

計算生命科学の基礎Ⅳ

—計算科学と生命科学の融合：基礎から医療・創薬・人工知能への応用まで—

生命系の分子動力学シミュレーション

2017年11月29日

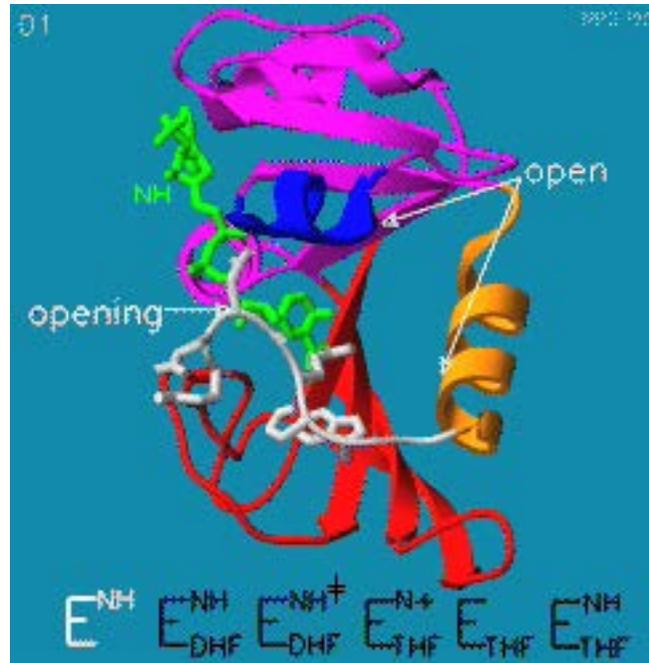
横浜市立大学 生命医科学研究科

池口 満徳

今日の内容

- イントロダクション: 生体分子の動きと機能
- 分子動力学シミュレーションとは
 - 基本アルゴリズム
 - 力場 (Force Field)、ソフトウェア
- 分子動力学シミュレーションの解析
 - 構造ゆらぎと構造変化
- 分子動力学シミュレーションの応用
 - 実験との連携
 - 超並列計算へ: レプリカ系の計算

