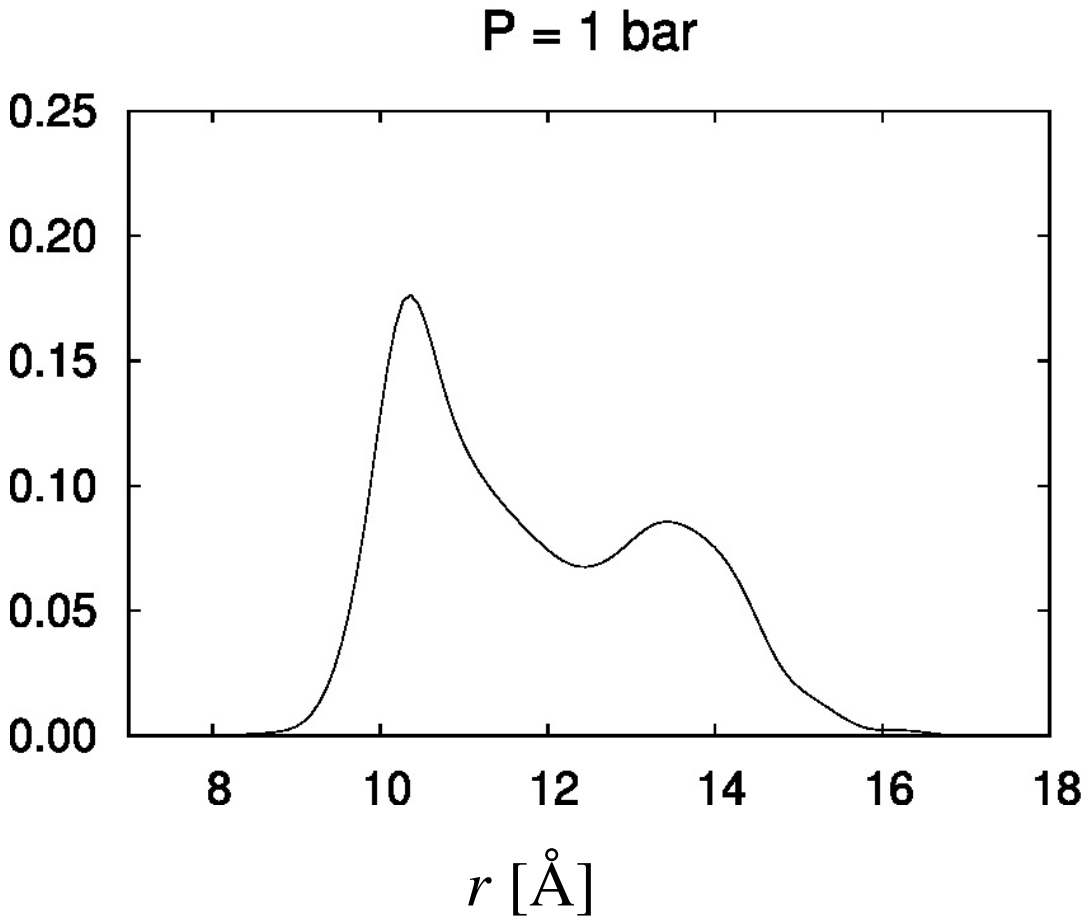
Large fluctuations observed in agreement with experiments.

Structural changes under high pressure

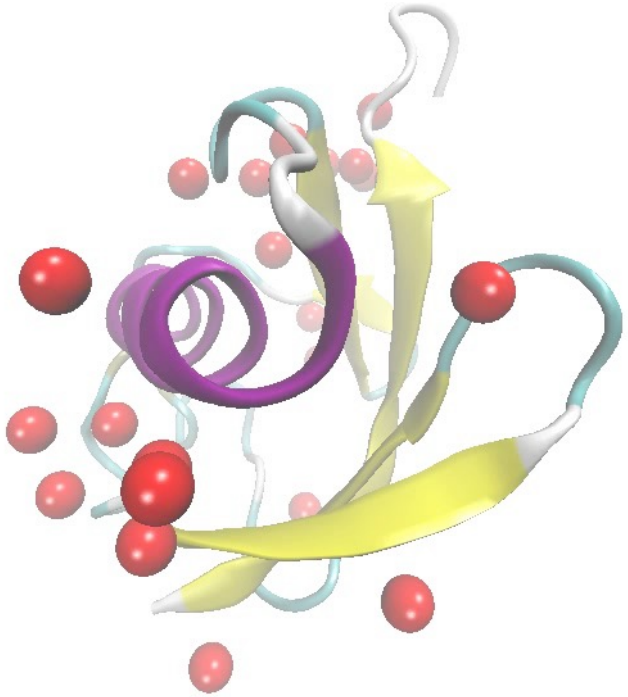


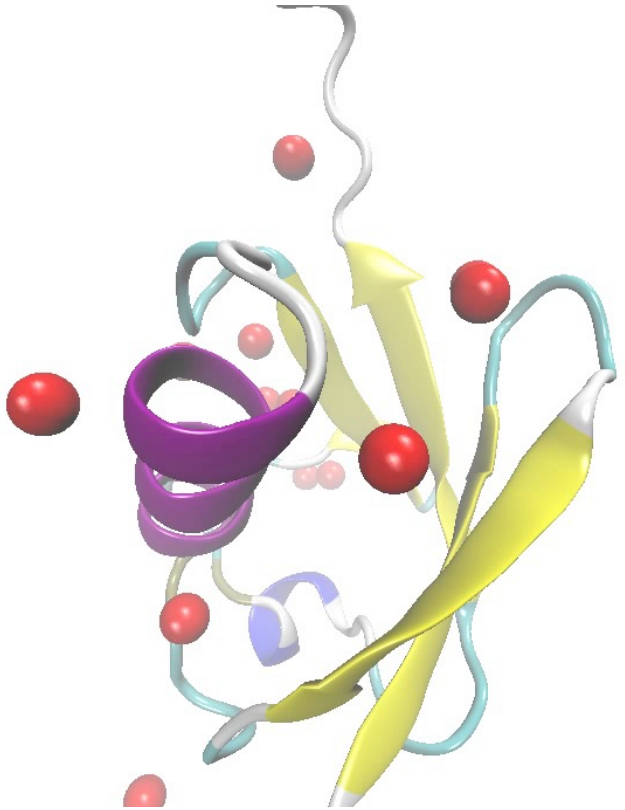


Simulation and movie by Y. Mori



Ubiquitin and water molecules





at low pressure

at high pressure

Simulation and movie by Y. Mori

Examples of Multidimensional REM, MUCA, and ST

A. Mitsutake & Y.O., *Phys. Rev. E* **79**, 047701 (2009);
J. Chem. Phys. **130**, 214105 (2009);
A. Mitsutake, *J. Chem. Phys.* **131**, 094105 (2009).

$$E_\lambda = E + \lambda V$$

4. Umbrella Sampling

$$H_k(q,p) = K(p) + E_0(q) + \sum_{l=1}^{L} \lambda^l V_l(q), \text{ where } V_l(\xi) = K_l\left(\xi(q) - d_l\right)^2.$$

* **REM: Replica-Exchange Umbrella Sampling (REUS)**
  random walk in reaction coordinate ξ
  Y. Sugita, A. Kitao & Y.O., *J. Chem. Phys.* **113**, 6042 (2000).
* **ST: Simulated Tempering Umbrella Sampling (STUS)**
  random walk in reaction coordinate ξ
  Y. Mori & Y.O., *Phys. Rev. E* **87**, 023301 (2013).
* **MUCA: Metadynamics (Wang-Landau in reaction coordinate)**
  random walk in reaction coordinate ξ
  A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. USA* **99**, 12562 (2002).

see also:
T. Huber, A.E. Torda, W.F. van Gunsteren, *J. Comp. Aid. Mol. Des.* **8**, 695 (1994);
H. Grubmuller, *Phys. Rev. E* **52**, 2893 (1995).

Prediction of Protein-Ligand Binding Structures by Replica-Exchange Umbrella Sampling

H. Kokubo, T. Tanaka & Y.O., *J. Comput. Chem.* **32**, 2810-2821 (2011);

H. Kokubo, T. Tanaka & Y.O., *J. Chem. Theor. Comput.* **9**, 4660-4671 (2013);

H. Kokubo, T. Tanaka & Y.O., *J. Comput. Chem.* **34**, 2601-2614 (2013);

Y.O., H. Kokubo & T. Tanaka, *J. Chem. Theor. Comput.* **10**, 3563-3569 (2014);

S. Tsukamoto, Y. Sakae, Y. Itoh, T. Suzuki & Y.O., *J Chem. Phys.* **148**, 125102 (2018);

H. Mishima, Y. Itoh, T. Kurohara, T. Suzuki, N. Asada, K. Kusakabe, and Y.O., *J. Comput. Chem.* **44**, 1604-1609 (2023).

Replica-Exchange Umbrella Sampling (REUS)

Y. Sugita, A. Kitao & Y.O., *J. Chem. Phys.* **113**, 6042 (2000).

$$H_k(q,p) = K(p) + E_0(q) + \sum_{l=1}^{L} \lambda^l V_l(q), \text{ where } V_l(\xi) = K_l\left(\xi(q) - d_l\right)^2.$$

- **Potential of Mean Force**

$$W_\beta(\xi) = -\frac{1}{\beta} \ln P_{\beta,\lambda=0}(\xi)$$

WHAM

$$P_{\beta,\lambda}(E,\xi) = \frac{\sum_{m=1}^{M} N_m(E,\xi) e^{-\beta\left(E + V_\lambda(\xi)\right)}}{\sum_{m=1}^{M} n_m e^{f_m - \beta_m\left(E + V_{\lambda_m}(\xi)\right)}},$$

$$e^{-f_m} = \sum_E \sum_\xi P_{\beta_m,\lambda_m}(E,\xi)$$

See also:

E.M. Boczko & C.L. Brooks, *J. Phys. Chem.* **97**, 4509 (1993).

B. Roux, *Comp. Phys. Commun.* **91**, 275 (1995).

Five test systems [ligand (protein)] ($T = 300$ K, $P = 1$ atm)

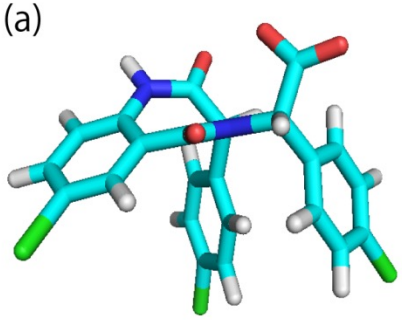
H. Kokubo, T. Tanaka & Y.O., *J. Comput. Chem.* **32**, 2810-2821 (2011).

benzodiazepine (protein: MDM2)

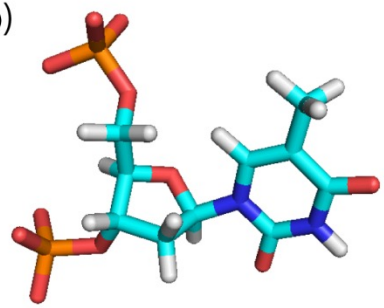
deoxythymidine 3',5'-bisphosphate (pdTp) (protein: staphylococcal nuclease)

tyrosine (protein: aldolase)

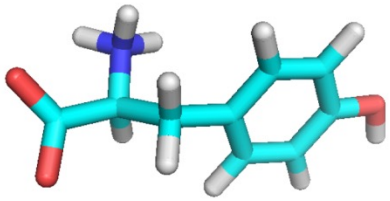
(a)



(b)



(c)

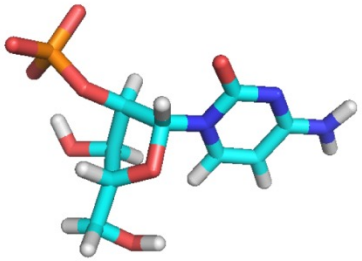


PDB code → 1T4E

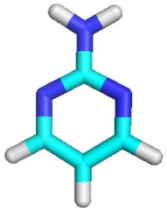
1SNC

1OF6

(d)



(e)



