

2006年4月4-5日

第2回「計算科学による新たな知の発見・統合・創出」シンポジウム
計算科学の戦略と次世代スーパーコンピュータ(つくば国際会議場)

超高速計算機による タンパク質の立体構造予測

岡本祐幸 (Yuko OKAMOTO)

名古屋大学・大学院理学研究科・物理学教室

e-mail: okamoto@phys.nagoya-u.ac.jp

URL: <http://jegog.phys.nagoya-u.ac.jp/tb/>

バイオサイエンス分野における 計算科学の次世代問題

1. 球状蛋白質の立体構造予測、機能予測
2. 膜蛋白質の立体構造予測、機能予測
3. アミロイド病(狂牛病、アルツハイマー病など)の発病機構
4. 薬品設計(創薬)問題

計算手法

1. 分子シミュレーション(分子動力学法、モンテカルロ法)
特に、拡張アンサンブル法
2. 溶媒理論
3. 電子状態理論
4. 1., 2., 3.の合体

2003年4月 ヒトゲノム解読完了宣言

約3万個の「遺伝子」(ORF)

ポストゲノム時代の最重要課題

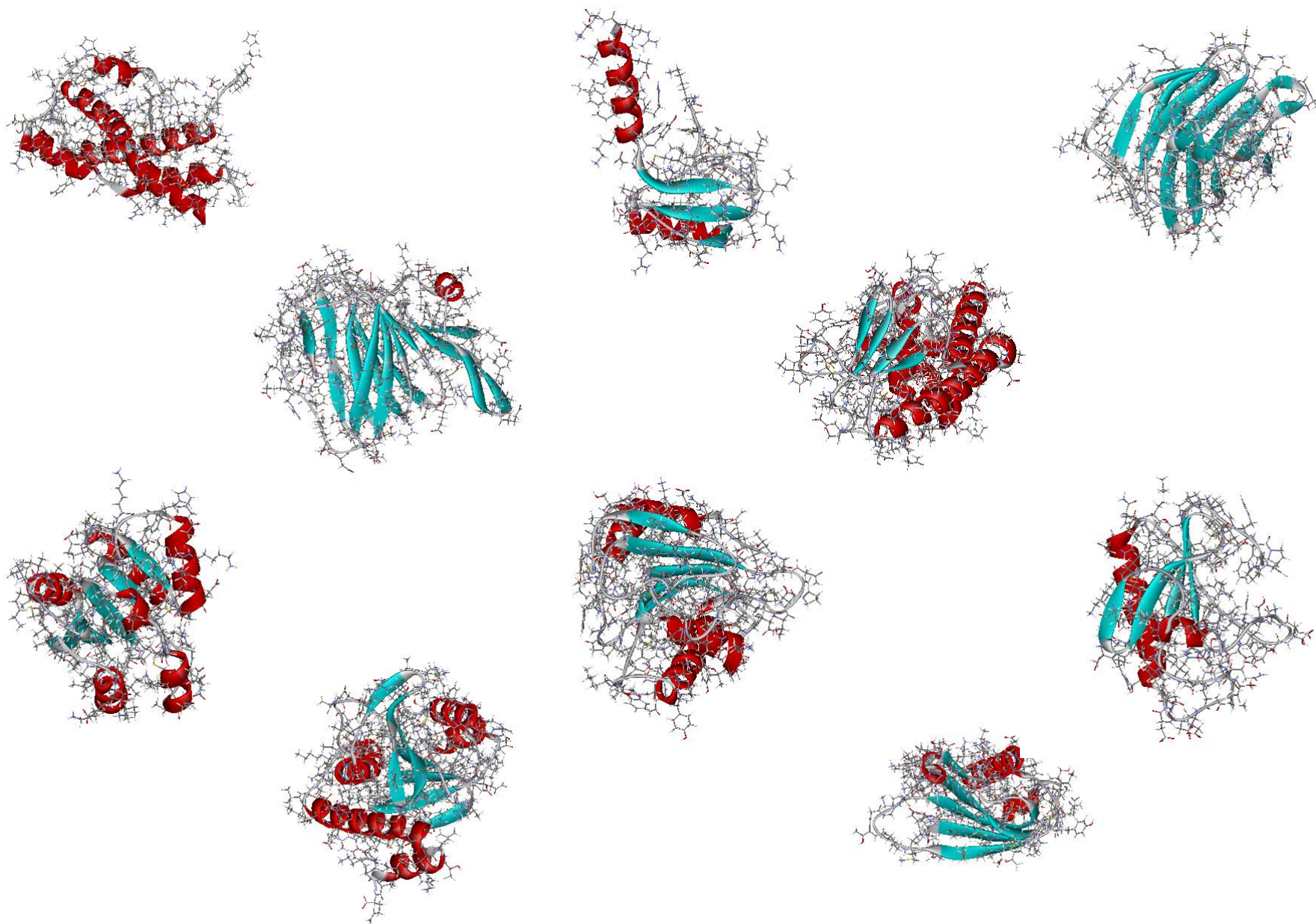
ヒトゲノム計画によって決定された、DNAの塩基配列
(すなわち、蛋白質のアミノ酸配列)の情報から、
それぞれの蛋白質がどういう立体構造(形)をして
いて、どういう機能(働き)をするかを探ること。

PDB (Protein Data Bank)

2006年1月3日現在

34,420個の構造が登録されている

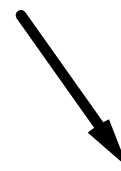
PDBに登録されているタンパク質の立体構造の例



戦略

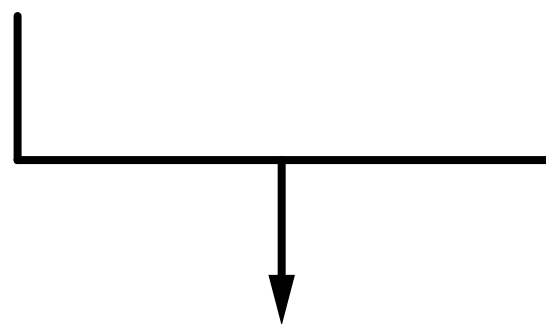
溶媒理論

電子状態理論



タンパク質系の
エネルギー関数

シミュレーション法



第一原理からのタンパク質の
立体構造予測

我々が開発(導入)してきた分子シミュレーション手法

1. 徐冷法 (simulated annealing; SA) Kawai, Kikuchi & Y.O. (1989)
2. マルチカノニカル法 (MUCA) Hansmann & Y.O. (1993)
3. マルチカノニカルMD法 (MUCAMD) Hansmann, Y.O. & Eisenmenger (1996)
4. 焼き戻し法 (simulated tempering; ST) Hansmann & Y.O. (1997)
5. 1/k法 Hansmann & Y.O. (1997)
6. Tsallis 統計に基づく拡張アンサンブル法 Hansmann & Y.O. (1997)
7. レプリカ交換MD法 (REMD) Sugita & Y.O. (1999)
8. 多変数レプリカ交換法 (MREM) Sugita, Kitao & Y.O. (2000)
9. レプリカ交換自由エネルギー摂動法およびレプリカ交換アンブレラサンプリング法
(REFEP & REUS) Sugita, Kitao & Y.O. (2000)
10. レプリカ交換マルチカノニカル法 (REMUCA) Sugita & Y.O. (2000)
11. マルチカノニカルレプリカ交換法 (MUCAREM) Sugita & Y.O. (2000)
12. レプリカ交換焼き戻し法 (REST) Mitsutake & Y.O. (2000)
13. NPT系のレプリカ交換法 Nishikawa, Sugita, Mikami & Y.O. (2000)
14. *Ab initio* レプリカ交換法 Ishikawa, Sugita, Nishikawa & Y.O. (2001)
15. 遺伝的交叉を用いた並列徐冷法 Hiroyasu, Miki, Ogura & Y.O. (2002)
16. マルチオーバーラップ法 Berg, Noguchi & Y.O. (2003)
17. マルチNPT法 (3種: MUBA, MUTH & MUBATH) Okumura & Y.O. (2003)
18. 焼き戻しレプリカ交換法 (STREM) Mitsutake & Y.O. (2004)
19. マルチNPT MD法 (MUBATHMD) Okumura & Y.O. (2004)
20. マルチオーバーラップMD法 Itoh & Y.O. (2004)

緑は他の分野から導入した手法

赤は独自に開発した手法

シミュレーション手法

- シミュレーションがエネルギー極小状態に留まらない強力な手法を用いる

次の手法を適用することを主張してきた

- * **拡張アンサンブル法 (Generalized-Ensemble Algorithms)**

エネルギー空間上のランダムウォークを実現してエネルギー極小状態に留まるのを避けるシミュレーション手法の総称

[例えば、**マルチカノニカル法 (Multicanonical Algorithm)**、

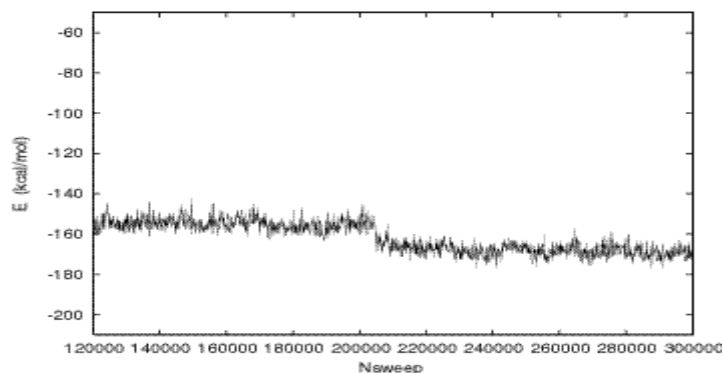
焼き戻し法 (Simulated Tempering)、**レプリカ交換法 (Replica-Exchange Method)**、**1/k Sampling**、**Tsallis 統計**, etc.]