タンパク質の機能と動き： タンパク質は硬い？

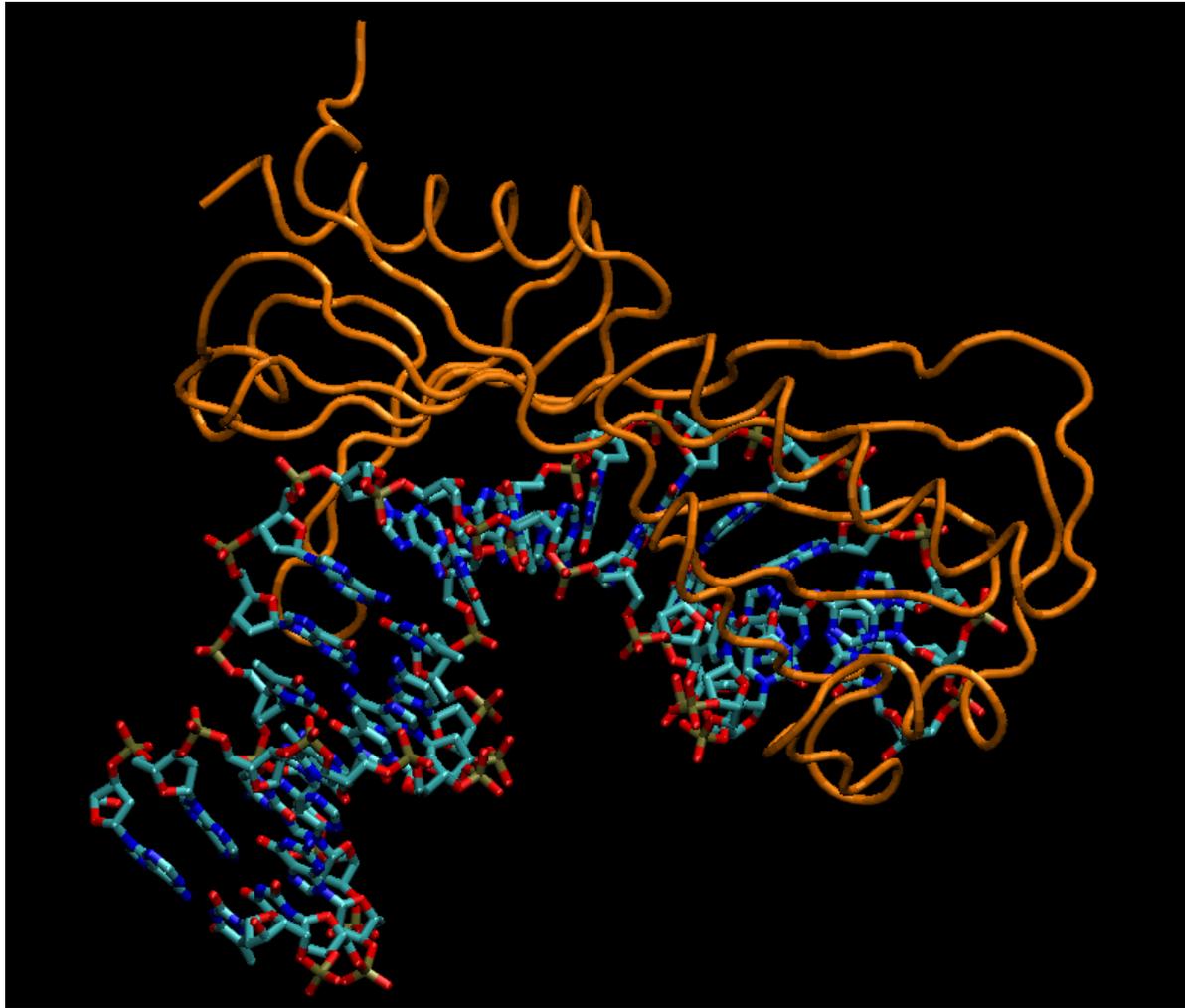


DHFR (dihydrofolate reductase): DHF → THFに還元する酵素

(<http://chem-faculty.ucsd.edu/kraut/dhfr.html>)