1ROB

2JJC

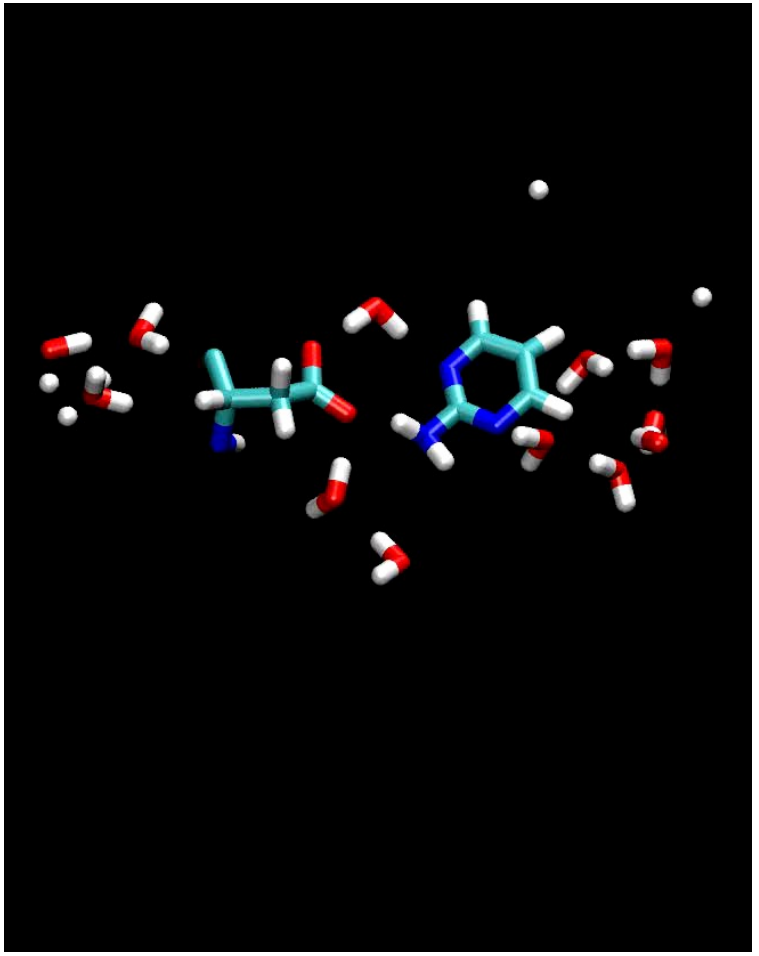
← PDB code

cytidylic acid (2'-CMP) (protein: ribonuclease A)

2-aminopyrimidine (protein: heat shock protein HSP90)

heat shock protein and
2-aminopyrimidine

Asp-78 and ligand

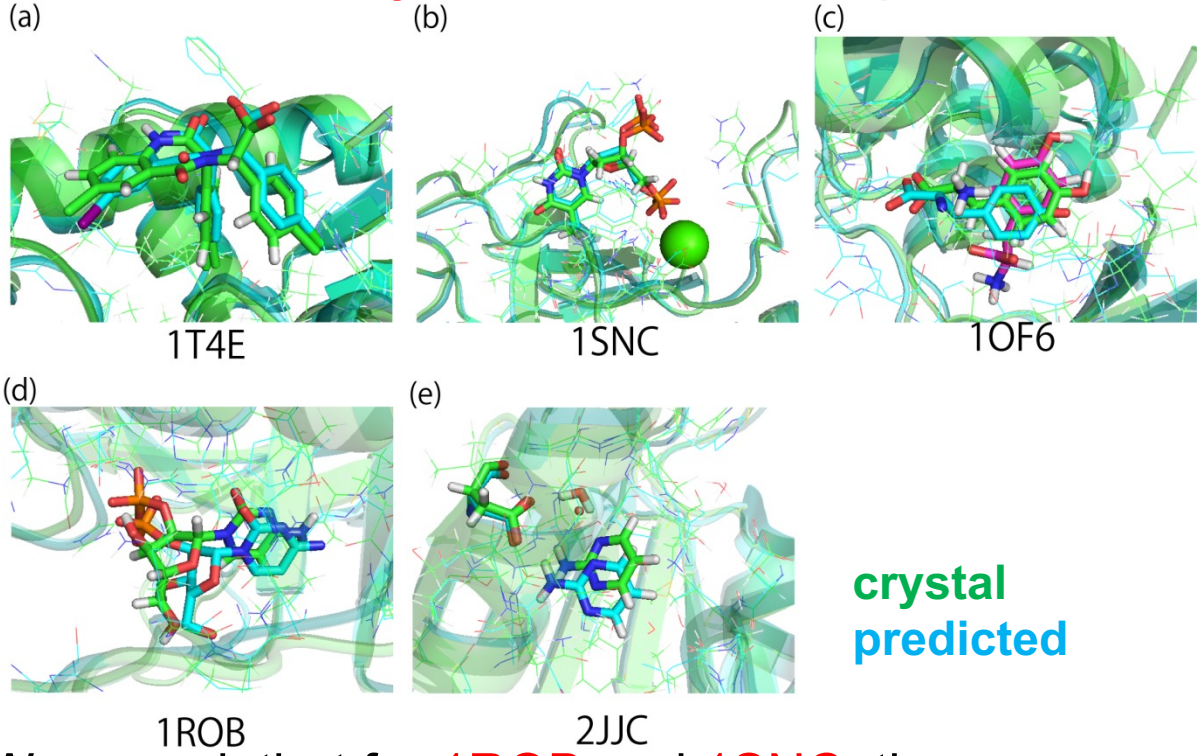


Simulation and movie
by H. Kokubo

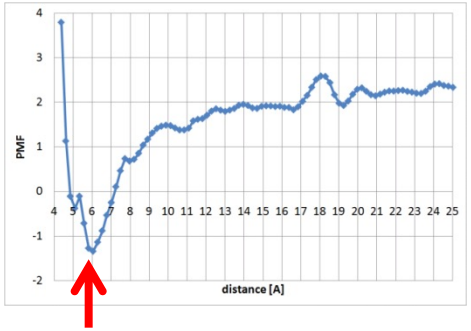
Results: REUS with 24 replicas

H. Kokubo, T. Tanaka & Y.O., *J. Comput. Chem.* **32**, 2810-2821 (2011).

- Starting from configurations in which protein and ligand are far away from each other in each system, our method predicted the ligand binding structures as the global minima in free energy (or, potential of mean force) in excellent agreement with the experimental data from PDB.



crystal
predicted



potential of mean force shows the most stable distance

We remark that for 1ROB and 1SNC, there are attempts by a popular existing docking program, *GOLD*, but they failed in the predictions (classified as significant errors).

$$\mathcal{W}_{T,\boldsymbol{\lambda}=\{0\}}(\xi) = -k_B T \ln \left[\sum_{E_0} P_{T,\boldsymbol{\lambda}=\{0\}}(E_0, \xi) \right]$$

Prediction of Protein-Ligand Binding Structures by 2-dim H-REMD: Replica-Exchange Umbrella Sampling (REUS) and Replica-Exchange Solute Tempering (REST): REUS/REST

H. Kokubo, T. Tanaka & Y.O., *J. Comput. Chem.* **34**, 2601-2614 (2013) .

Cf. REUS/gREST: S. Re, H. Oshima, K. Kasahara, M. Kamiya & Y.Sugita, *PNAS* **116**, 18404-18409 (2019).

Potential Energy:

$$E_{mn}(q^{[i]}) = \frac{\beta_n}{\beta_0} U_{ll}(q^{[i]}) + \sqrt{\frac{\beta_n}{\beta_0}} U_{ls}(q^{[i]}) + U_{ss}(q^{[i]}) + V_m(q^{[i]})$$

l: ligand

s: protein/water

Total no. of replicas: $M \times N$

REST parameters (n=1, 2, …, N)

REUS parameters, i.e., umbrella potentials (m=1, 2, …, M)

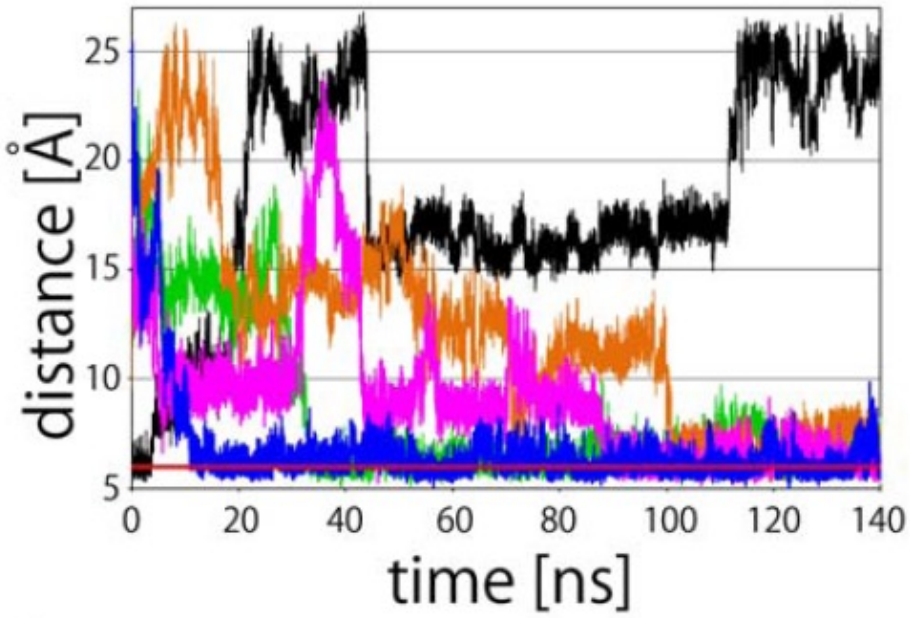
REST: P. Liu, B. Kim, R.A. Friesner, & B.J. Berne, *PNAS* **102**, 13749 (2005).
T. Terakawa, T. Kameda, and S. Takada, *J. Comput. Chem.* **32**, 1228 (2011).

REUS: Y. Sugita, A. Kitao & Y.O., *J. Chem. Phys.* **113**, 6042 (2000).

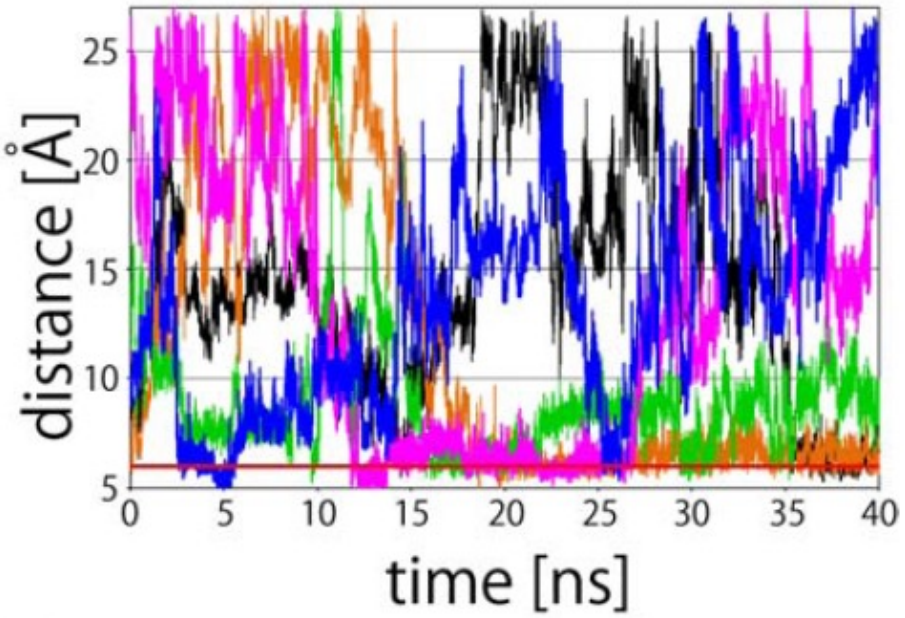
2-dim H-REMD: REUS (*M*=24) + REST (*N*=8) (192 replicas)

H. Kokubo, T. Tanaka & Y.O., *J. Comput. Chem.* **34**, 2601-2614 (2013) .

REUS (1T4E)



REUS/REST (1T4E)



No. of Random Walk Cycles

	1T4E		2JJC	
	REUS	REUS/REST	REUS	REUS/REST
from min to max	1	44	4	31
from max to min	3	68	4	34
either direction	4	112	8	65
return	0	28	1	11

Prediction of Ligand Binding Affinity by 2-dim H-REMD: REUS/REST simulation

Y.O., H. Kokubo & T. Tanaka, *J. Chem. Theory Comput*. **10**, 3563-3569 (2014).