利点

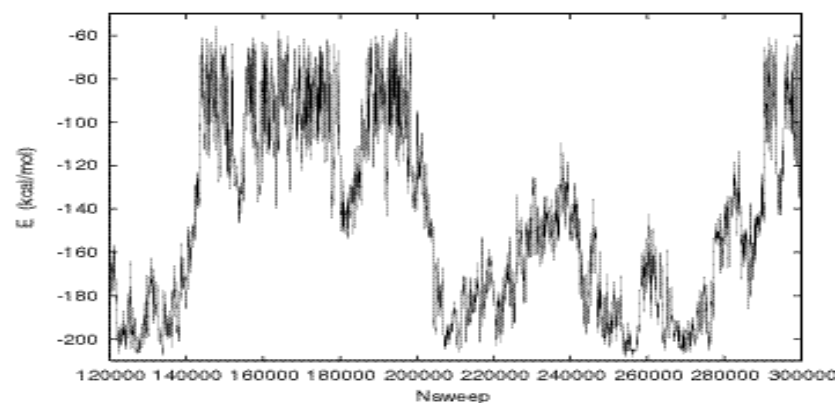
唯一回のシミュレーションの結果から、最小エネルギー状態ばかりでなく、幅広い温度領域における熱力学量を計算できる

17-Residue Helical Peptide (120000-300000 MC Sweeps)

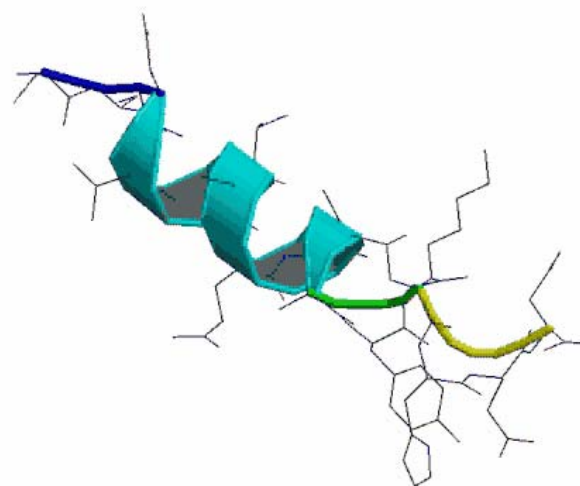
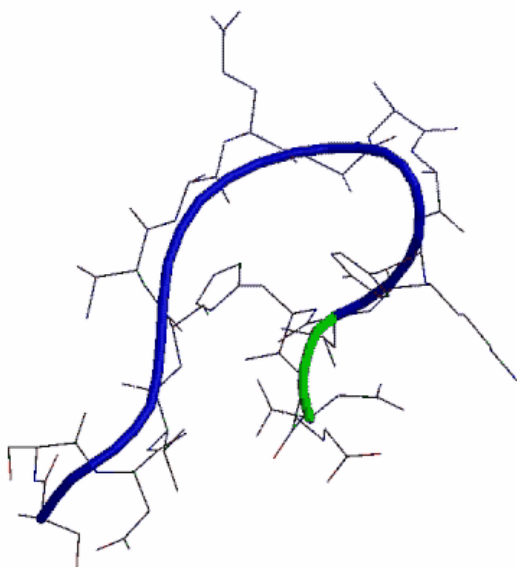
Simulation and movie by A. Mitsutake



Canonical MC: $T = 200$ K



Multicanonical MC

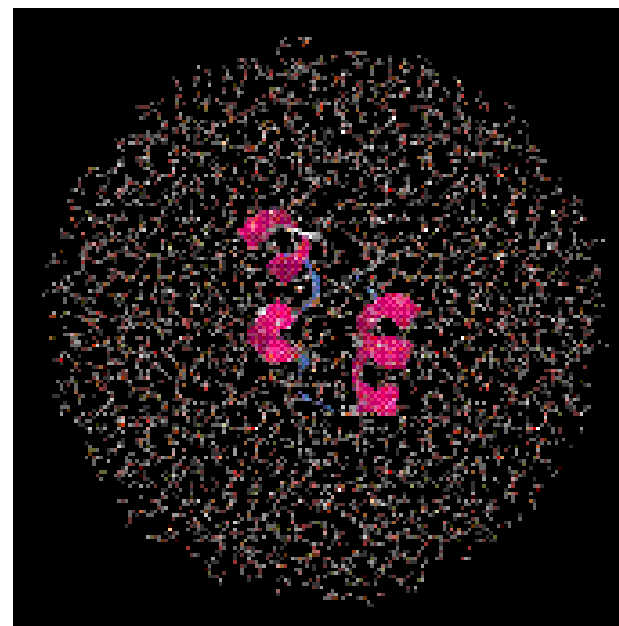


球状タンパク質の立体構造予測

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

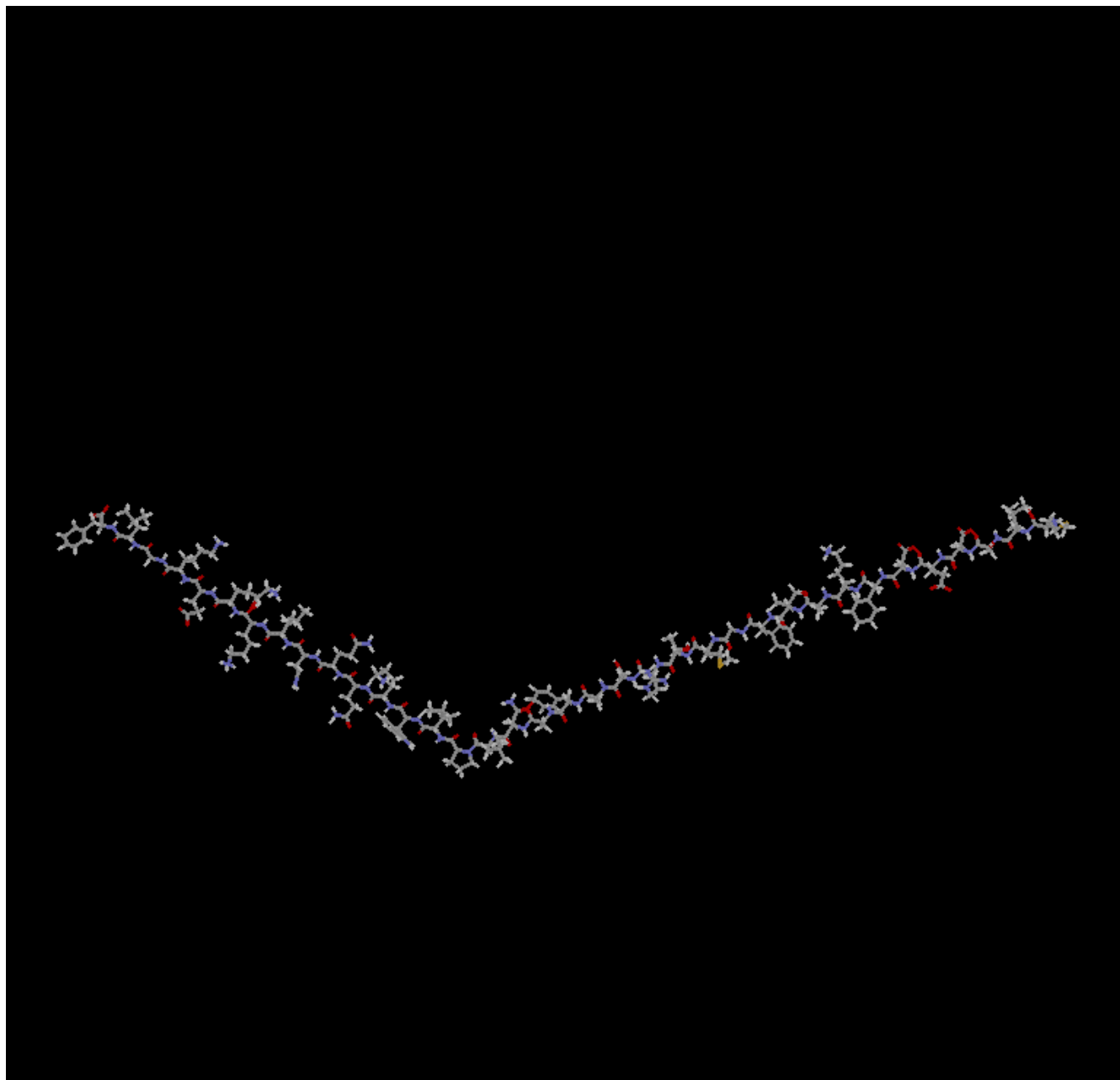
Villin headpiece subdomain
(36 amino acids; 596 atoms)
in sphere of water of radius
30 Å (3513 water molecules);
altogether 11,135 atoms