タンパク質の機能は、**動くこと**によって起こる

DNAは硬い？



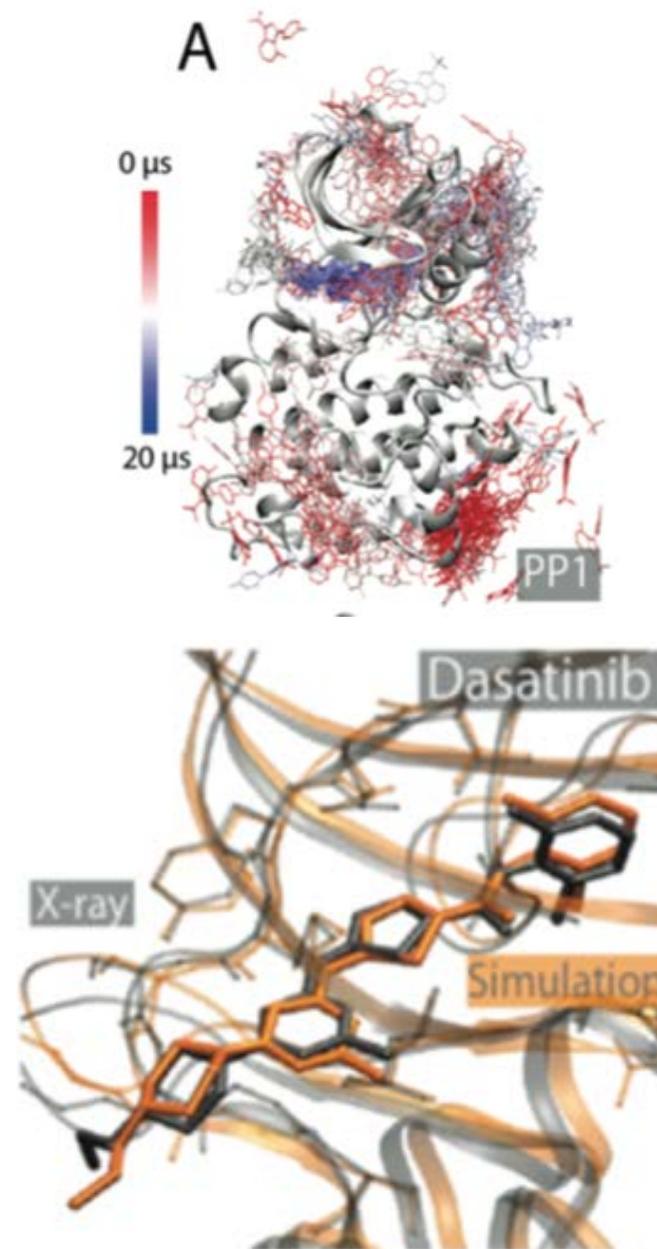
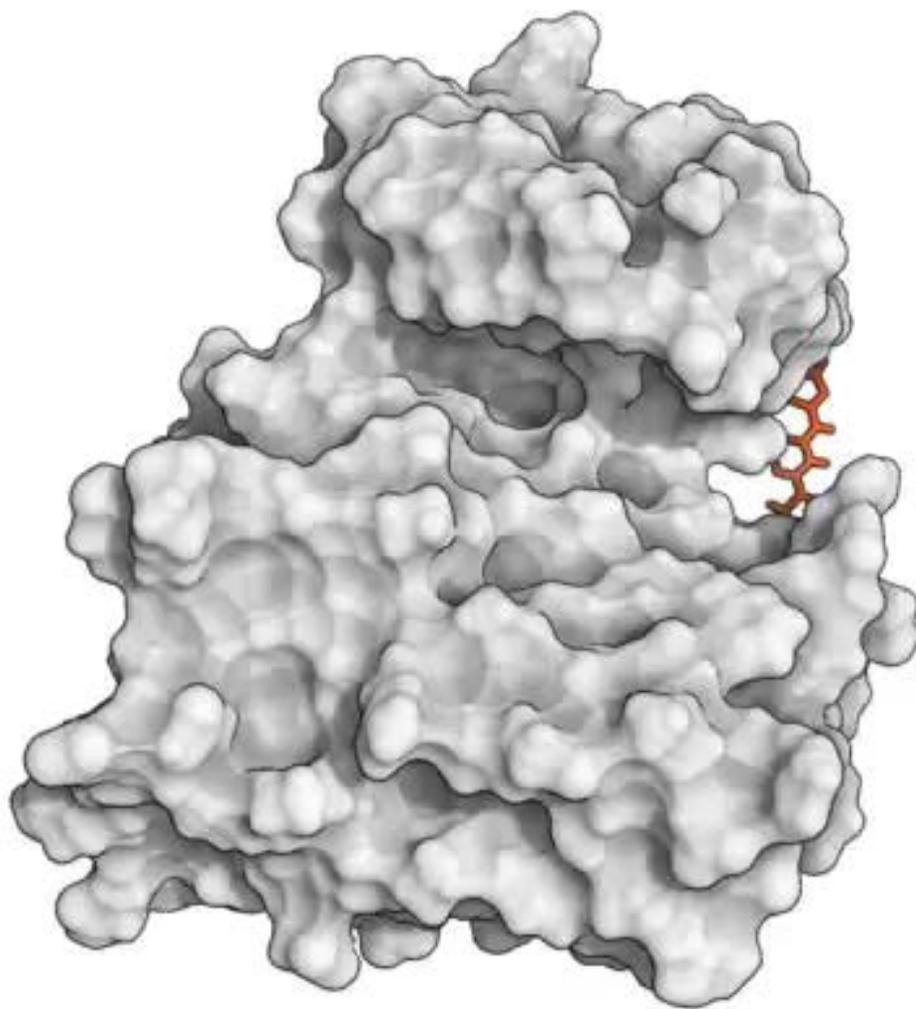
TATA binding protein
in TFIIB
DNA 複合体