Table 2. Calculation of the Absolute Binding Free Energy of Compound 29 with MDM2 by the Double Decoupling Method[a]

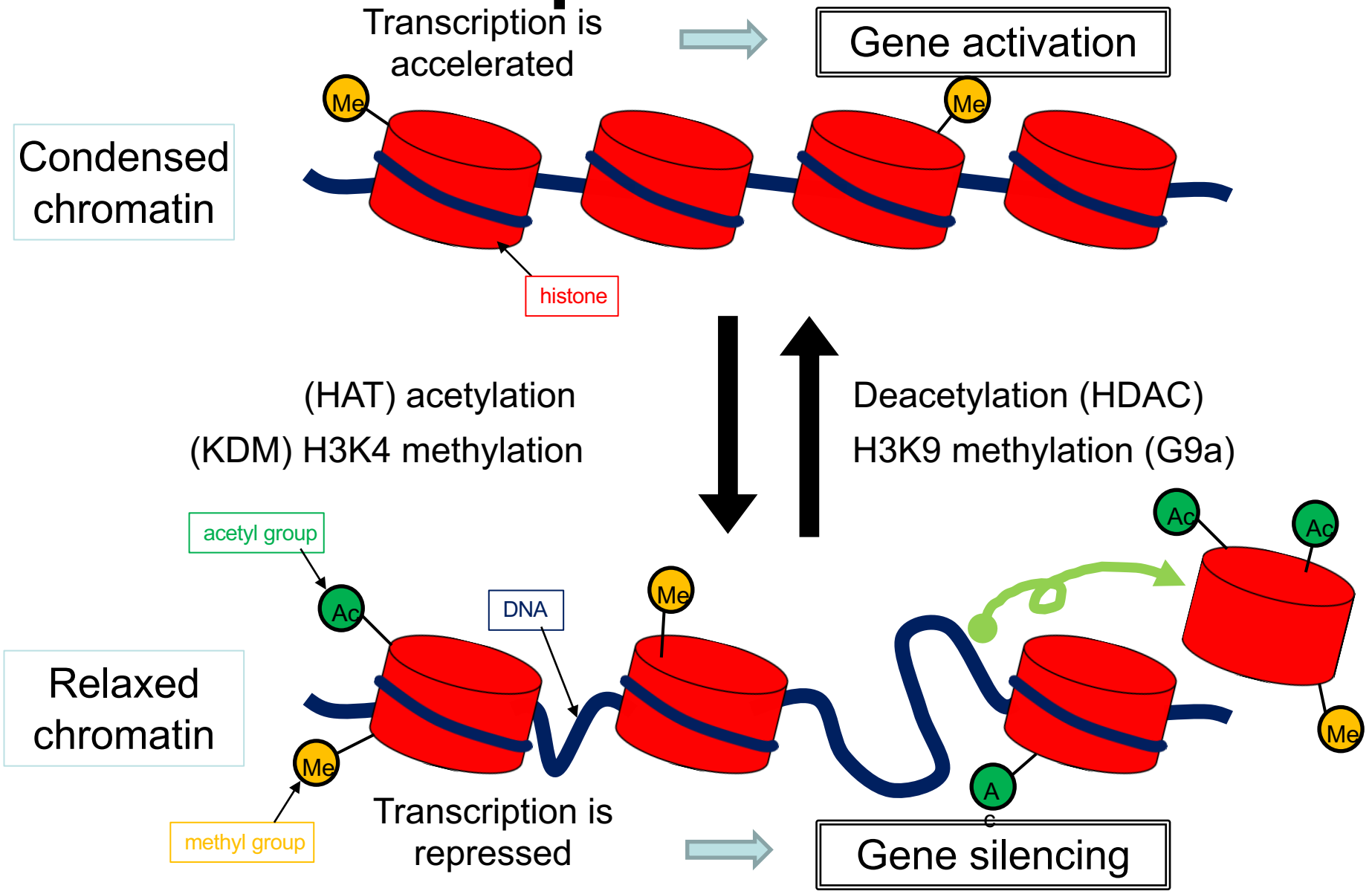
	ΔG_{total}	ΔG_{elec}	ΔG_{vdW}	$\Delta G_{restraint}$
from gas to water	−81.49	−83.51	2.02	0.00
from gas to complex	−92.88	−80.44	−16.92	4.48
from water to complex (binding)	−11.39	3.07	−18.94	4.48

[a] $\Delta G_{total} = \Delta G_{elec} + \Delta G_{vdW} + \Delta G_{restraint}$ for each process. Unit is kcal/mol. Calculated binding free energy is −11.39 kcal/mol, which is comparable with the experimental value −12.30 kcal/mol.

Studies for Ligand Binding Selectivity of Histone Deacetylase Inhibitor by Molecular Dynamics Simulation

S. Tsukamoto, Y. Sakae, Y. Itoh, T. Suzuki, Y.O.,
J. Chem. Phys. **148**, 125102 (2018).

Epigenetic regulations of gene expression



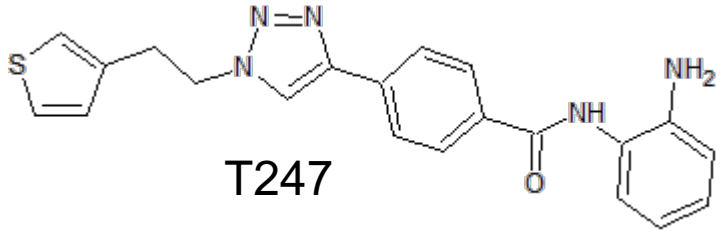
HDAC isozymes

There are 18 HDAC isozymes. They are divided into 4 classes by sequence homology and functions. The HDACs which belong to the same class have very similar sequences and structures. When we use HDAC inhibitors for therapy, we should consider selectivity of the inhibitors. Unnecessary inhibition may cause critical side effect.

- Class 1: HDAC1, HDAC2, HDAC3, HDAC8
- Class 2: HDAC4, HDAC5, HDAC7, HDAC9, HDAC6, HDAC10
- Class 3: SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6, SIRT7
- Class 4: HDAC11

Yellow: Crystal structure of HDAC2 (pdbID: 3max).
Blue: Crystal structure of HDAC3 (pdbID: 4a69)

Purpose of this study



T247

T. Suzuki, Y. Kasuya, Y. Itoh, Y. Ota, P. Zhan, K. Asamitsu, H. Nakagawa, T. Okamoto, N. Miyata *PLoS ONE*, **8**, e68669 (2013)

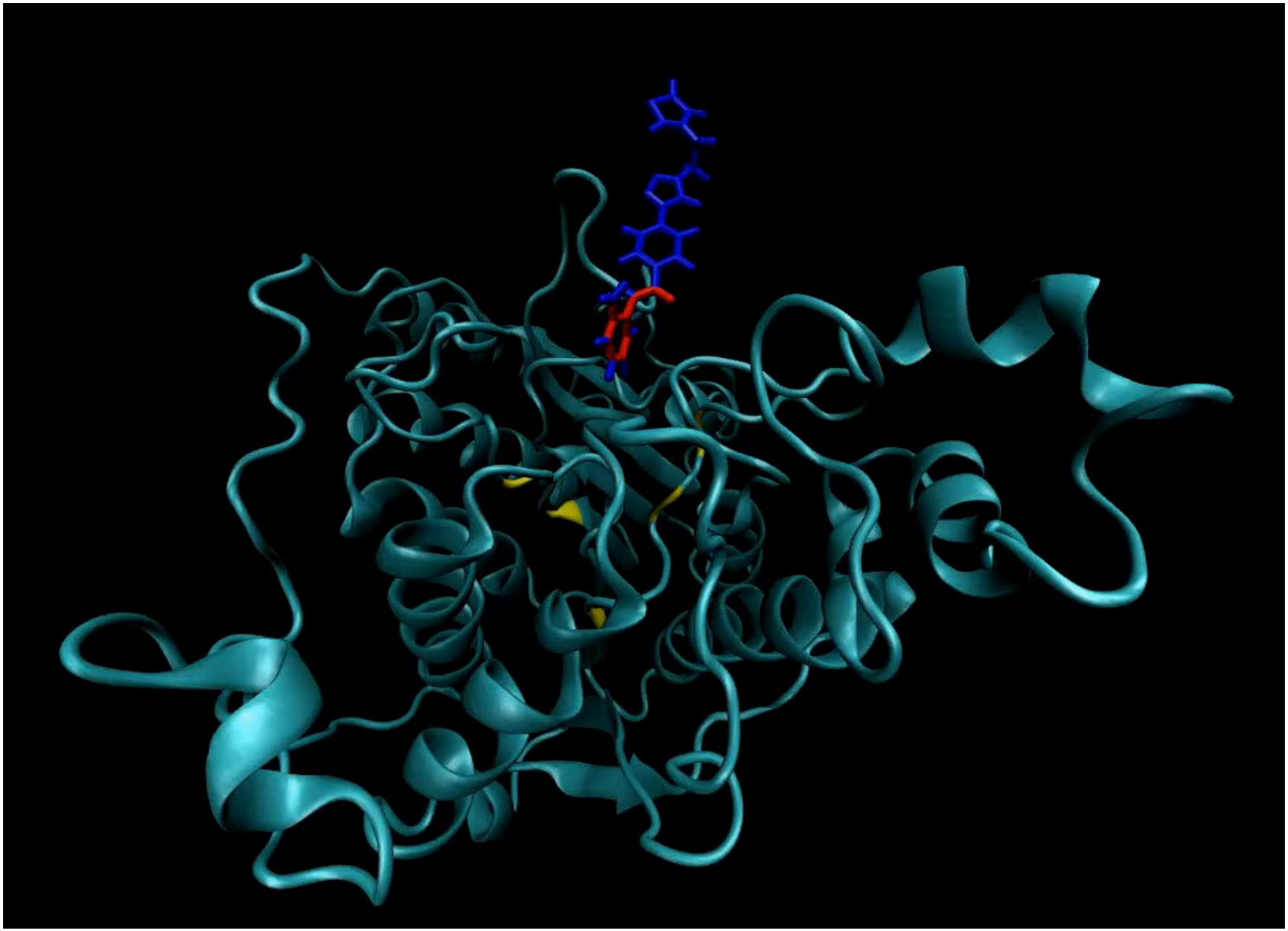
Compound	IC_{50} (μM)					
	class I				class IIa	class IIb
	HDACs	HDAC1	HDAC3	HDAC8	HDAC4	HDAC6
vorinostat (3)	0.073	0.39	0.27	0.66	>10	0.34
1	>100	>100	19	>100	>100	>100
T247	>100	19	0.24	>100	>100	>100
T326	>100	>100	0.26	>100	>100	>100

[a]Values are means of at least three experiments.
doi:10.1371/journal.pone.0068669.t002

1. To reproduce the selectivity of T247 by using molecular dynamics simulations
2. To elucidate the reason why T247 has the selectivity for HDAC3
3. To propose new selective inhibitors

REUS MD Simulation (One of the Replicas)

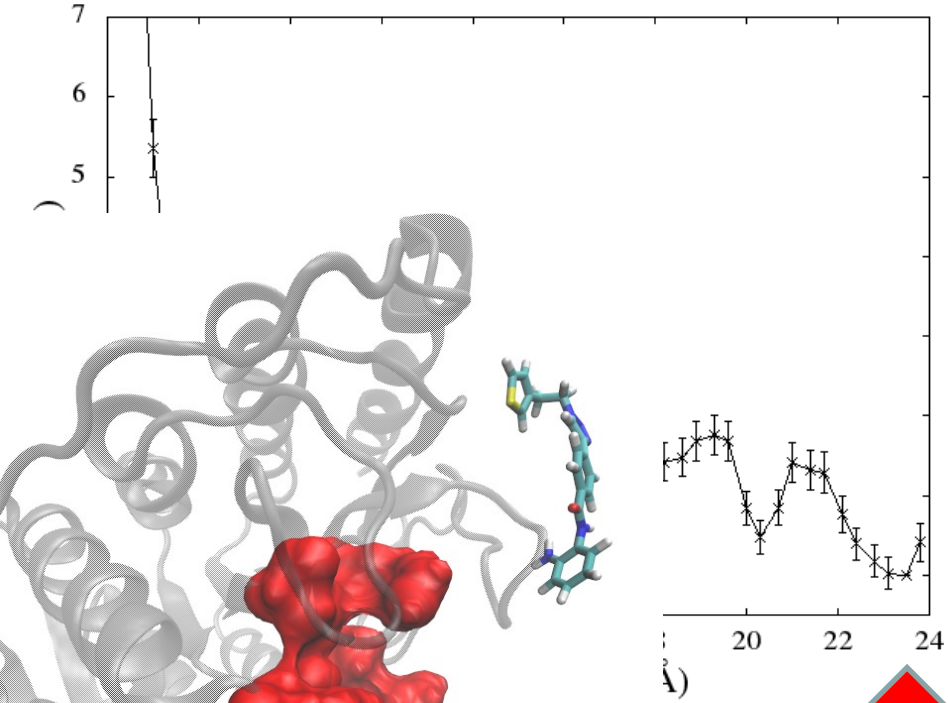
Simulation and movie by S. Tsukamoto



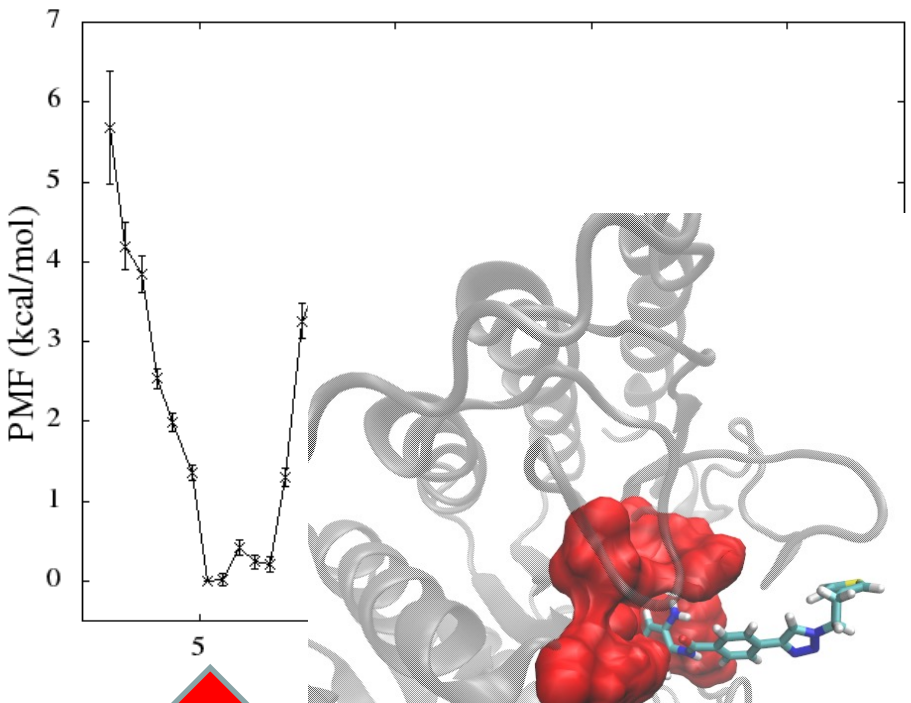
Potential of mean force

We succeeded in reproducing the selectivity.