R = 30Å, 3550wat, 11246atom



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

Initial Conformation



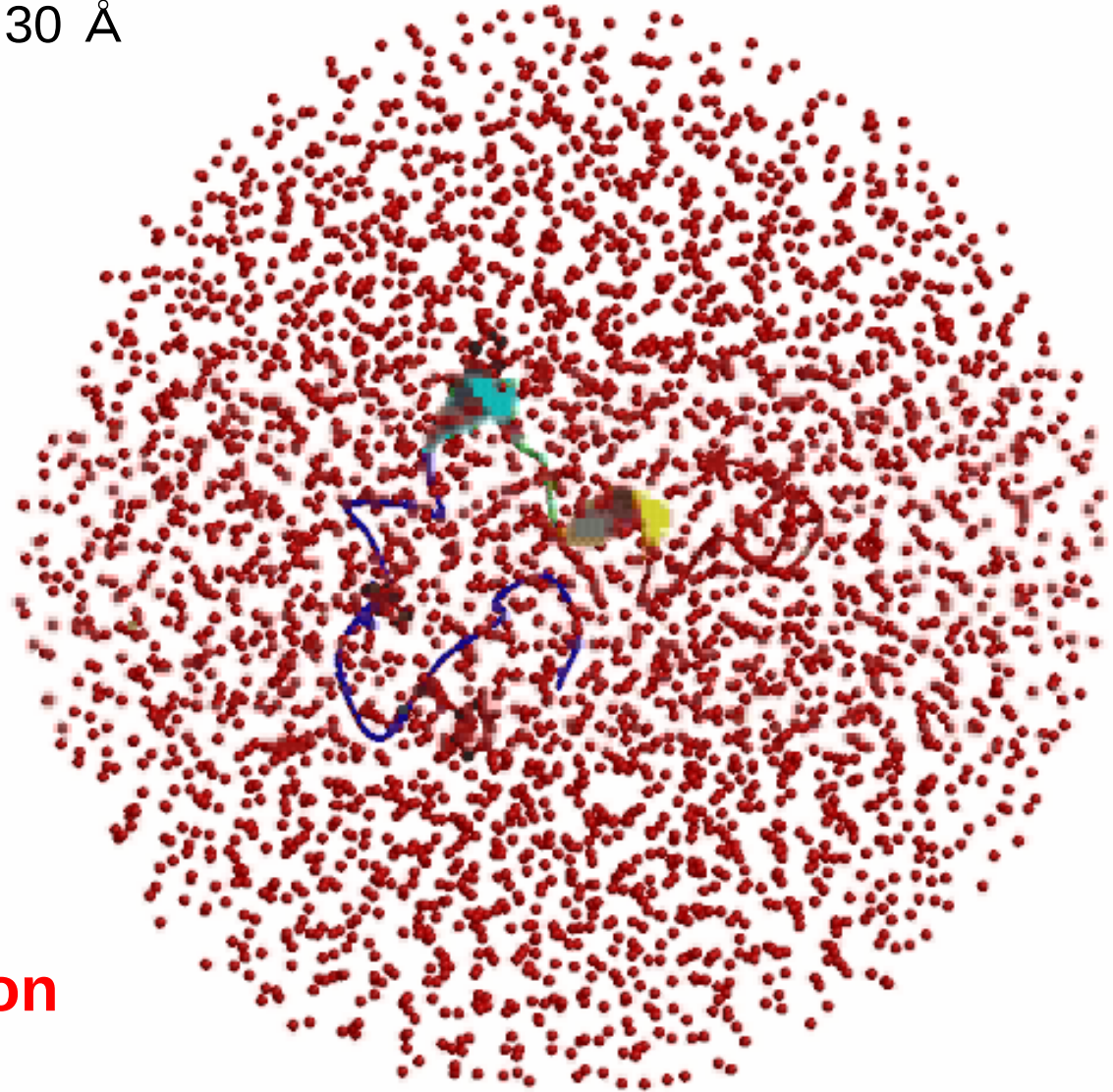
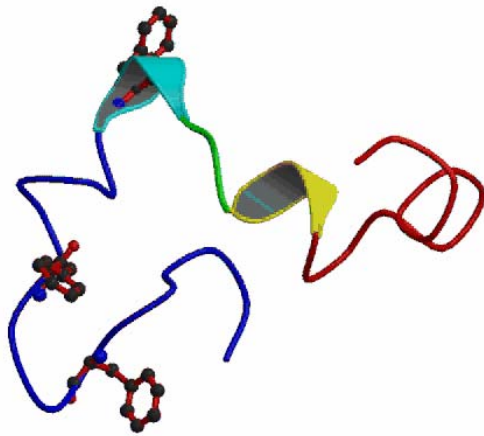
First **REMD** in
gas phase with
96 replicas

Then **REMD** in
water with
96 replicas

Then
MUCAREM in
water with
8 replicas

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

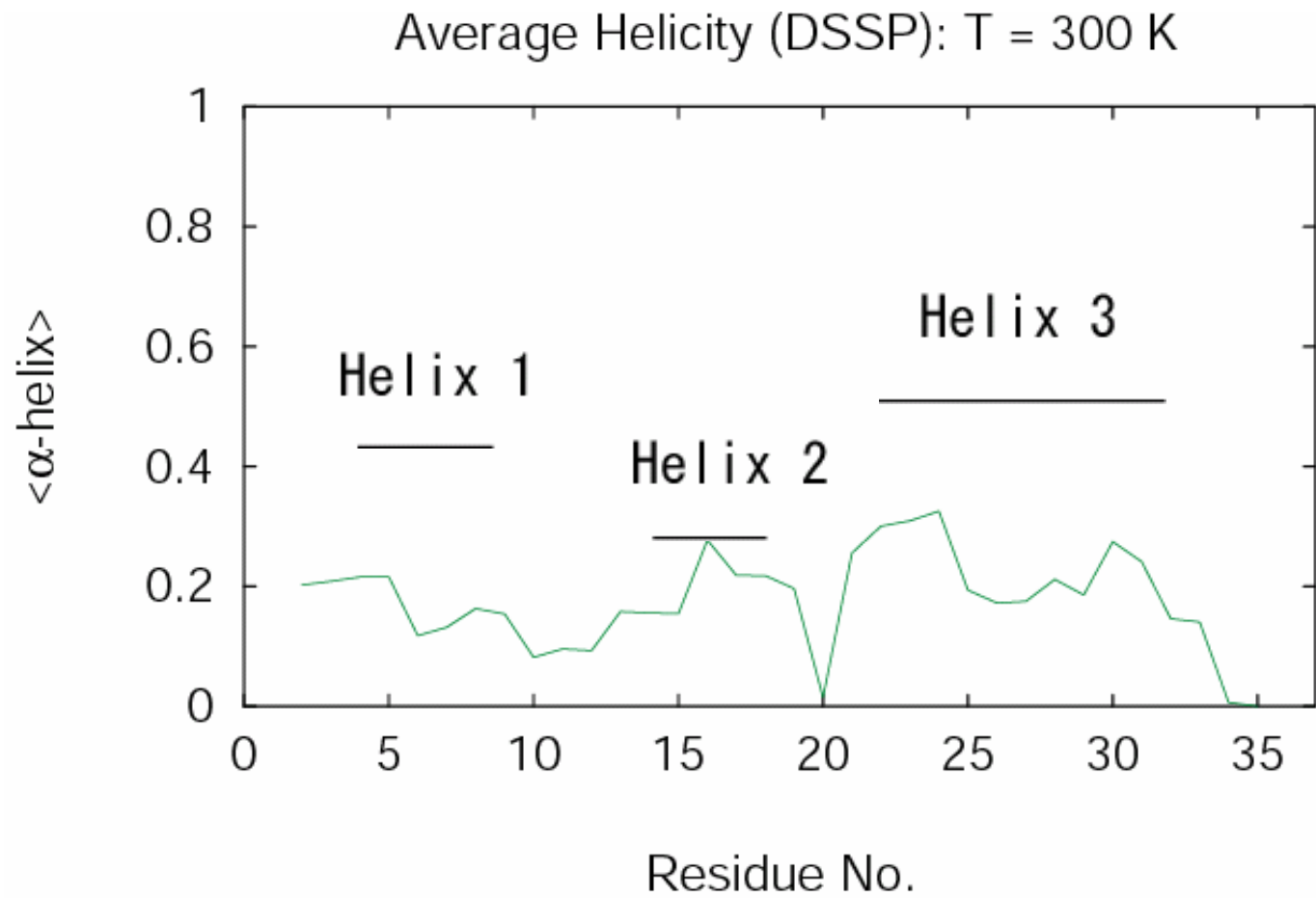
Villin headpiece subdomain
(36 amino acids; 596 atoms)
in sphere of water of radius 30 Å
(3513 water molecules);
altogether 11,135 atoms



MUCAREM simulation

Native Helices:

1. Res. No. 4 ~ 8
2. Res. No. 15 ~ 18
3. Res. No. 23 ~ 32

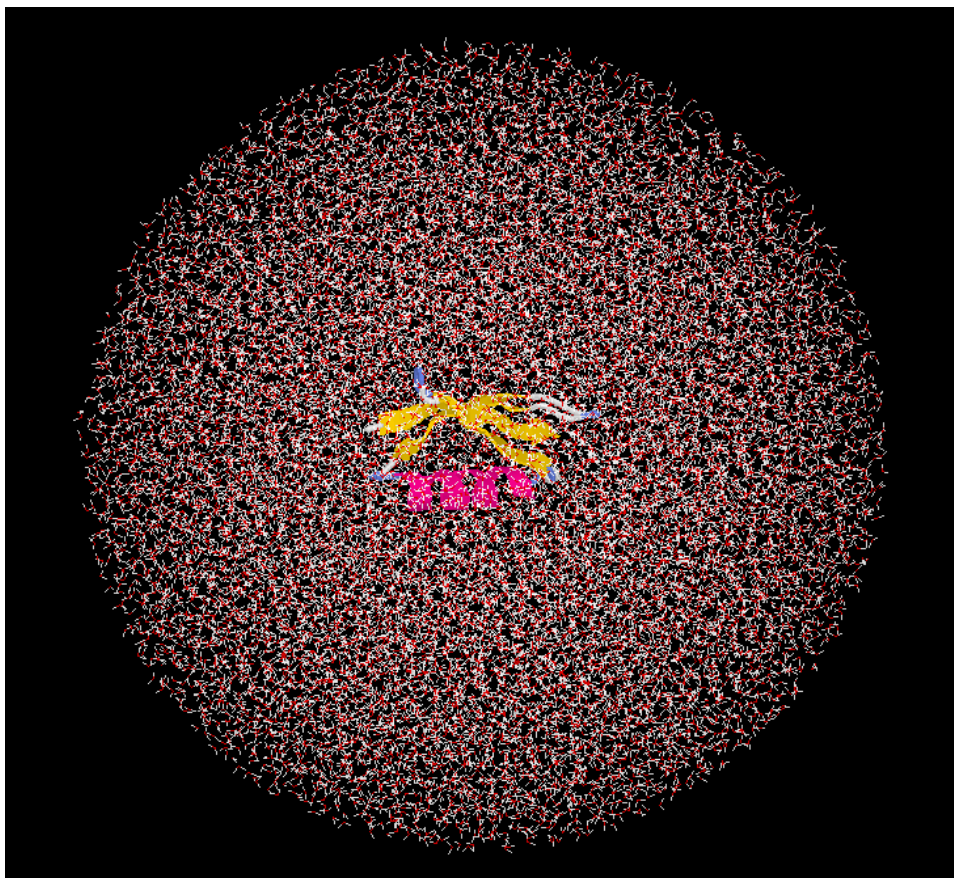


T. Yoda, Y. Sugita & Y.O., in preparation.

Protein G (56 amino acids; 786 atoms) 地球シミュレータ利用

sphere of radius 50 Å (number of water molecules: 17,187)