DNAは曲がっている

PDB ID: 1AIS

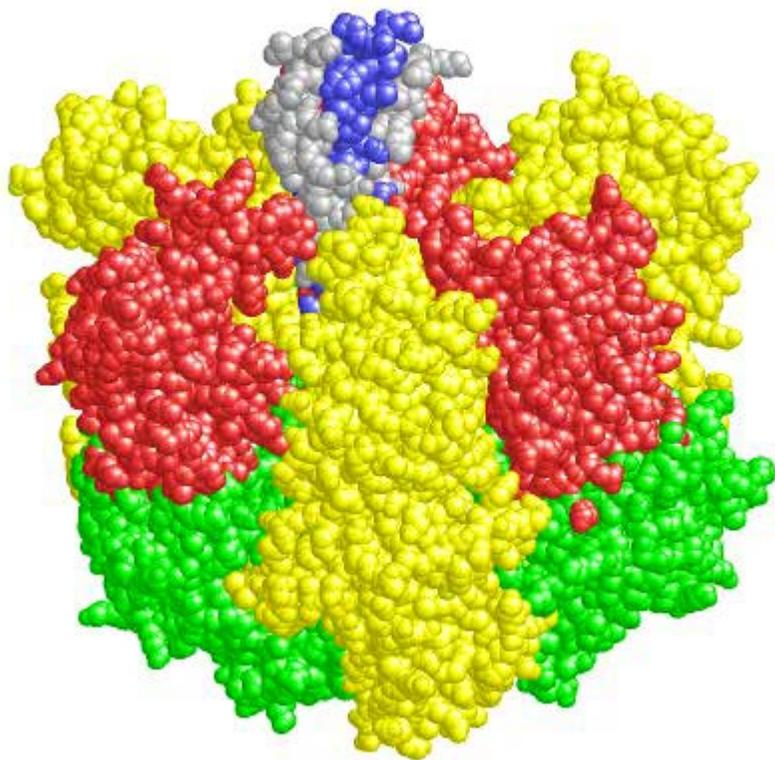
生体分子は、大きな**立体構造変化**をおこす

薬剤の結合過程の計算

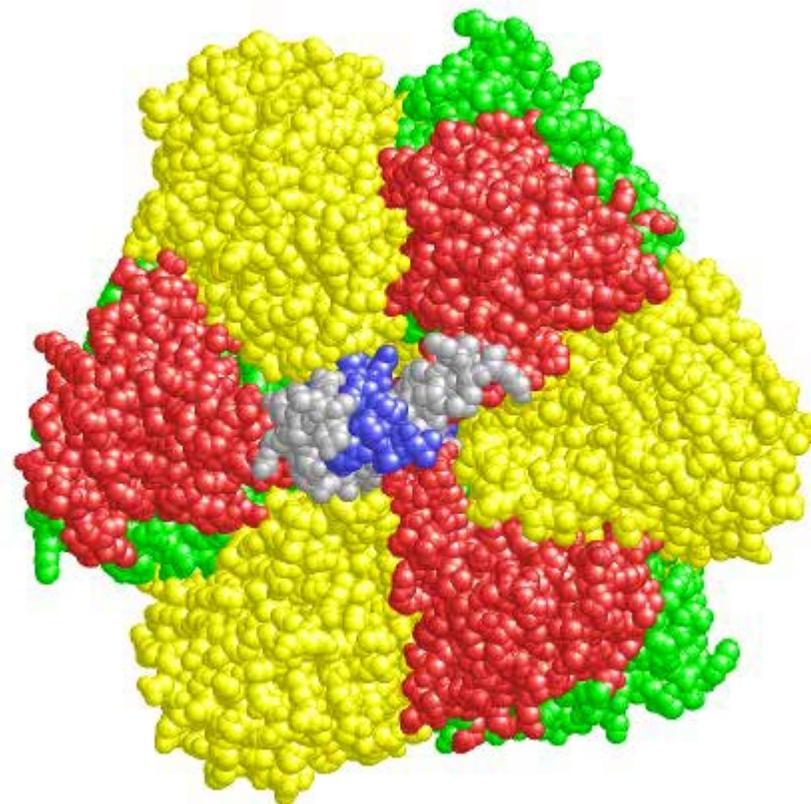


Shaw, J. Am. Chem. Soc. 133, 9181 (2011)

F₁-ATPase: 回転分子モーター



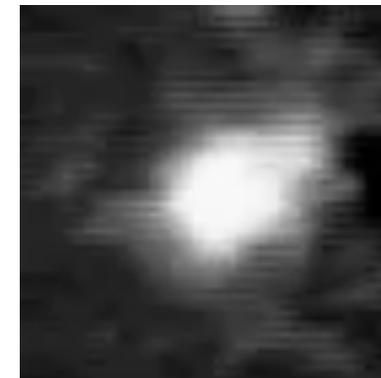
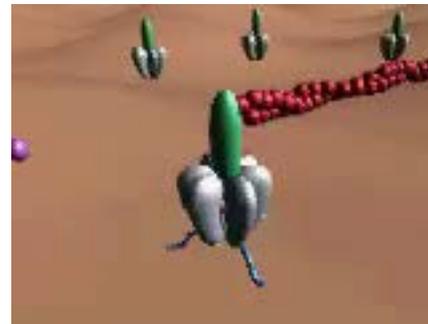
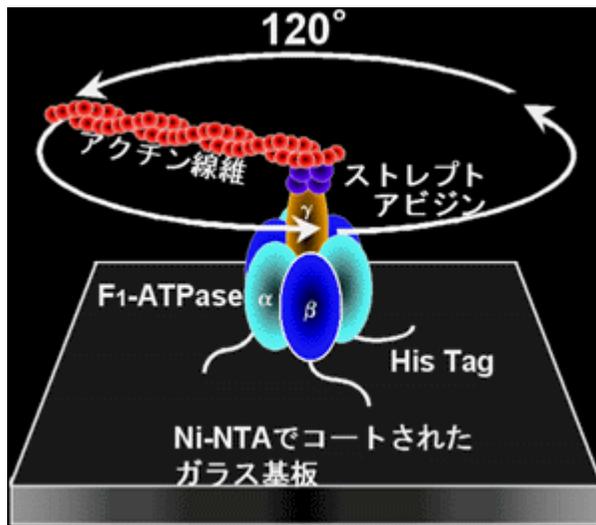
View from Side