HDAC2



HDAC3

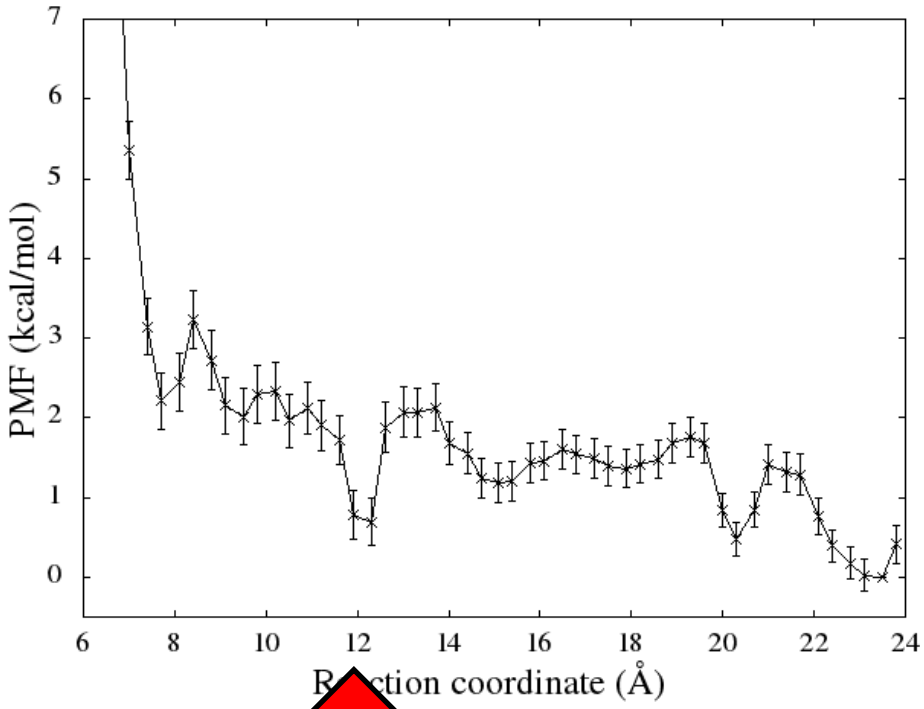


The most stable state

Potential of mean force

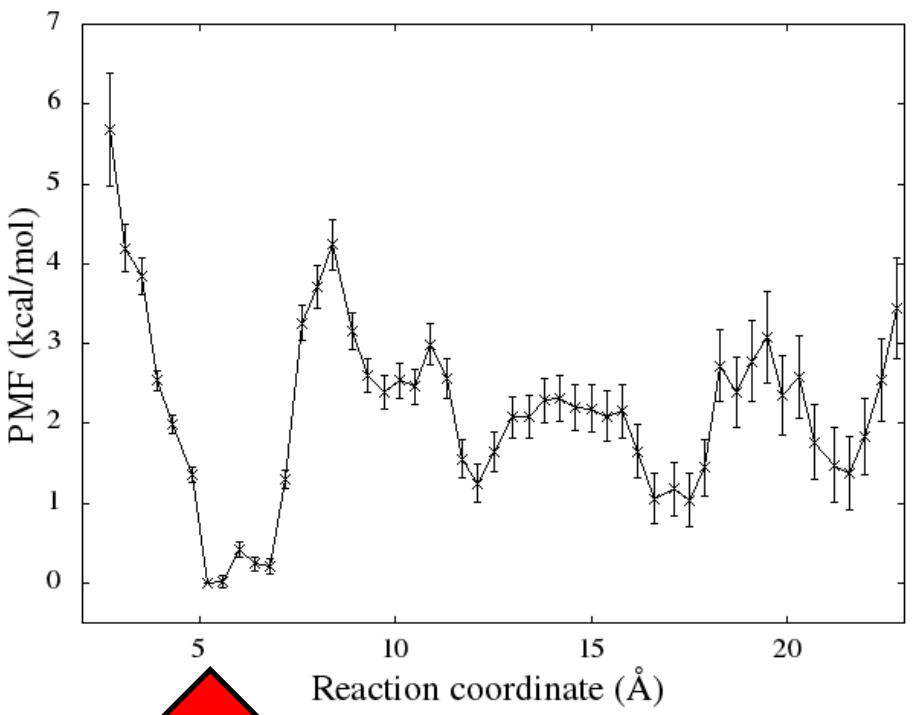
S. Tsukamoto, Y. Sakae, Y. Itoh, T. Suzuki, Y.O., *J. Chem. Phys.* **148**, 125102 (2018).

HDAC2



Docking state

HDAC3

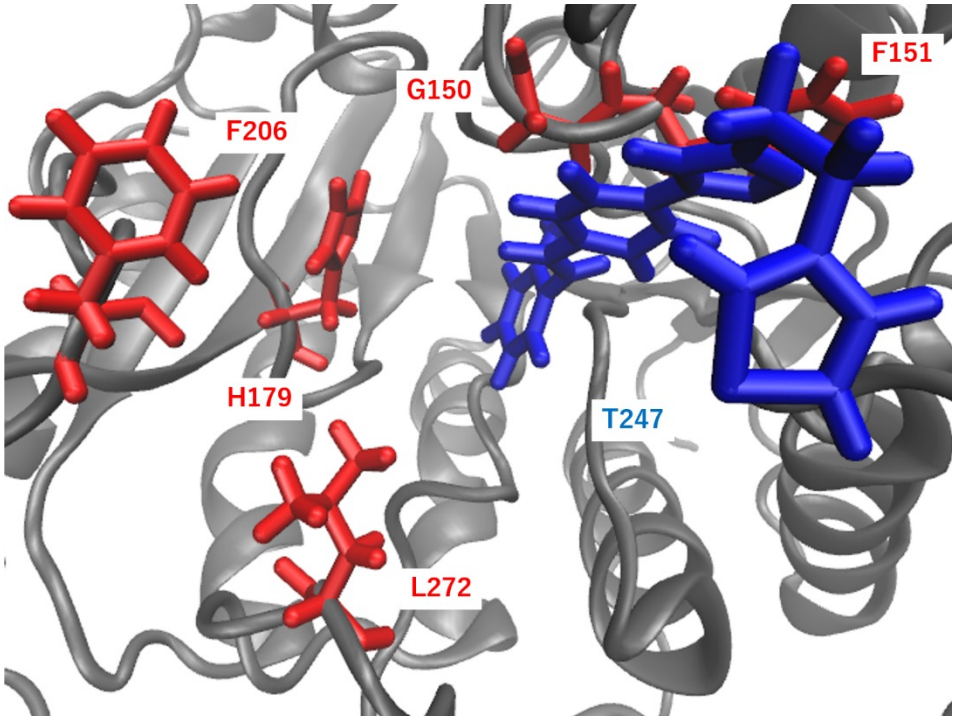


Docking state

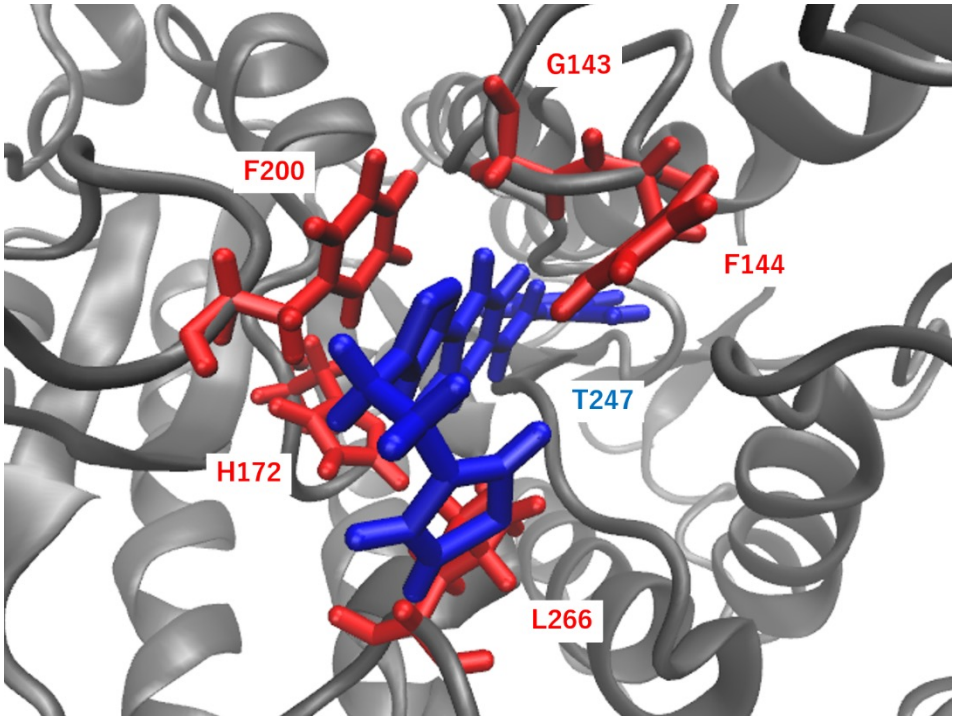
Hydrophobic structures at binding site

S. Tsukamoto, Y. Sakae, Y. Itoh, T. Suzuki, Y.O., *J. Chem. Phys.* **148**, 125102 (2018).

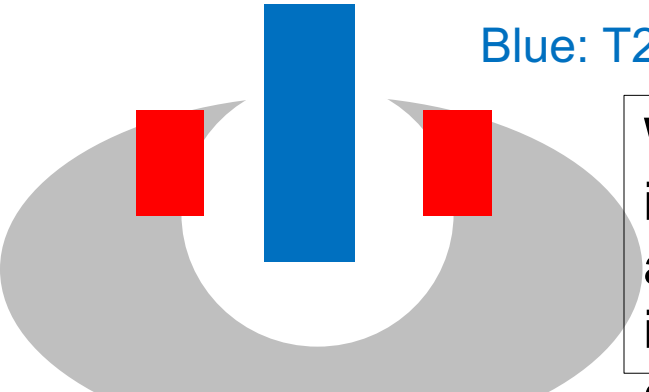
HDAC2



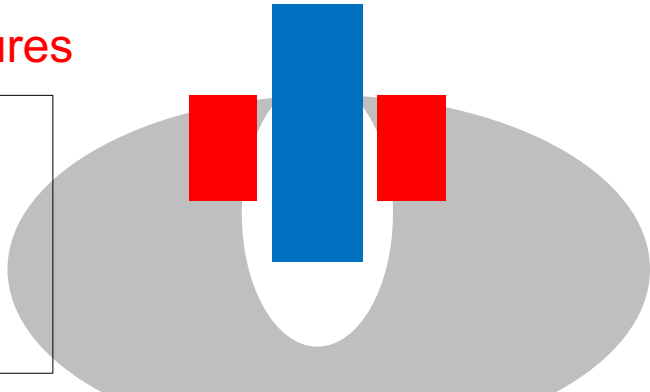
HDAC3



Blue: T247 Red: hydrophobic structures

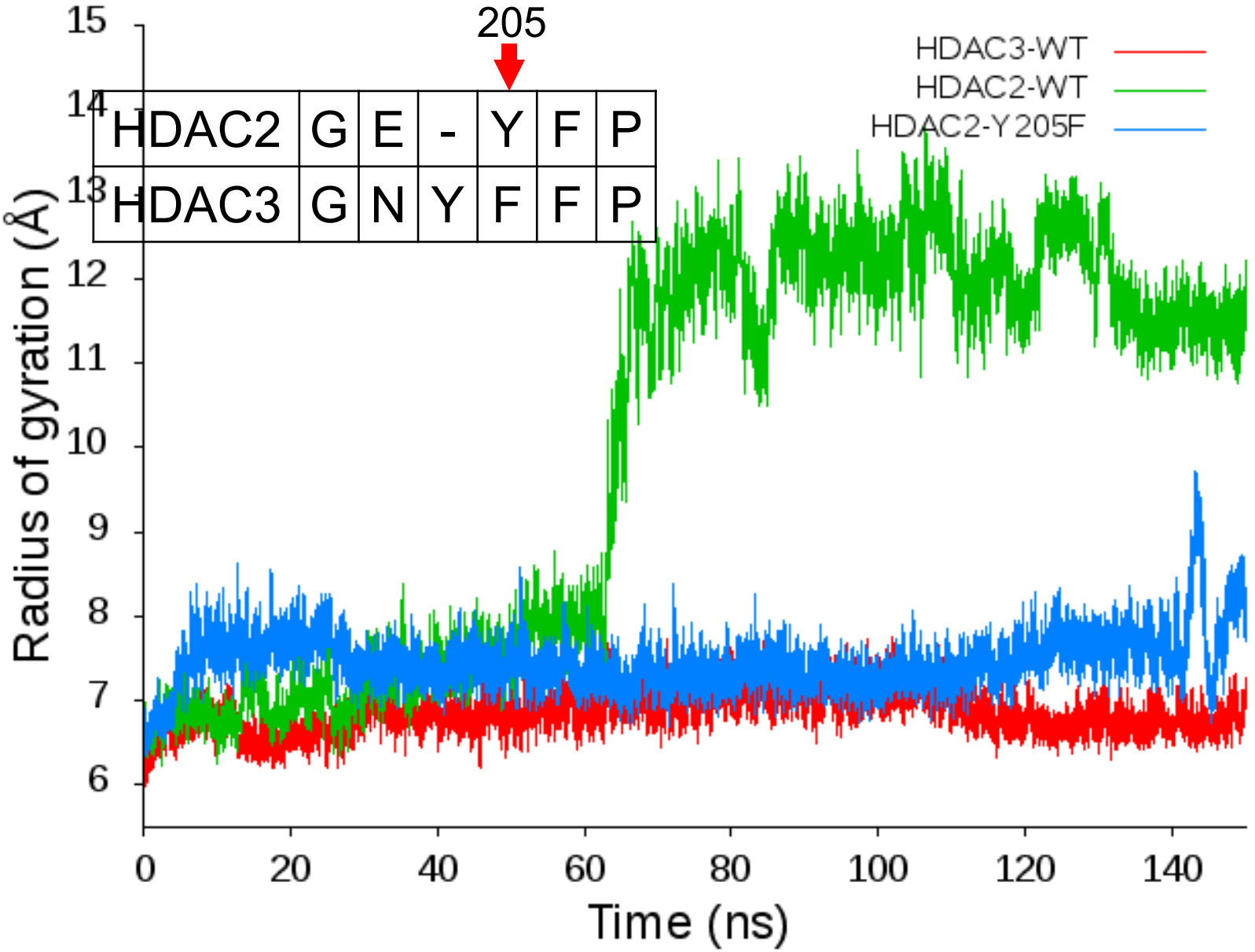


Weak hydrophobic interaction among T247 and hydrophobic residues in HDAC2 causes weak affinity