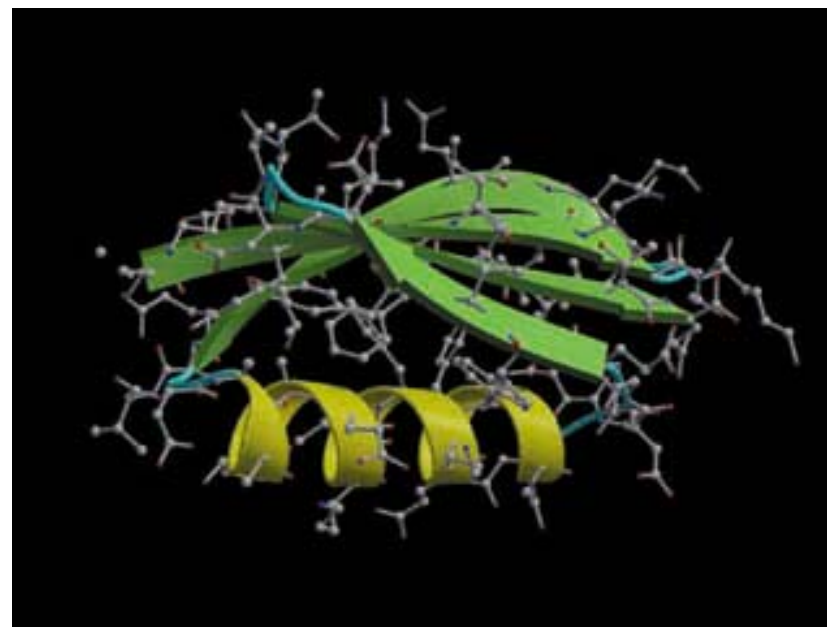
Total number of atoms in the system: 52,416)



Initial Conformation:

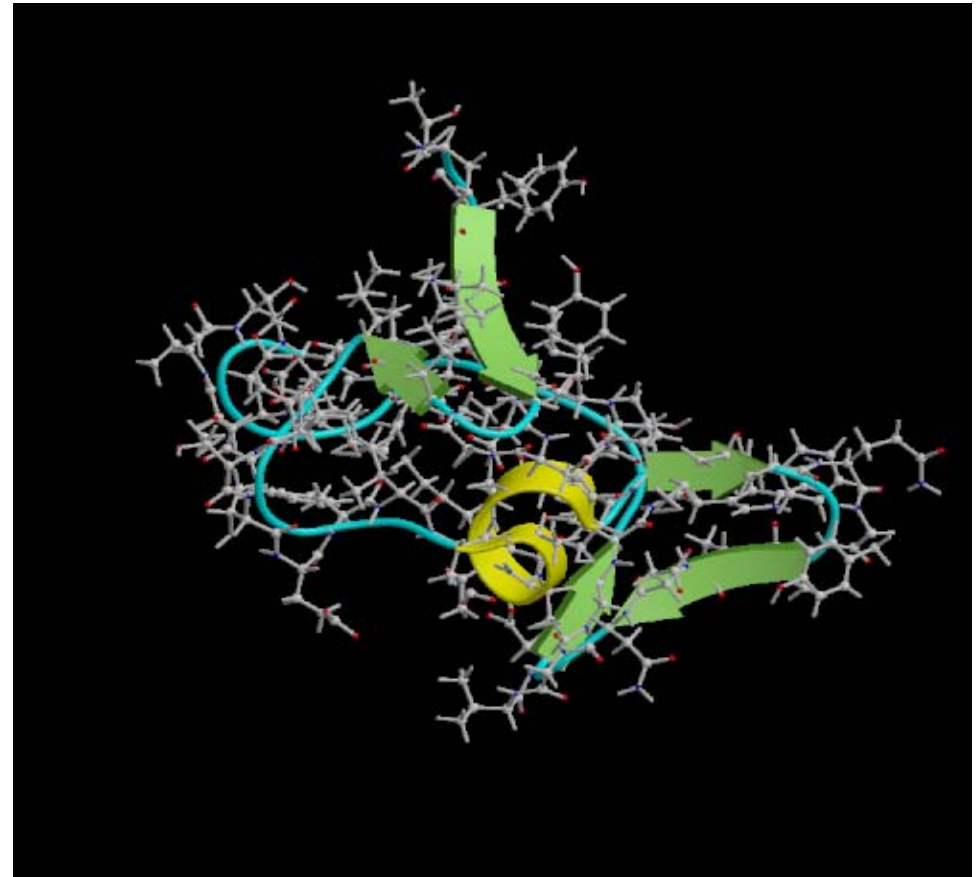
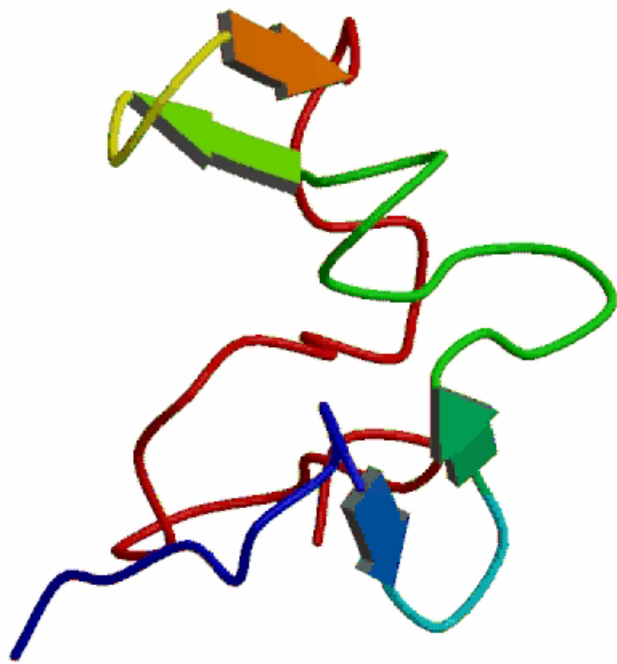


Native Conformation:



A. Mitsutake, T. Yoda, Y. Sugita, T. Nishikawa, Y. Sakae & Y.O., in preparation.

Snapshots 地球シミュレータ利用



A. Mitsutake, T. Yoda, Y. Sugita, T. Nishikawa, Y. Sakae & Y.O., in preparation.

膜タンパク質の立体構造予測

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).

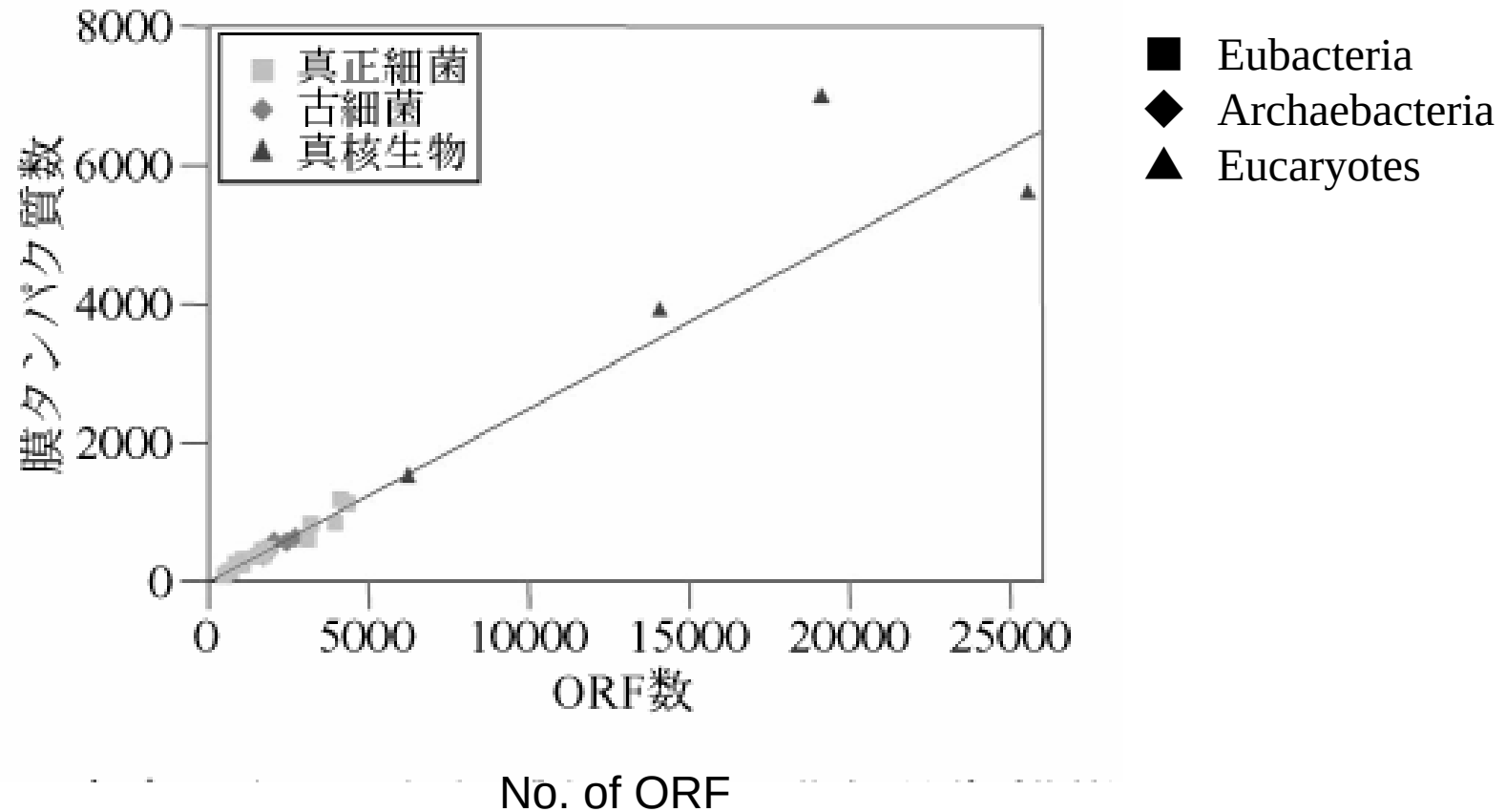
H. Kokubo & Y.O., *Chem. Phys. Lett.* **392**, 168-175 (2004).