View from Top

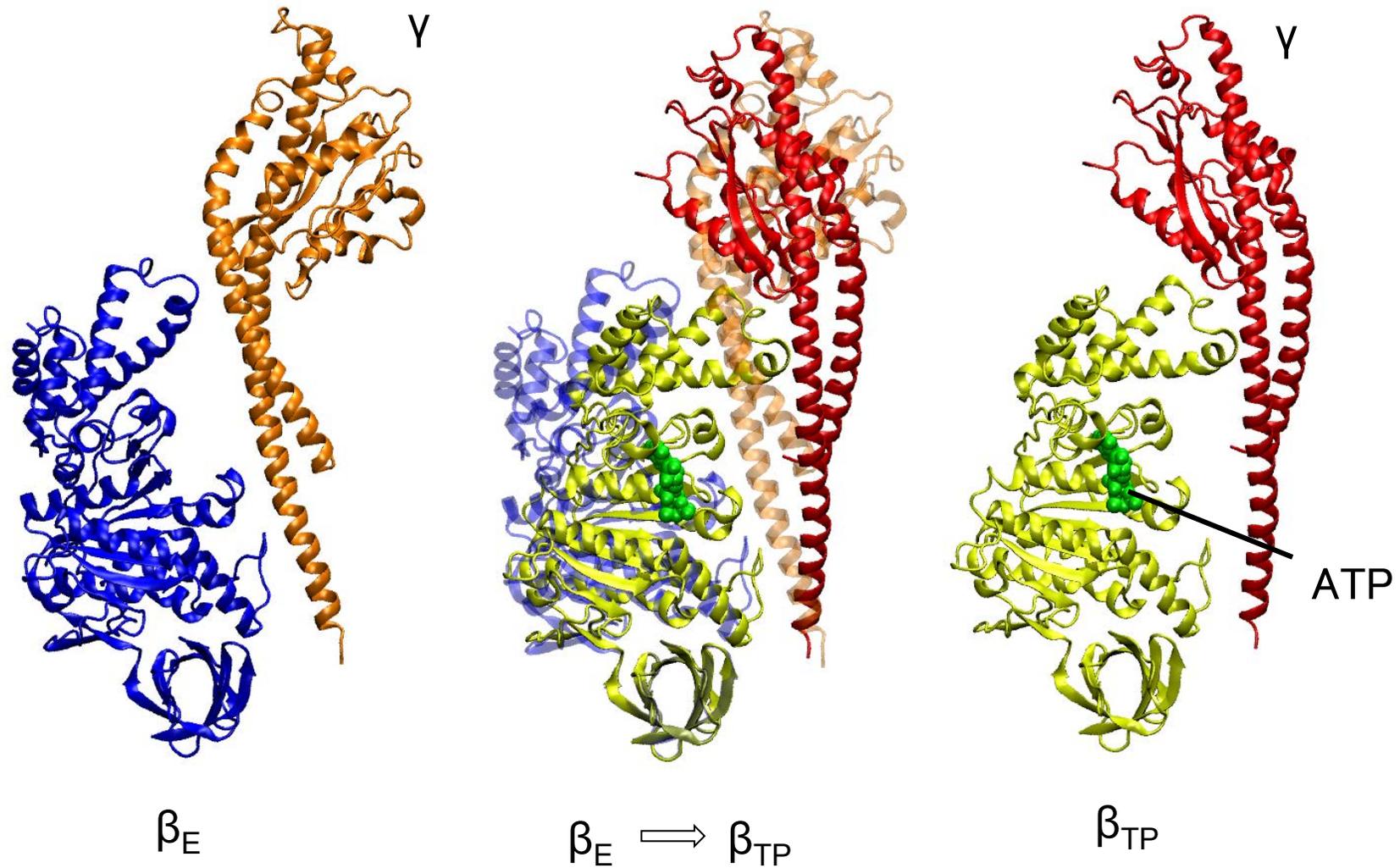
Web site of Oster's lab

Direct Observation of Rotation in F₁-ATPase

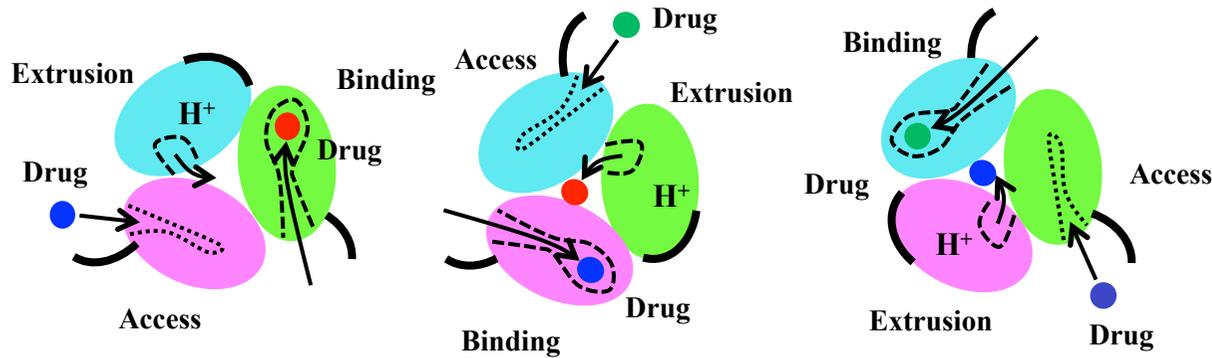