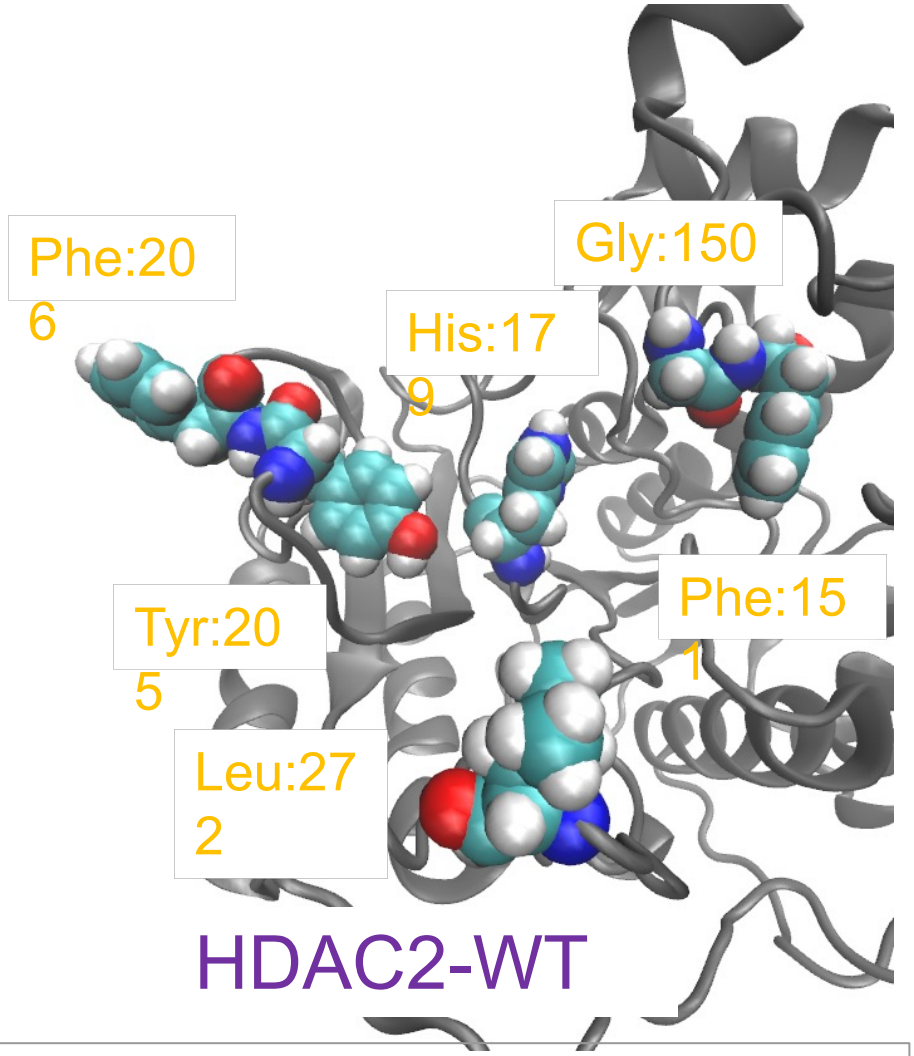


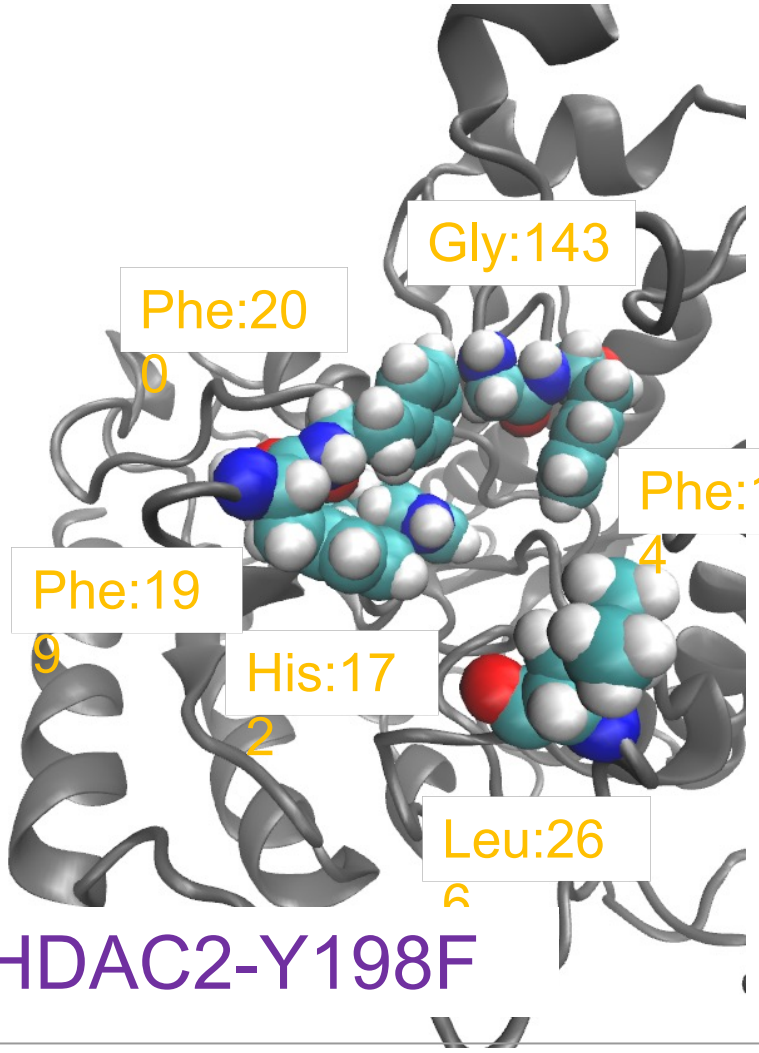
Radii of gyration of the hydrophobic structures

S. Tsukamoto, Y. Sakae, Y. Itoh, T. Suzuki, Y.O., *J. Chem. Phys.* **148**, 125102 (2018).



Structure of wild type and mutant of HDAC2





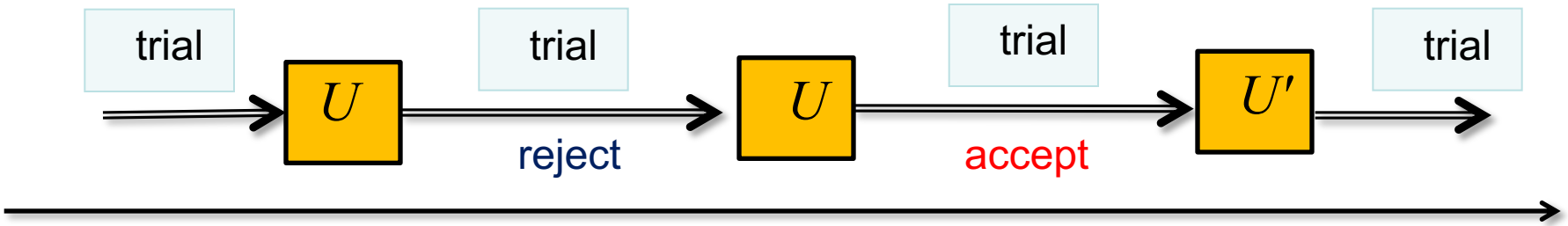
Hydrophobic structure was opened

Hydrophobic structure was closed

焼き戻し傘サンプル法

Simulated Tempering Umbrella Sampling (STUS)

Y. Mori & Y.O., *Phys. Rev. E* **87**, 023301 (2013).



アンブレラポテンシャル Umbrella Potential

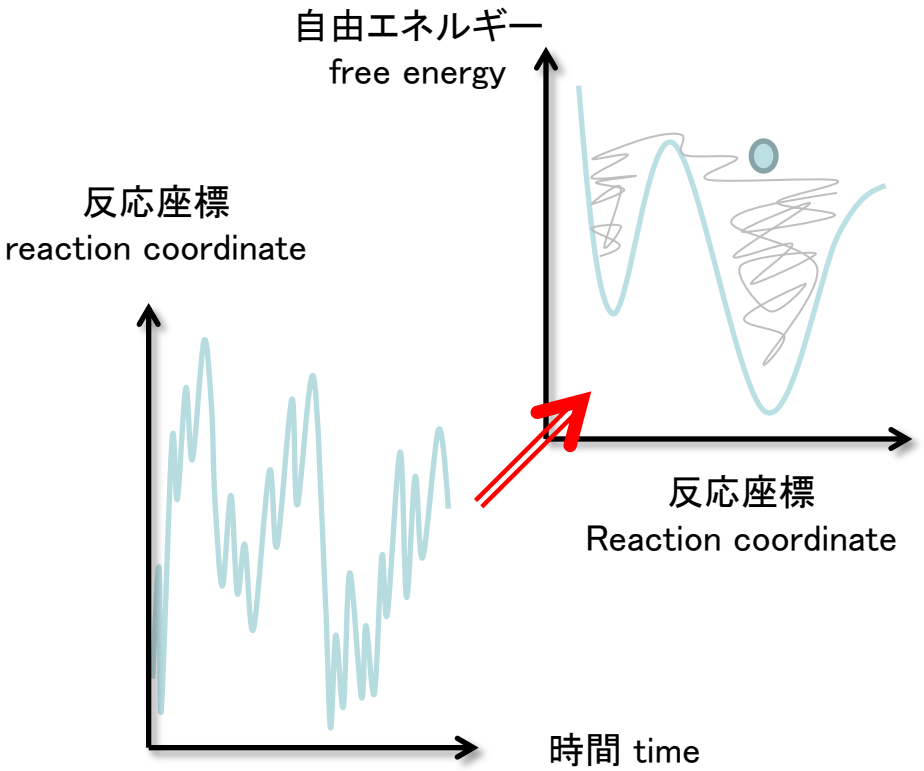
$$U[\xi(r)] = k\left[\xi(r) - \xi_0\right]^2$$

遷移確率 Transition Probability

$$w(X \to X') = \min\left[1, \exp(-\Delta)\right]$$
$$\Delta = \beta(U' - U) - (f' - f)$$

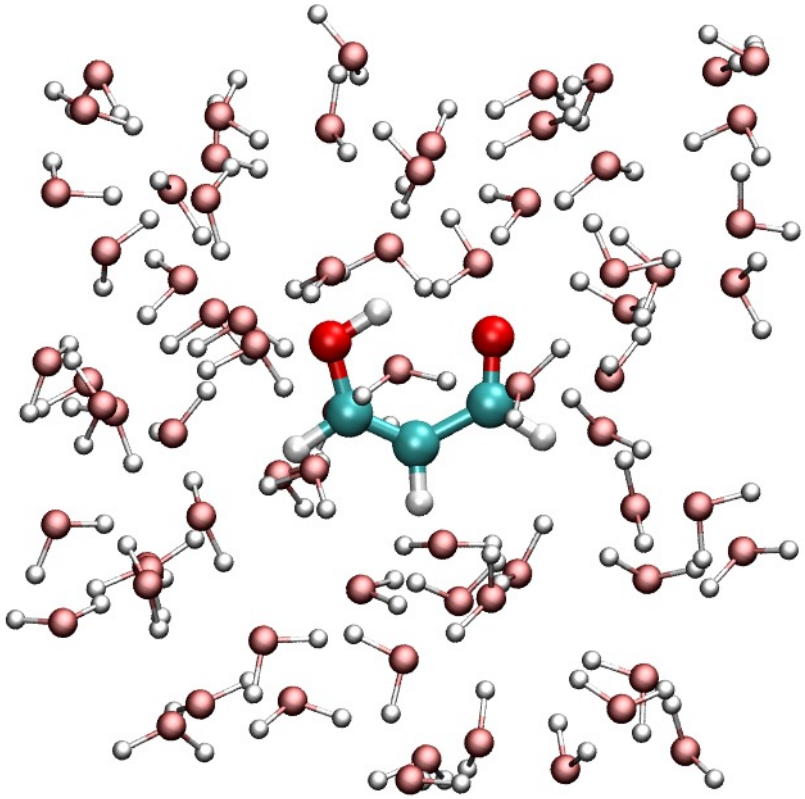
無次元化された自由エネルギー Dimensionless Free Energy

$$f = -\ln\left\{\int_{-\infty}^{\infty} dr \exp\left\{-\beta\left[E(r) + U(r)\right]\right\}\right\}$$