PDB (Protein Data Bank) Database

2006年1月3日現在

34,420個の構造が登録されている

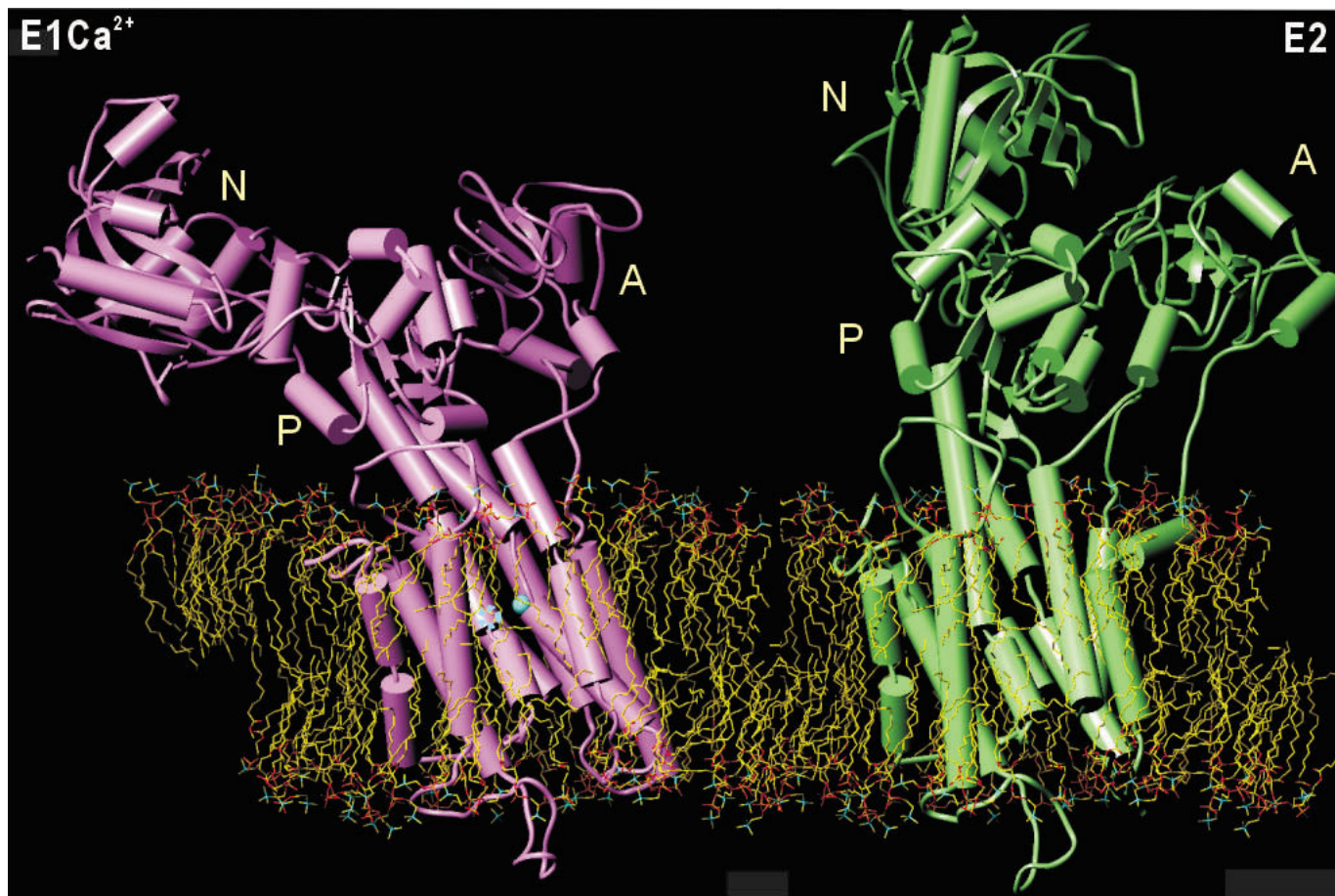
No. of Membrane Proteins



Ref.: 美宅成樹, 生物物理 **42**, 104 (2002).

See also, A. Krogh et al., *J. Mol. Biol.* **305**, 567 (2001).