Yoshida et al. *Nature Reviews Molecular Cell Biology* 2 669–677 2001

Conformational Changes of β subunit upon ATP binding

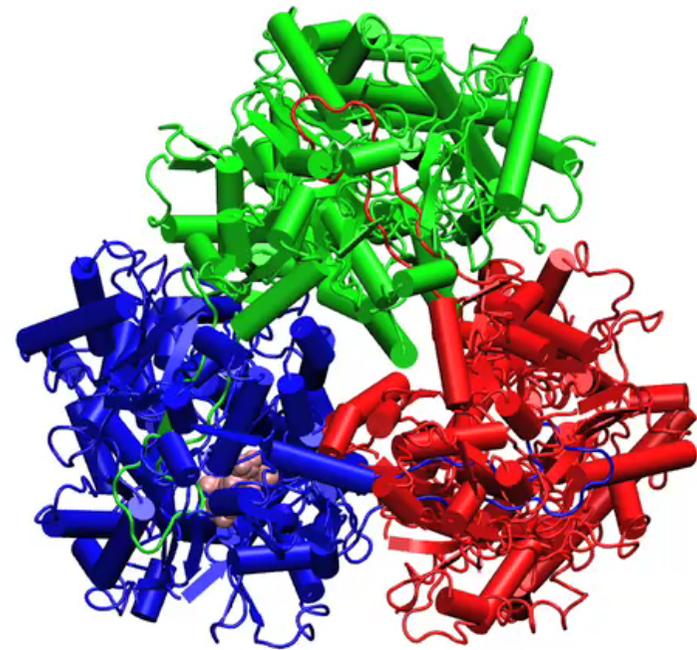
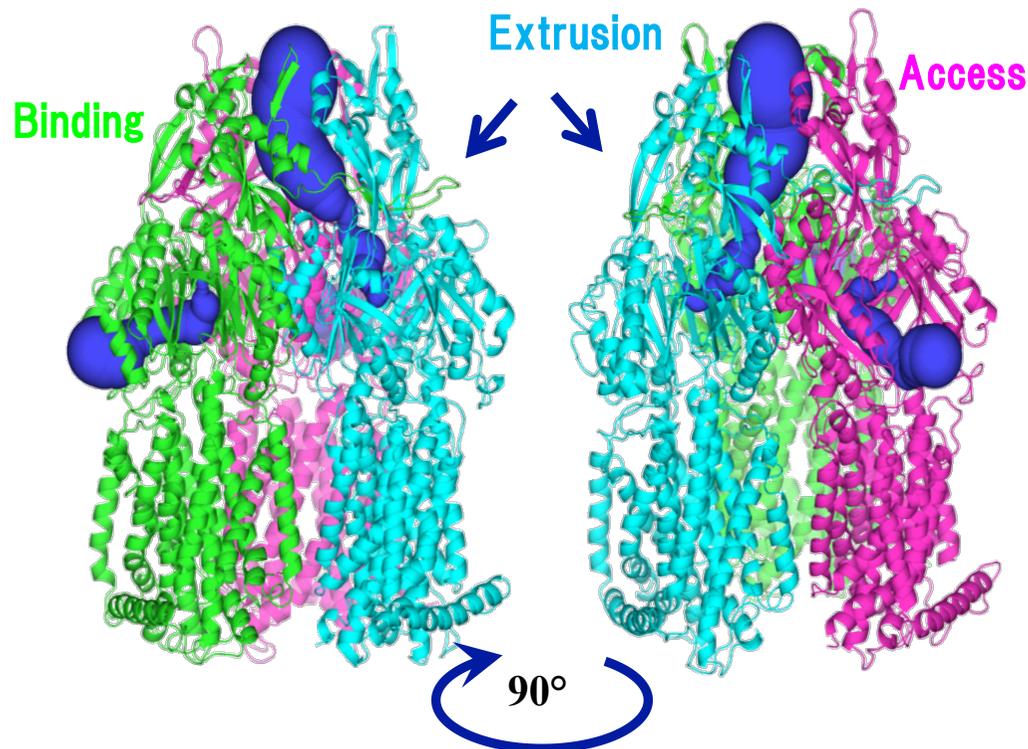