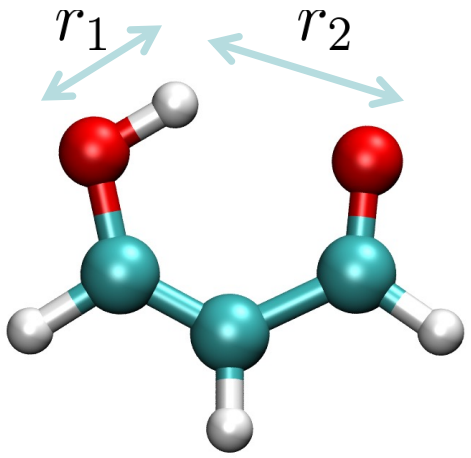
マロンアルデヒドの分子内プロトン移動
Proton Transfer in Malonaldehyde

Y. Mori & Y.O., *Phys. Rev. E* **87**, 023301 (2013).

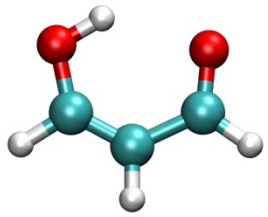


反応座標の定義
Definition of reaction coordinate

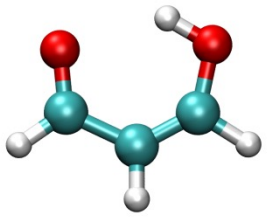
$$\xi = r_2 - r_1$$



$\xi > 0$



$\xi < 0$



マロンアルデヒド malonaldehyde：1 分子 molecule
水 water：71 分子 molecules
総原子数　Total No. of Atoms：222

Simulation Conditions（シミュレーション条件）

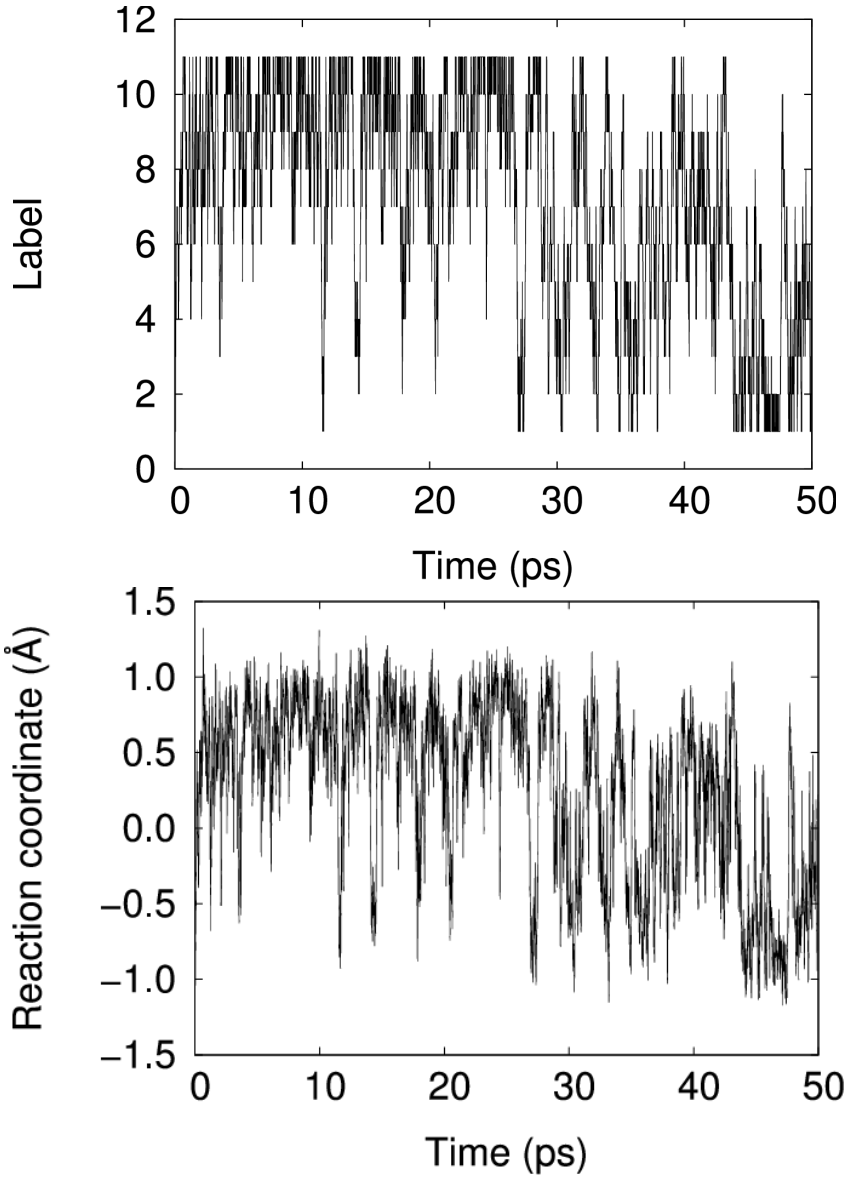
Y. Mori & Y.O., *Phys. Rev. E* **87**, 023301 (2013).

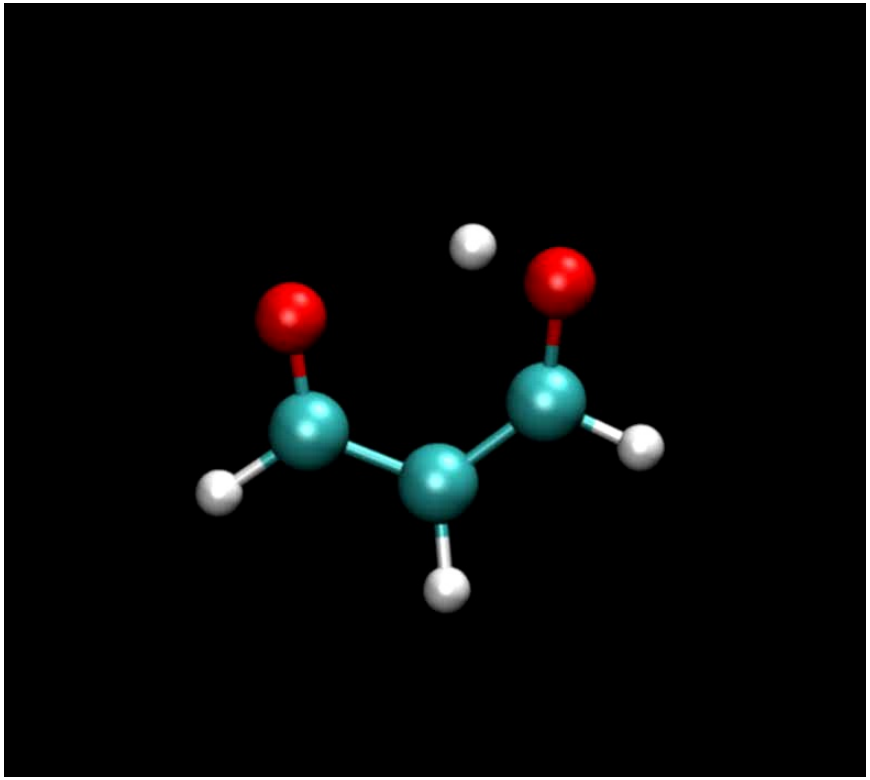
計算条件	
プログラム Program	CP2K (ver. 2.1) *
交換相関汎関数 Functional	BLYP
基底 Bases	Gaussian and plane wave
擬ポテンシャル Pseudo Potential	Goedecker-Teter-Hutter
アンサンブル Statistical Ensemble	*NVT*
温度制御法 Thermostat	Nosé-Hoover chains (# of chains = 3)
温度 Temperature	300 K
体積 Volume	13.82 × 13.82 × 13.82 (Å^3)
時間ステップ Time Step	0.5 fs
総シミュレーション時間 Total Sim. Time	200 ps (50 ps × 4 runs)
Umbrella Potentials	
アンブレラポテンシャル（k）	0.01 (hartree·bohr^{-2})
アンブレラポテンシャル（ξ_0） 11個のパラメータを使用	-1.0, -0.8, -0.6, -0.4, -0.2, 0.0, 0.2, 0.4, 0.6, 0.8, 1.0 (Å)

* http://cp2k.berlios.de/

シミュレーション結果 Results

Y. Mori & Y.O., *Phys. Rev. E* **87**, 023301 (2013).

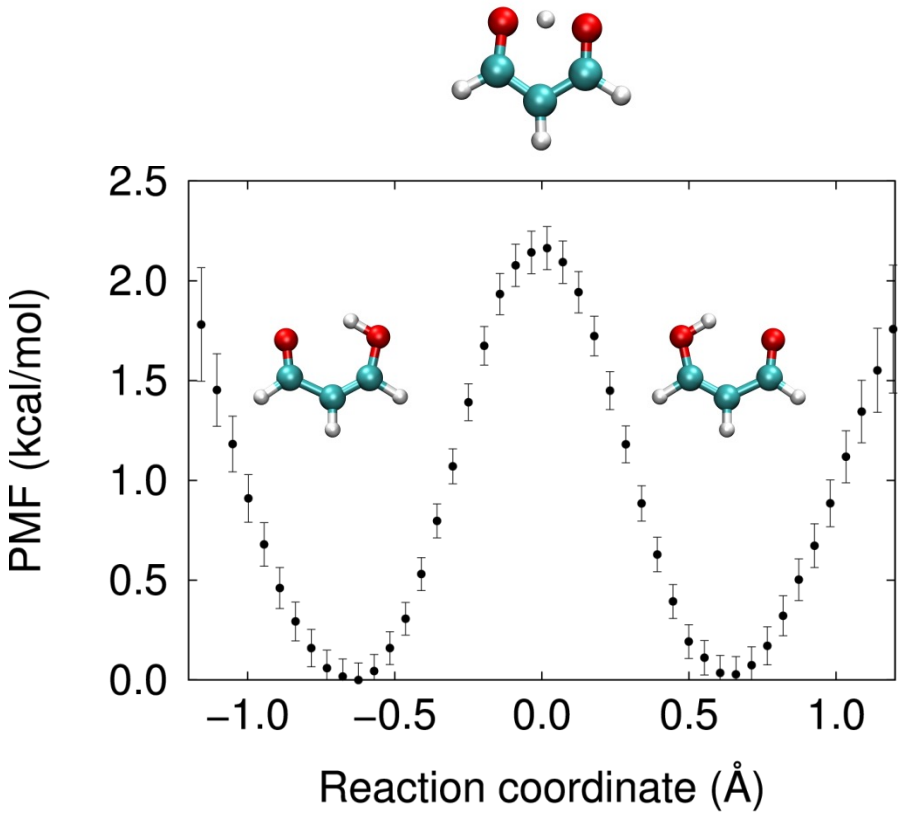




Simulation and movie by Y. Mori

Potential of Mean Force（平均力ポテンシャル）

Y. Mori & Y.O., *Phys. Rev. E* **87**, 023301 (2013).



平均力ポテンシャル
Potential of Mean Force: PMF

$$\mathcal{F}(\xi) \propto -k_{\mathrm{B}}T \ln\left\{\frac{1}{Q}\int dr\ \delta\left[\xi'(r)-\xi\right] e^{-\beta E(r)}\right\}$$

平均力ポテンシャルを再重法により計算
Reweighting by Multistate Bennett Acceptance Ratio estimator: MBAR*

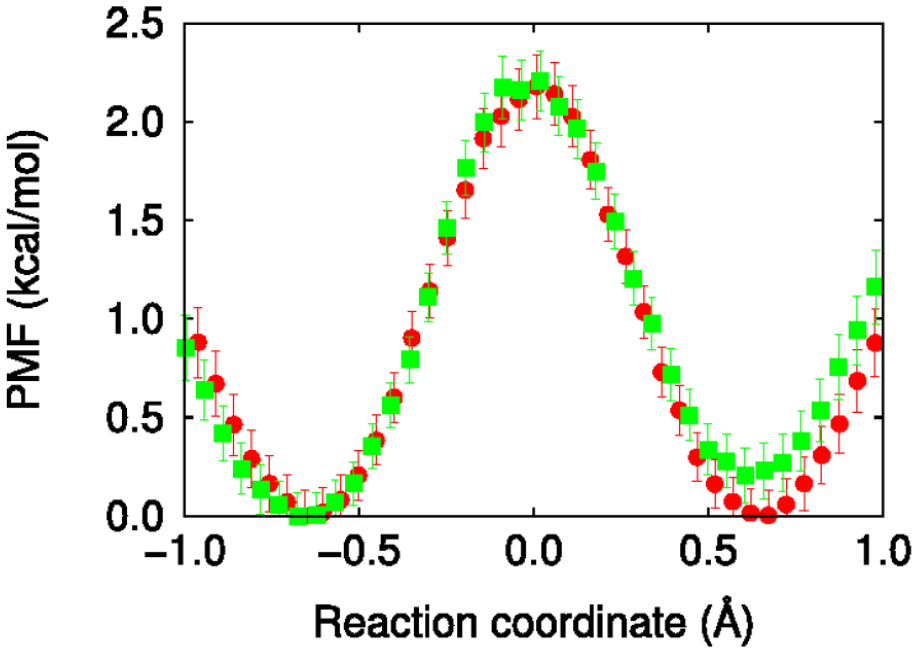
$$\hat{A}_a = \sum_{j=1}^{K}\sum_{n=1}^{N_j} \frac{A(\boldsymbol{x}_{jn})\exp[\hat{f}_a - \beta E(\boldsymbol{x}_{jn})]}{\sum_{k=1}^{K} N_k \exp\{\hat{f}_k - \beta[E(\boldsymbol{x}_{jn}) + U_k(\boldsymbol{x}_{jn})]\}}$$

$$\hat{f}_i = -\ln \sum_{j=1}^{K}\sum_{n=1}^{N_j} \frac{\exp[-\beta E(\boldsymbol{x}_{jn})]}{\sum_{k=1}^{K} N_k \exp\{\hat{f}_k - \beta[E(\boldsymbol{x}_{jn}) + U_k(\boldsymbol{x}_{jn})]\}}$$

* M. R. Shirts and J. D. Chodera, *J. Chem. Phys.* **129**, 124105 (2008).