Ca ATPase

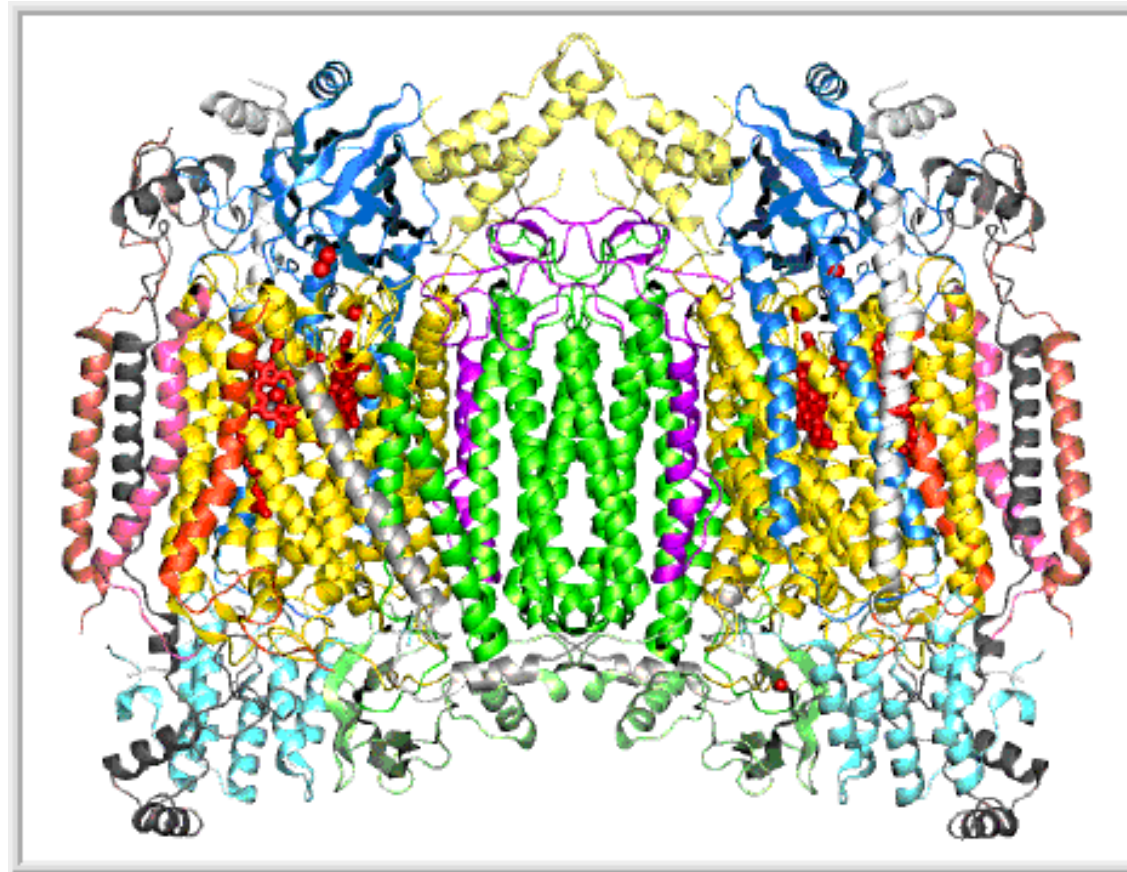


PDB code: 1eul

From: 東大分生研豊島研のホームページ

<http://www.iam.u-tokyo.ac.jp/StrBiol/index.html>

Cytochrome c Oxidase



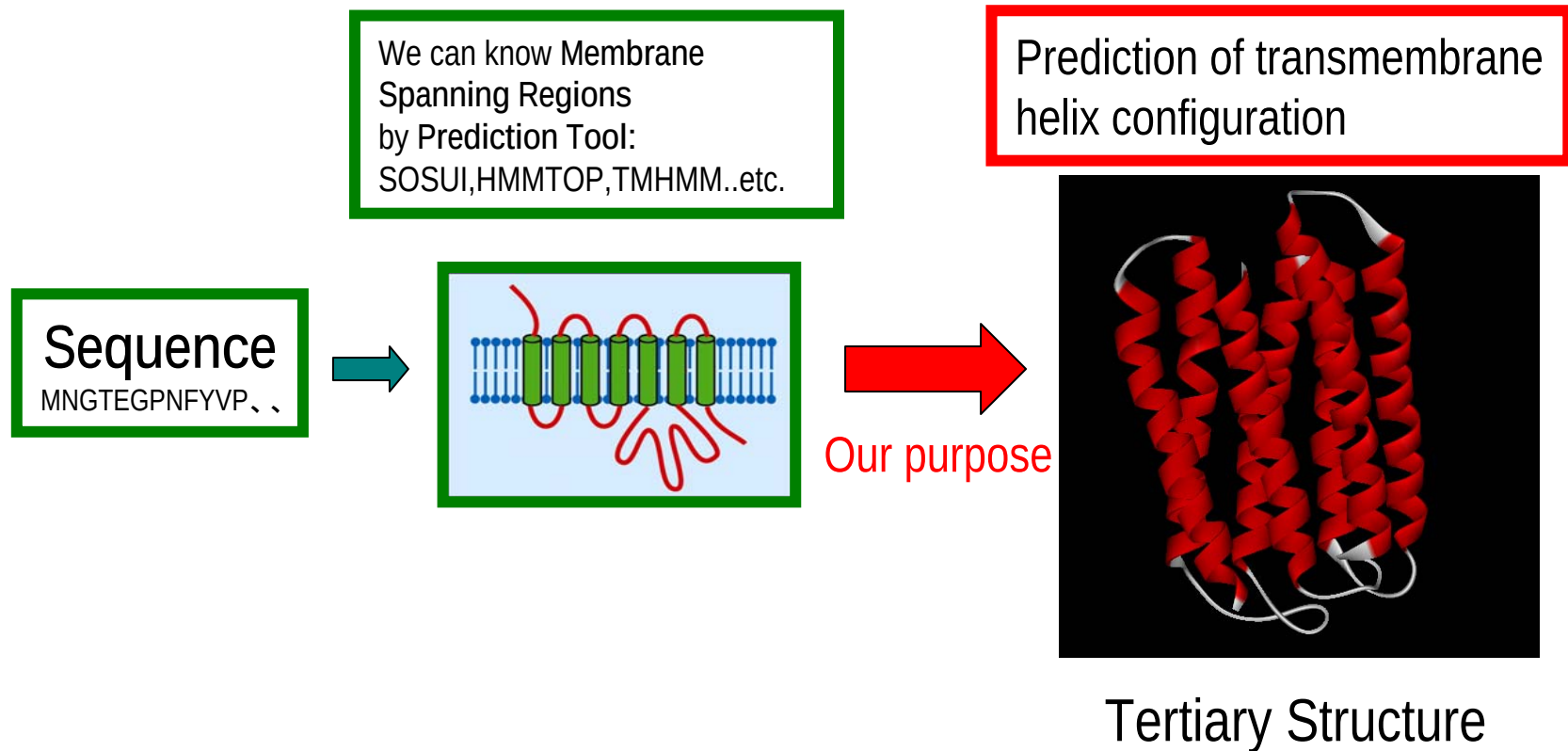
PDB code: 2occ

From: 阪大蛋白研月原研のホームページ

<http://www.protein.osaka-u.ac.jp/crystallography/>

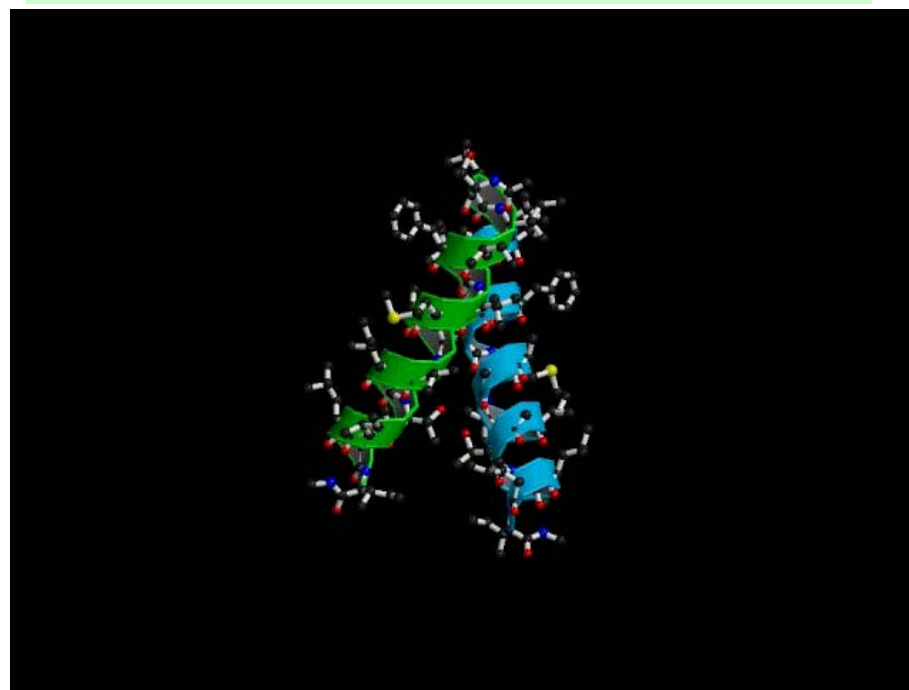
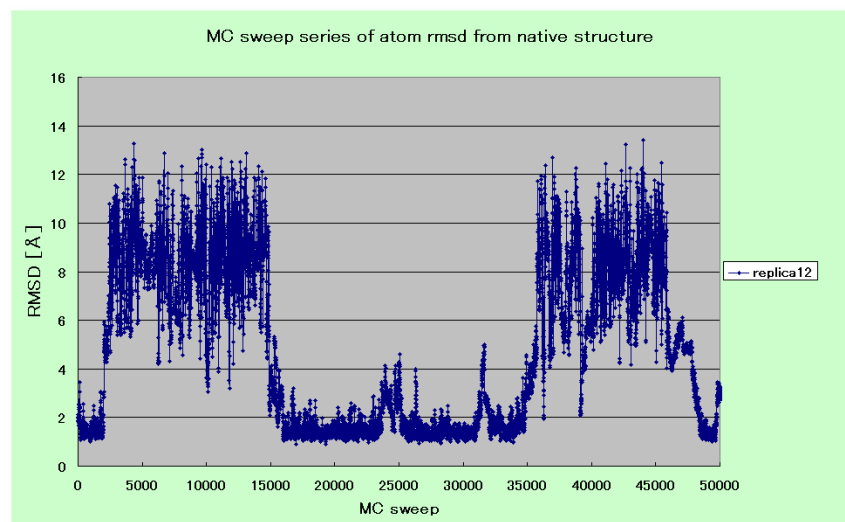
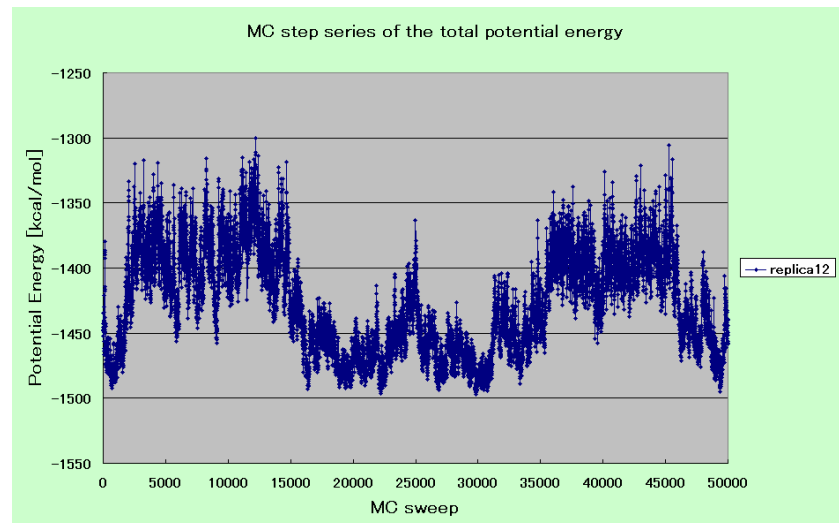
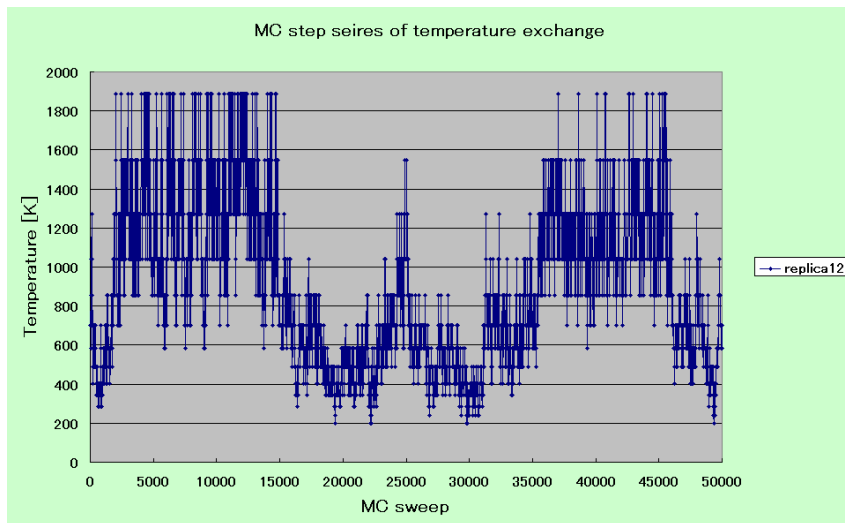
膜タンパク質の立体構造予測