多剤排出トランスポーターAcrB

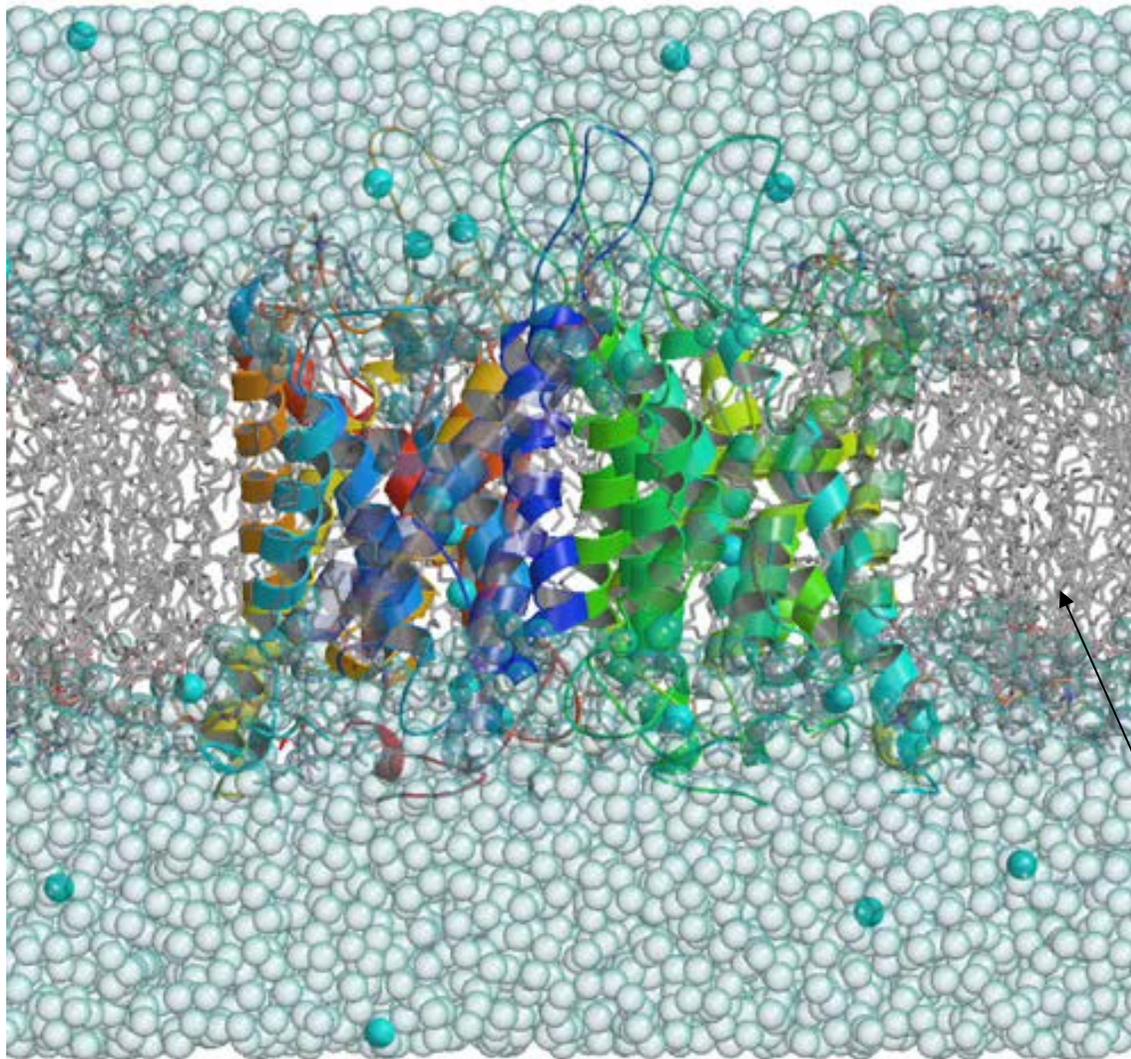


順番に状態が入れ替わることで、薬剤を取り込み、排出する

S. Murakami, *PNE*, 52 (2007), 406-414



膜タンパク質：水やイオンの動き（透過）



生体膜に埋もれている
タンパク質もある

膜蛋白質の機能：
膜内外の物質輸送

アクアポリン：
水を輸送するタンパク質

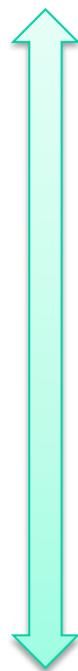
水やイオンの動きが重要

生体膜

by 橋戸公則

アクアポリンの水の透過の分子動力学シミュレーション

生体分子の機能： 何かが動く



	動くもの	機能	シミュレーション手法
小	電子	電子移動, 光励起	電子状態計算 (QM)
	プロトン	プロトン移動, プロトン化	QM, 全原子MD
	原子	構造変化など	全原子MD
	低分子	リガンド結合, 膜輸送	全原子MD
	側鎖	分子 (低分子, 高分子) 結合	全原子MD
	ドメイン	構造変化, 分子モーター	全原子, 疎視化MD
	サブユニット	超分子システム	全原子, 疎視化MD
	蛋白質全体	多蛋白質システム	疎視化MD
	分子濃度	細胞機能	システム生物学
	細胞	血流	流体力学, 構造力学
大	組織	臓器運動	流体力学, 構造力学

必要に応じて見るスケールを変えることが重要

分子動力学(MD)シミュレーション

MDの基本アルゴリズム

原子間力の計算

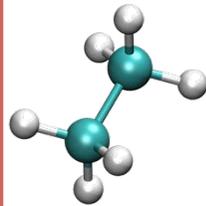


運動方程式による
原子位置の更新

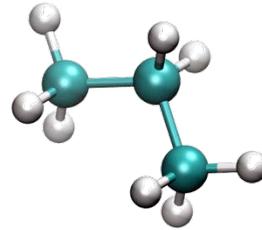


MDの力場ポテンシャル関数

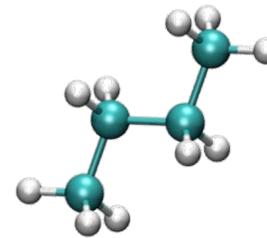
Bond



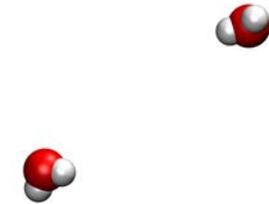
Angle



Torsion



Nonbond (Elec, vdW)