Potential of Mean Force（平均力ポテンシャル）

Y. Mori & Y.O., *Phys. Rev. E* **87**, 023301 (2013).



red : STUS

green : conventional US

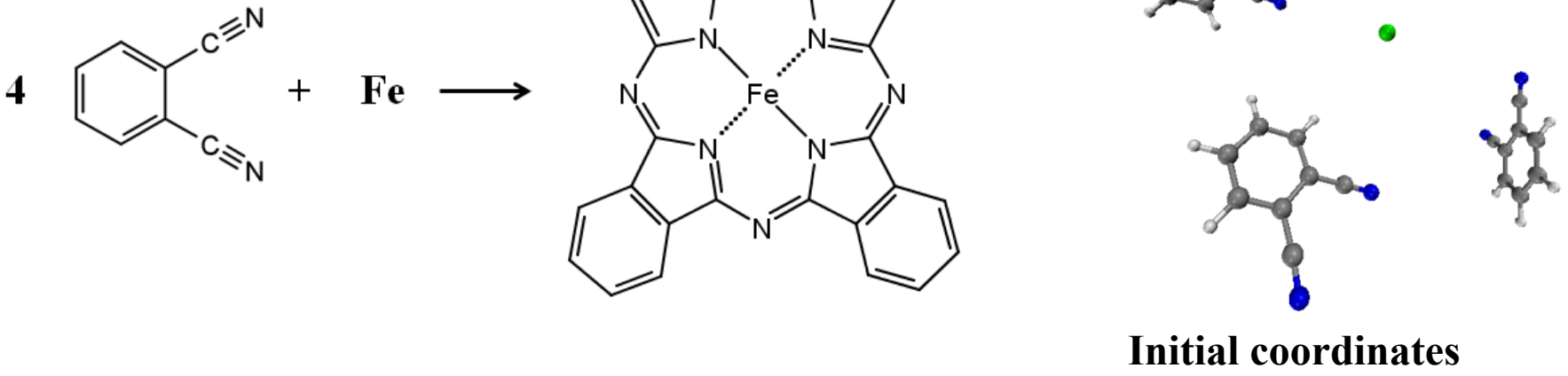
Proton transfer energy barriers for malonaldehyde

Method	Geometry	Barrier (kcal/mol)
MP2	MP2	2.8
CCSD (T)	MP2	3.9
CCSD (T)	VSXC	3.9
BLYP	BLYP	2.0
PBE	PBE	0.9
B3LYP	B3LYP	3.0
B1B95	B1B95	2.8
PBE1PBE	PBE1PBE	2.0
VSXC	VSXC	3.7

S. Sadhukhan, D. Munoz, C. Adamo, & G.E. Scuseria, *Chem. Phys. Lett.* **306** (2008) 83.

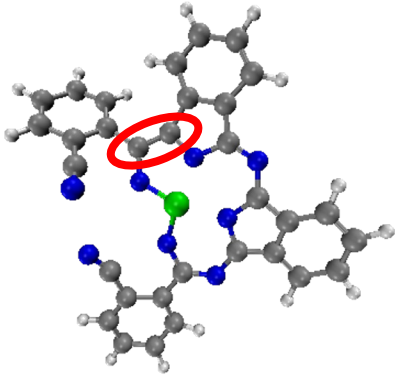
Formation of one phthalocyanine from four phthalonitriles and one iron atom with REUS and Density Functional Tight Binding (DFTB) method

S. Ito, S. Irle, and Y.O., *Comput. Phys. Commun.* **204**, 1-10 (2016);
S. Ito, Y. Wang, Y.O. & S. Irle, *J. Chem. Phys.* **149**, 072332 (2018).



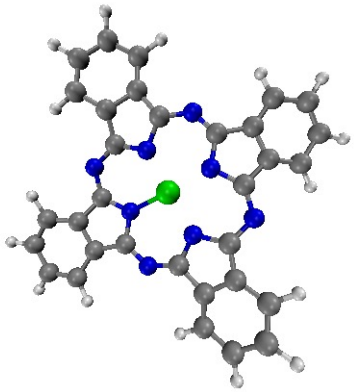
Initial coordinates

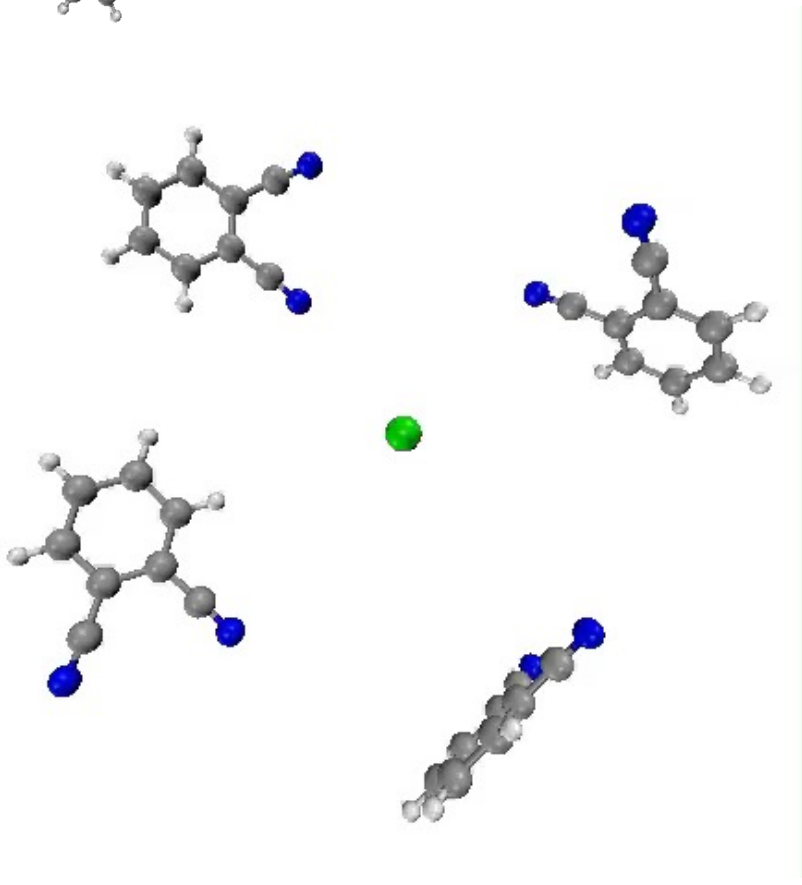
- A simulation was already tried with only DFTB, but the phthalocyanine was not formed.
- Therefore, we carried out the simulation in this mechanism with REUS and DFTB by DFTB+/REM.

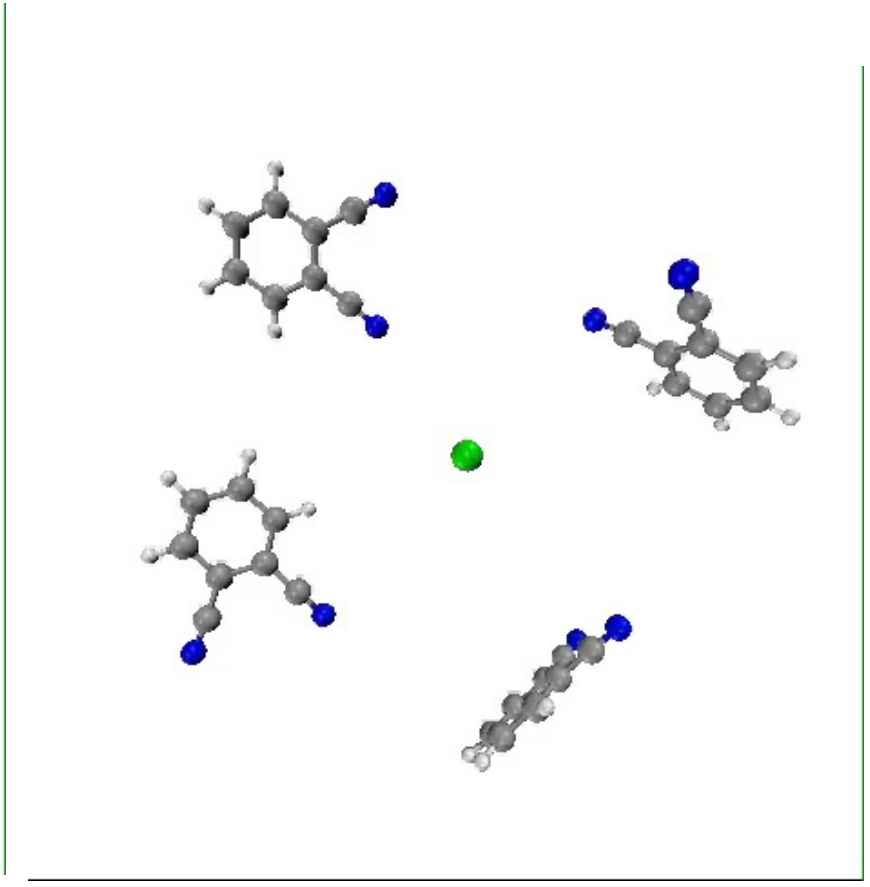


Movies of trajectories

Simulation and movie by S. Ito



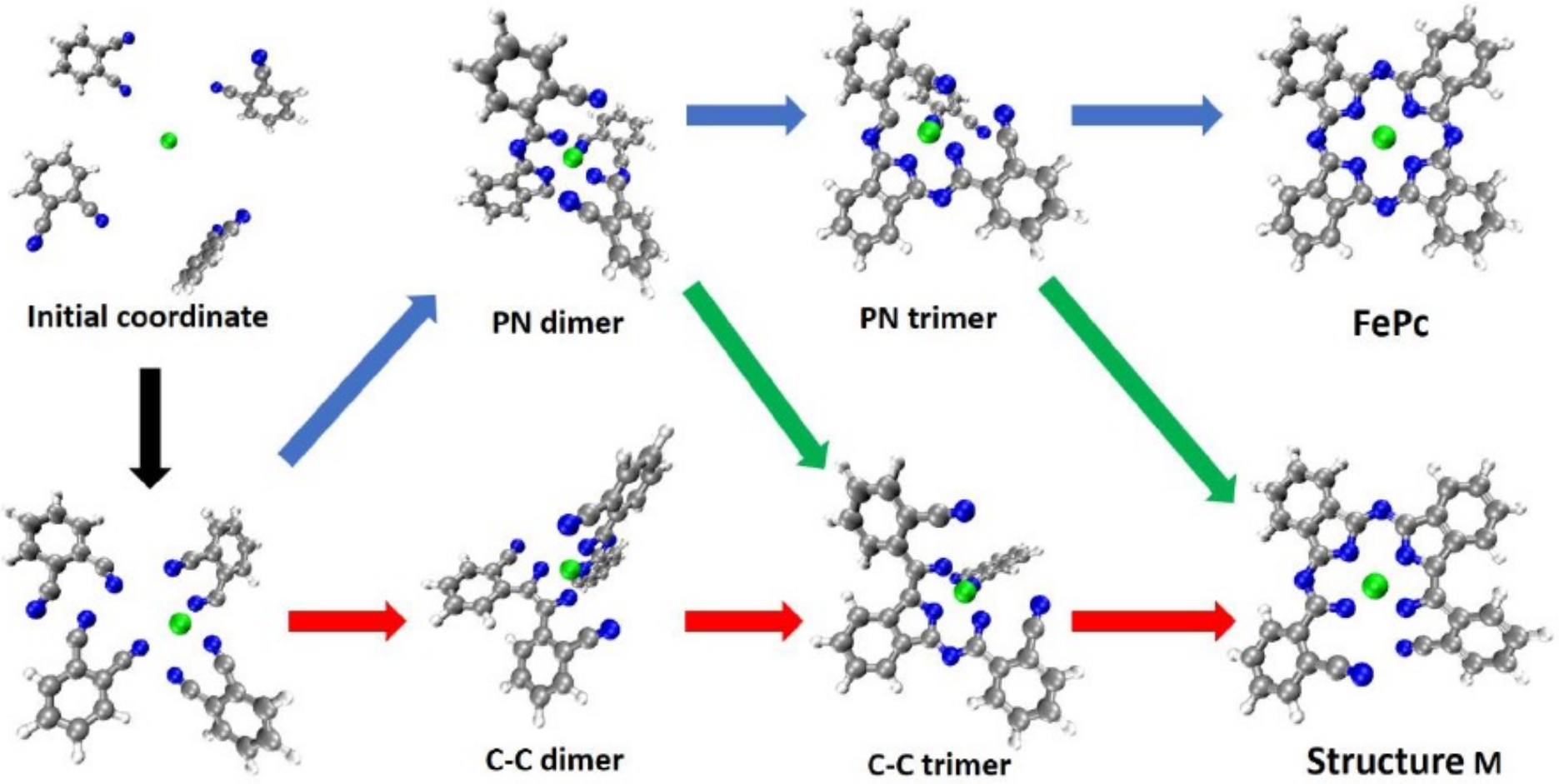




Product: Structure M

Product: Structure P (Phthalocyanine)