- 膜タンパク質の「アミノ酸配列情報と膜貫通領域部分がどこかという情報」から「3次構造」を計算機シミュレーションによって予測する



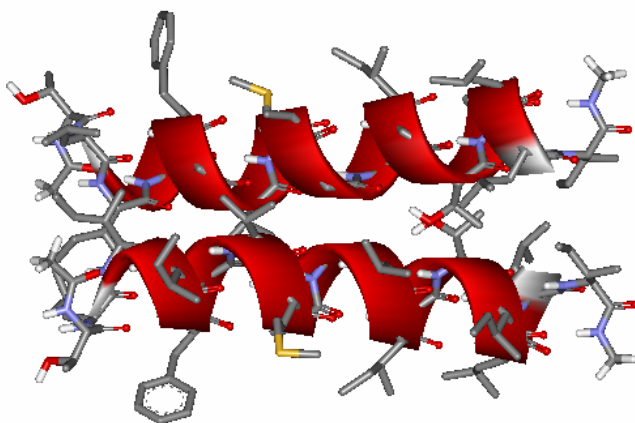
DIMERIC TRANSMEMBRANE DOMAIN OF HUMAN GLYCOPHORIN A

Simulation and movie by H. Kokubo



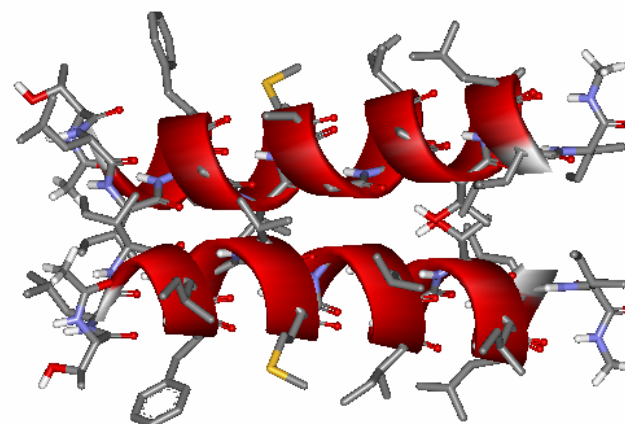
レプリカ交換シミュレーションの結果 実験構造と予測構造の比較

実験で求められた構造



PDB code: 1AFO

レプリカ交換シミュレーションで
予測された最小エネルギー構造

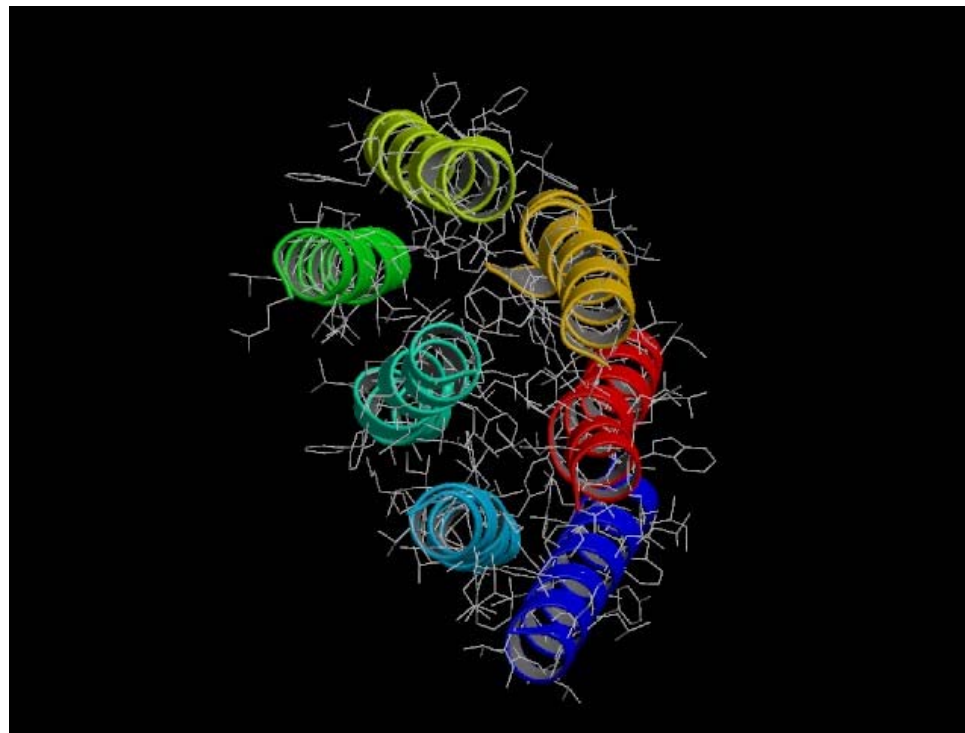


RMSD 0.59 Å ($\epsilon = 1.0$)

Bacteriorhodopsin (Case of 7 Helices)

H. Kokubo & Y.O., *Chem. Phys. Lett.* **392**, 168-175 (2004).

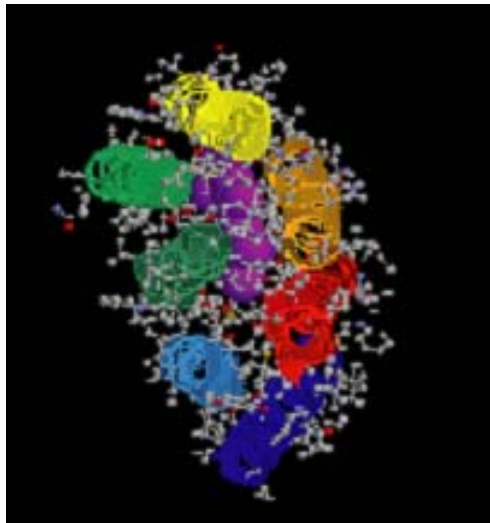
Native Structure



Bacteriorhodopsin

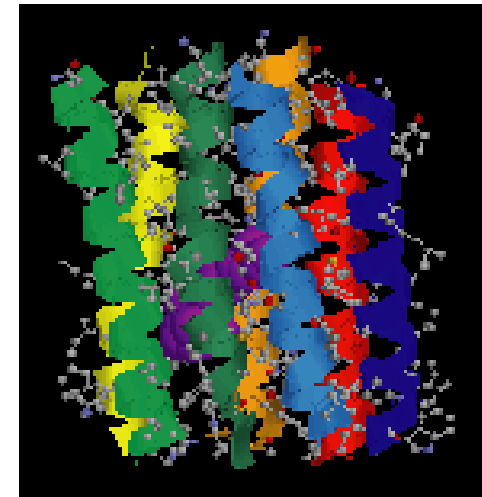
PDB structure

(top)



D **E**
C **F**
B **G**
A

(side)



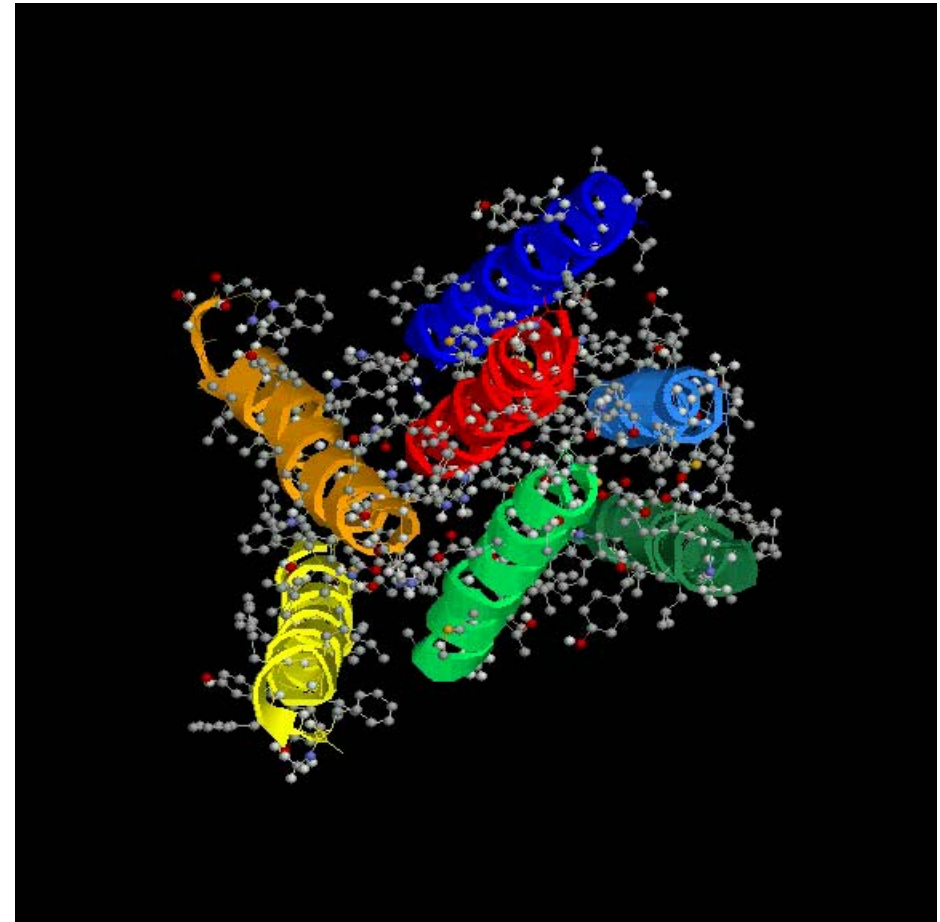
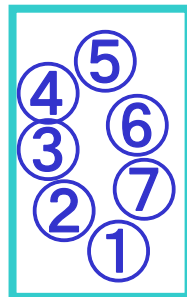
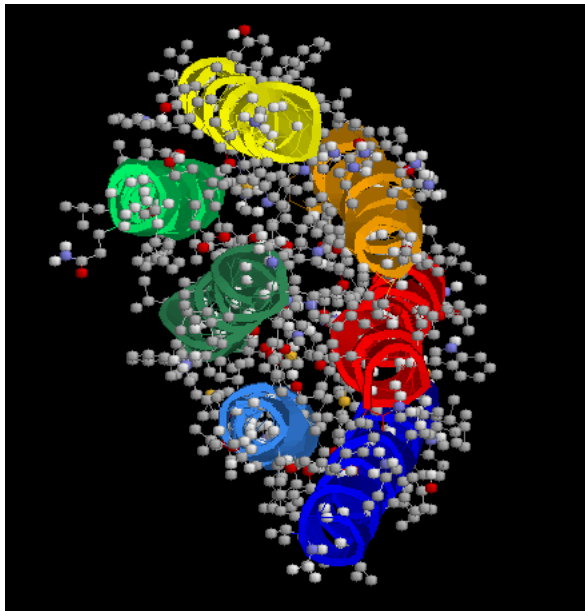
G F E
D C B A

Native structure (PDB code: 1C3W). Retinal is **purple**.
Seven transmembrane helices from **A** to **G**.
This figure was drawn without lipid molecules and loop regions.

REMC

(from above)