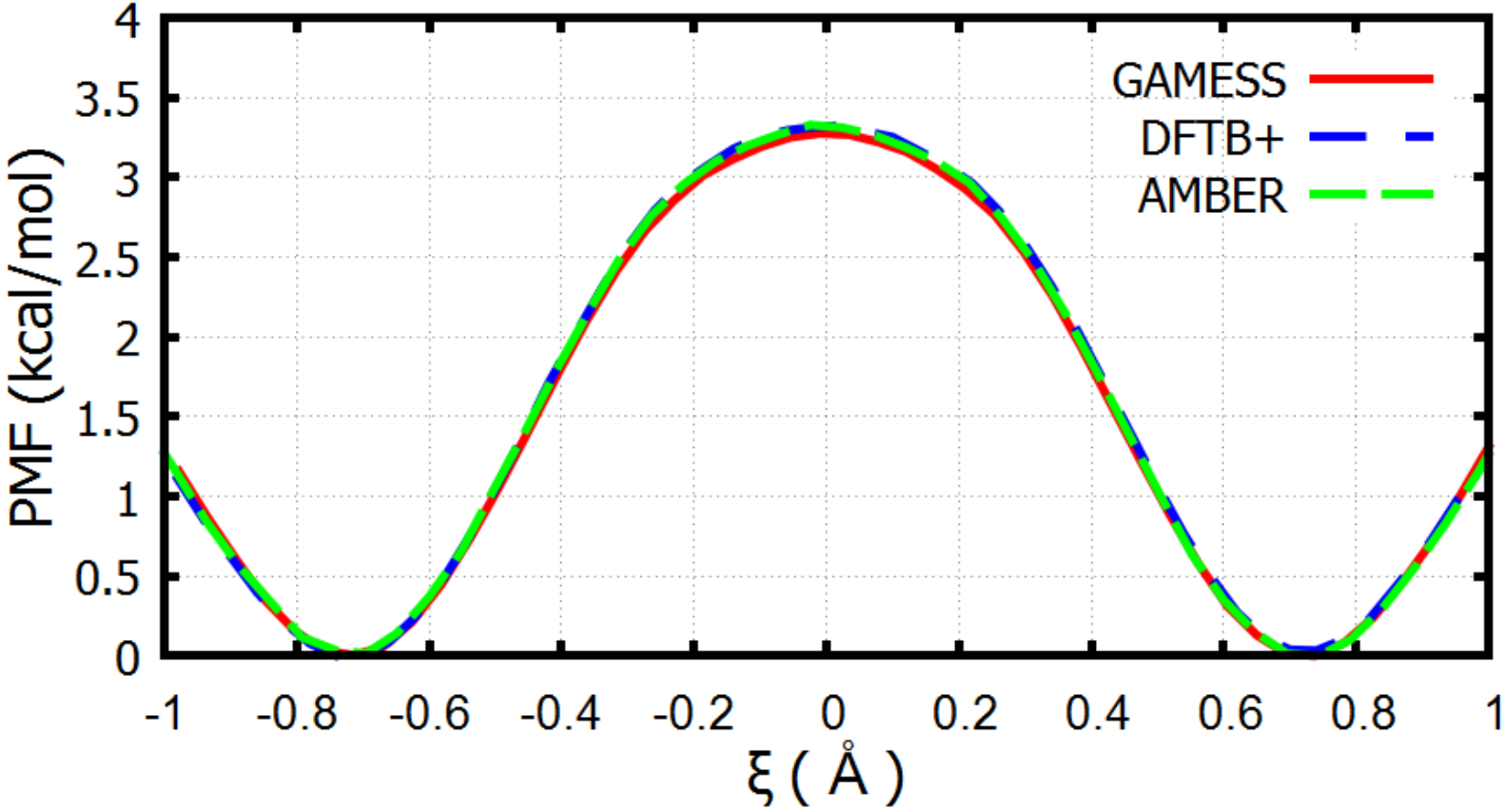
Conjecture on scheme for forming phthalocyanine



- S. Ito, S. Irle, and Y.O., *Comput. Phys. Commun.* **204**, 1-10 (2016);
- S. Ito, Y. Wang, Y.O. & S. Irle, *J. Chem. Phys.* **149**, 072332 (2018).

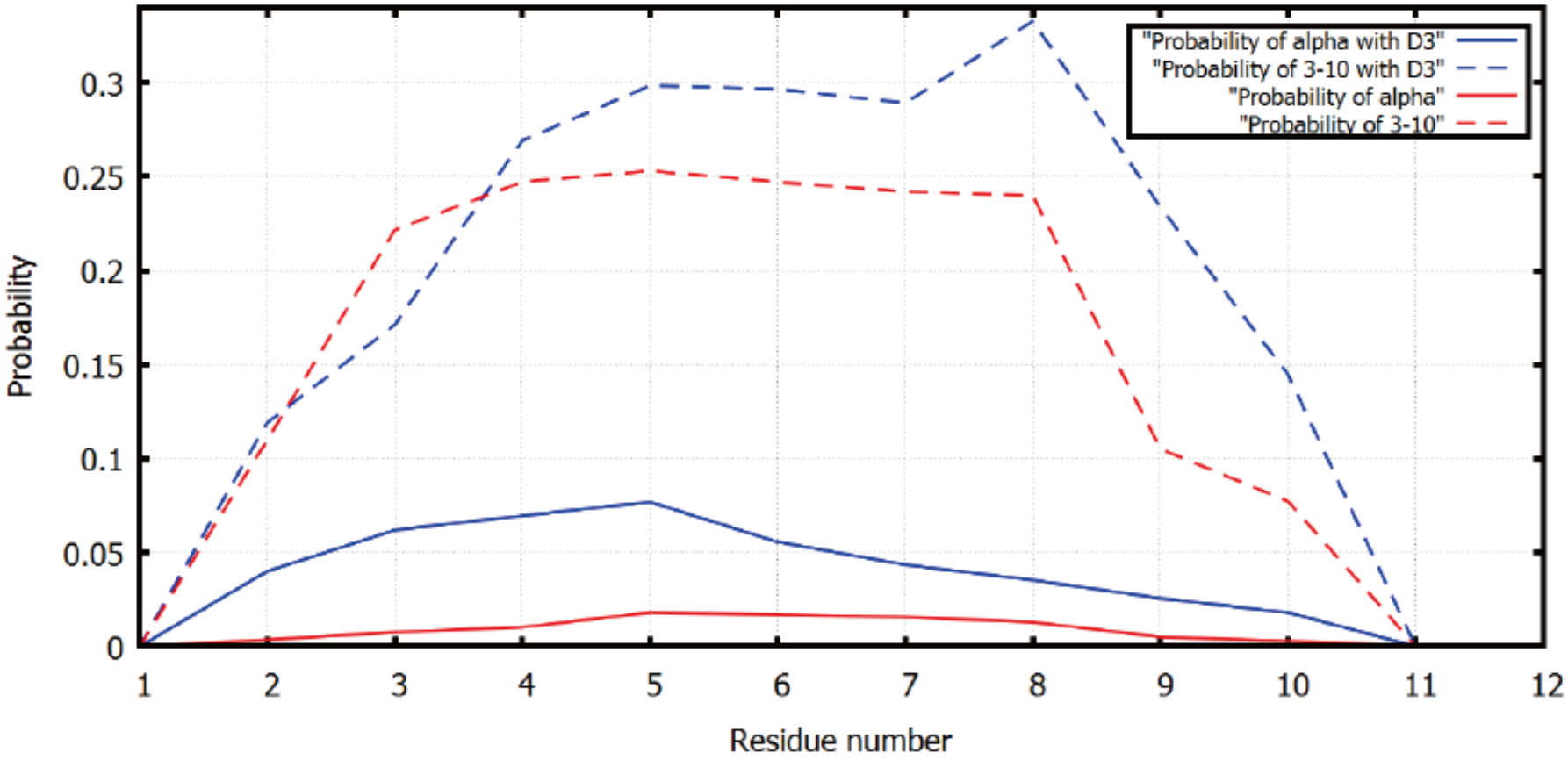
Implementation of REUS to GAMESS

- S. Ito, Fedorov, Y.O. & S. Irle, *Comput. Phys. Commun.* **228**, 152 (2018).



DFTB-REMD Simulations of Deca-Alanine with/without D3 Dispersion Corrections

- S. Ito, Fedorov, Y.O. & S. Irle, *Comput. Phys. Commun.* **228**, 152 (2018).



Red: DFTB; Blue: DFTB with D3 dispersion corrections

SUMMARY

We have shown that generalized-ensemble algorithms are particularly suitable for biomolecular simulations.

COLLABORATORS (Former Group Members)

Ayori MITSUTAKE [IMS → Keio Univ. → Meiji Univ.]
Ulrich H.E. HANSMANN [IMS → Michigan Tech. Univ. → Univ. of Oklahoma]
Masato MASUYA [IMS → Kagoshima Univ.]
Yuji SUGITA [IMS → Univ. Tokyo → RIKEN]
Takao YODA [IMS → Nagahama Inst. Bio-Science]
Hiroshi NOGUCHI [IMS → Juelich → Univ. of Tokyo]
Hisashi OKUMURA [IMS → Nagoya Univ. → Rutgers Univ. → IMS]
Satoru G. ITOH [IMS → Nagoya Univ. → NIH → IMS]
Yoshitake SAKAE [IMS → Hiroshima Univ. → Nagoya Univ. → Temple Univ. → RIST]
Hironori KOKUBO [IMS → Univ. of Houston → Takeda Pharm. → Chugai Pharm.]
Yoshiharu MORI [Nagoya Univ. → IMS → Kitasato Univ. → Kobe Univ.]
Tetsuro NAGAI [Nagoya Univ. → Ritsumeikan Univ. → Univ. of Tokyo → Fukuoka Univ.]
Ryo URANO [Nagoya Univ. → Boston Univ. → Nagoya Univ. → Okayama Univ.]
Shuichiro TSUKAMOTO [Nagoya Univ.]
Shingo ITO [Nagoya Univ. → Boston Univ. → RIKEN]
Yuhei TACHI [Nagoya Univ. → Deloitte Analytics]
Daiki MATSUBARA [Nagoya Univ. → RIKEN]
Takuya HAYASHI [Nagoya Univ. → Kureha]
Kohei NODA [Nagoya Univ. → JSR]
Ryuga SUGANO [Nagoya Univ. → NHK]

COLLABORATORS

Hikaru KAWAI [Cornell Univ. → Univ. of Tokyo → KEK → Kyoto Univ. → National Taiwan Univ.]
Takeshi KIKUCHI [Cornell Univ. → Univ. of Tokyo → Ciba-Geigy → Ritsumeikan Univ.]
Masataka FUKUGITA [Kyoto Univ. → Univ. of Tokyo]
Takashi NAKAZAWA [Nara Women's Univ.]
Fumio HIRATA [IMS → Ritsumeikan Univ. → Toyota RIKEN]
Masahiro KINOSHITA [Kyoto Univ.]
Jose ONUCHIC [Univ. of San Diego → Rice Univ.]
Bernd A. BERG [Florida State U.]
Chizuru MUGURUMA [Chukyo Univ.]
Tomoyuki HIROYASU [Doshisha Univ.]
Mitsunori MIKI [Doshisha Univ.]
Katsuya ISHII [Nagoya Univ.]
Takahiro KATAGIRI [Nagoya Univ.]
Satoshi OHSHIMA [Nagoya Univ.]
Ichiro TAKAHASHI [Nagoya Univ.]
Toshimasa TANAKA [Takeda Pharm.]
Ryuichi UEOKA [Sojo Univ.]
Wolfhard JANKE [Univ. of Leipzig]
John STRAUB [Boston Univ.]
Takayoshi SUZUKI [Kyoto Prefect. Univ. of Medicine→ Osaka Univ.]
Yukihiro ITOH [Kyoto Prefect. Univ. of Medicine→ Osaka Univ.]
Stephan IRLE [Nagoya Univ. → Oak Ridge National Lab.]
Dmitri G. FEDOROV [AIST]