Replica 14

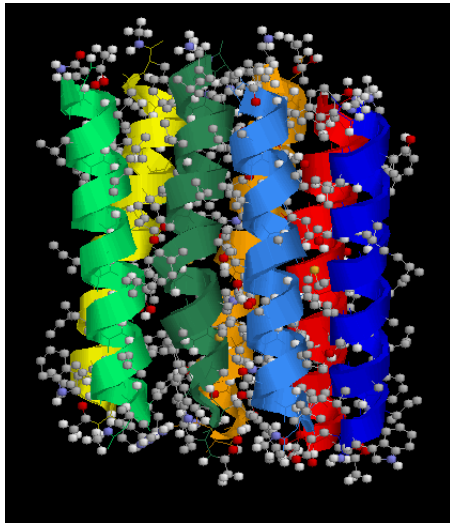


Native structure

REMC

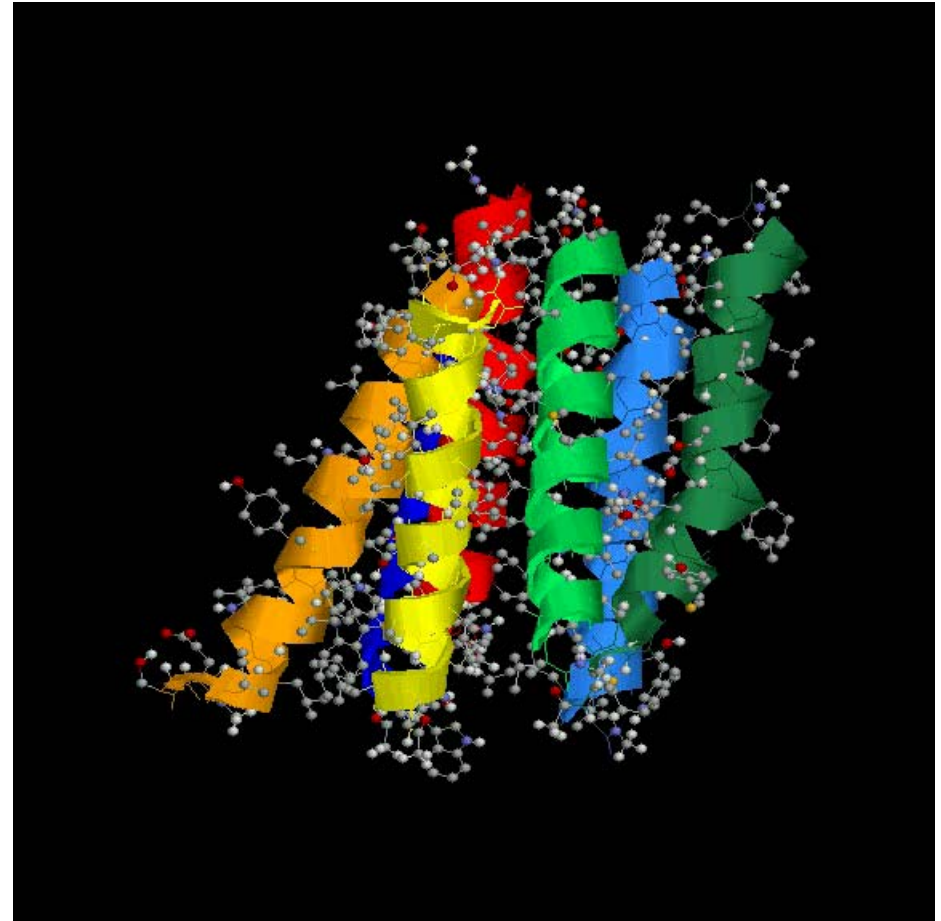
(from above)

Replica 14

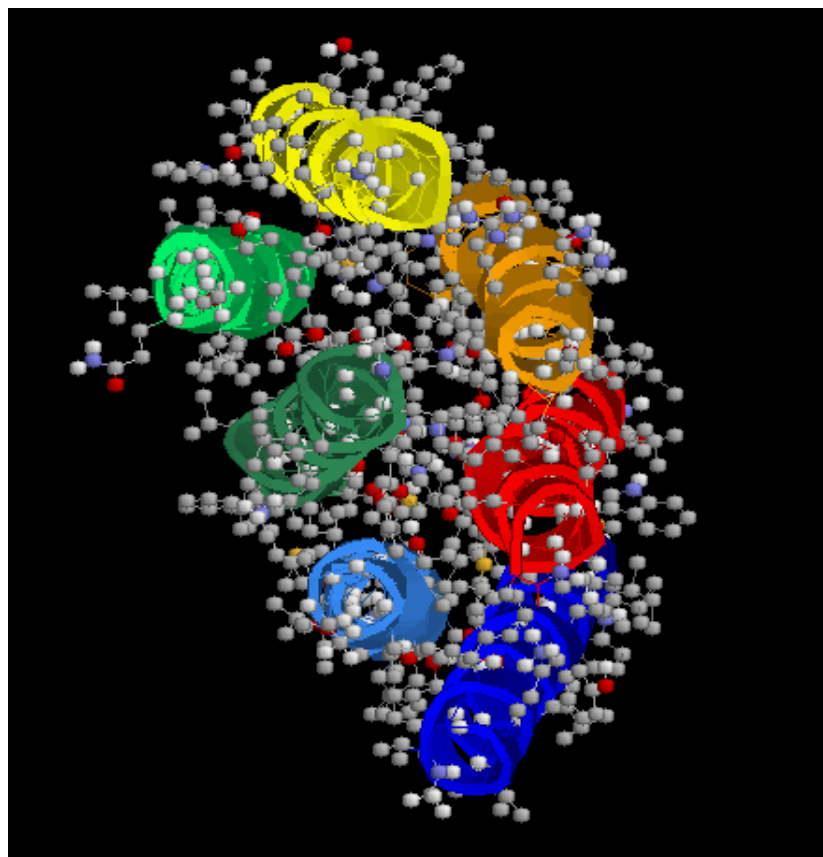


5 6 7
4 3 2 1

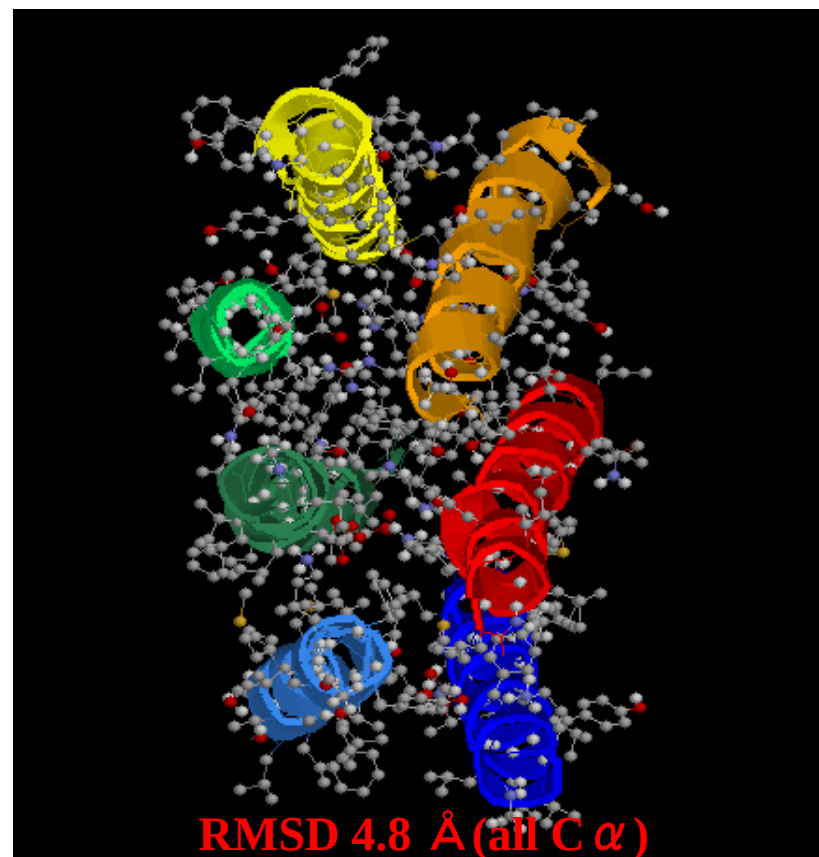
Native structure



Local Minimum Structure

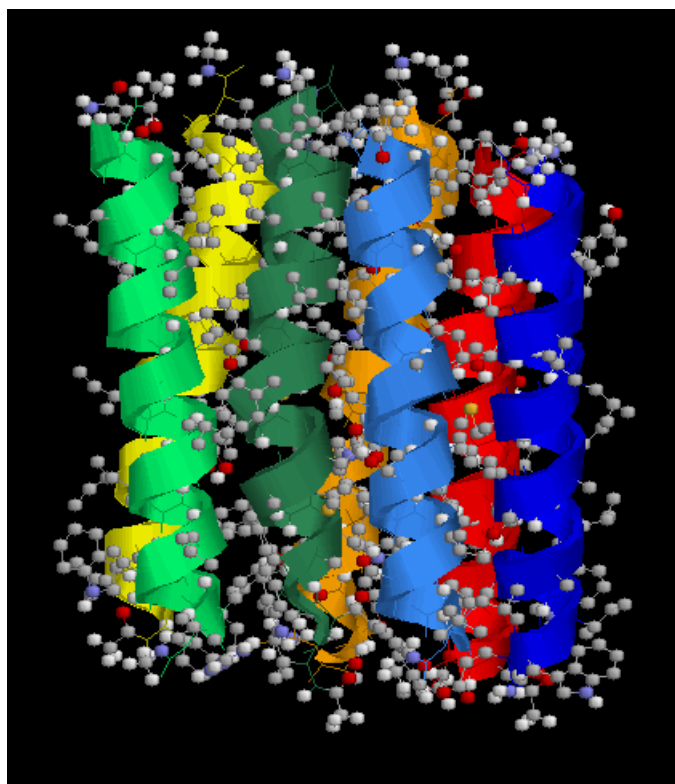


Native Structure

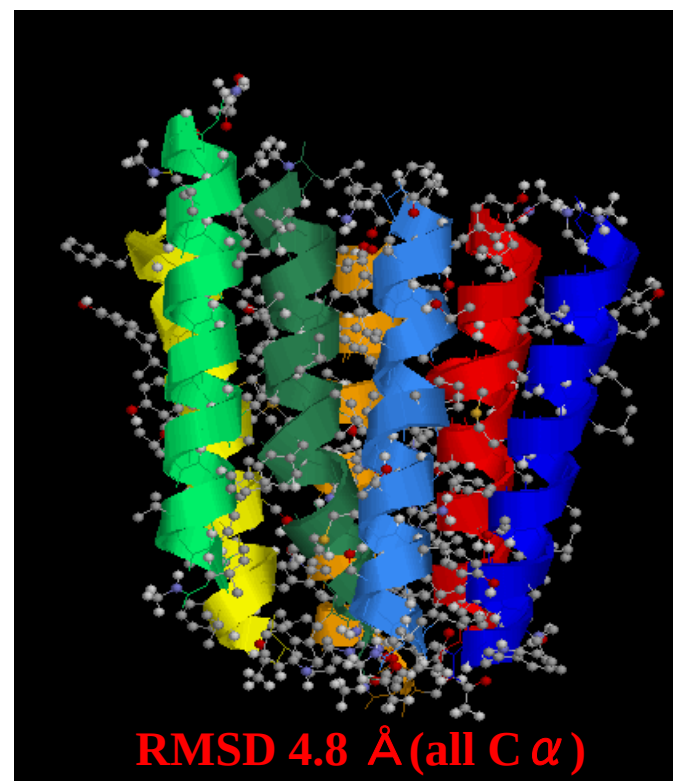


Replica 14

Local Minimum Structure



Native Structure



Replica 14

おわりに

バイオ・ライフサイエンス分野においては、
まだまだ計算機パワーが小さ過ぎるで扱え
る系がとても小さく制限されているが、少し
ずつ、実験に代わる予言能力が備わりつ
つある。