初のタンパク質MDを行ったKarplusらが
2013年ノーベル化学賞を受賞

The Nobel Prize in Chemistry 2013

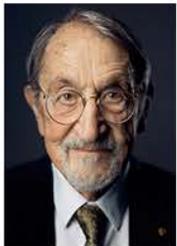


Photo: A. Mahmoud
Martin Karplus



Photo: A. Mahmoud
Michael Levitt



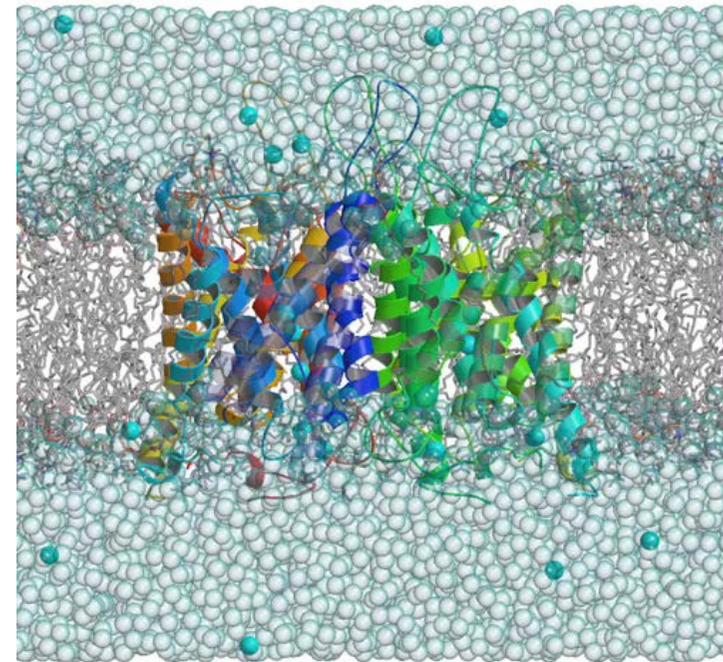
Photo: A. Mahmoud
Arieh Warshel

(<http://www.nobelprize.org/>)

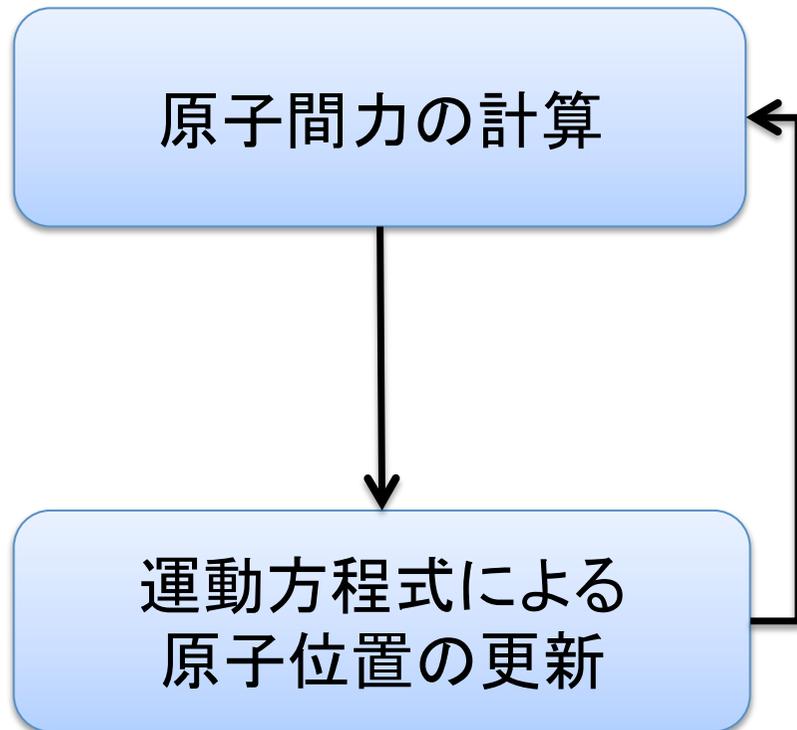
アクアポリンの 生体膜中でのMD

- ✓ 全原子数~16万
- ✓ 数千万~数十億
ステップの計算
- ✓ スパコンが必要

Hashido, Ikeguchi, Kidera,
FEBS. Lett., 579, 5549, 2005
Hashido, Kidera, Ikeguchi,
Biophys. J., 93, 373, 2007



分子動力学シミュレーションの基本アルゴリズム

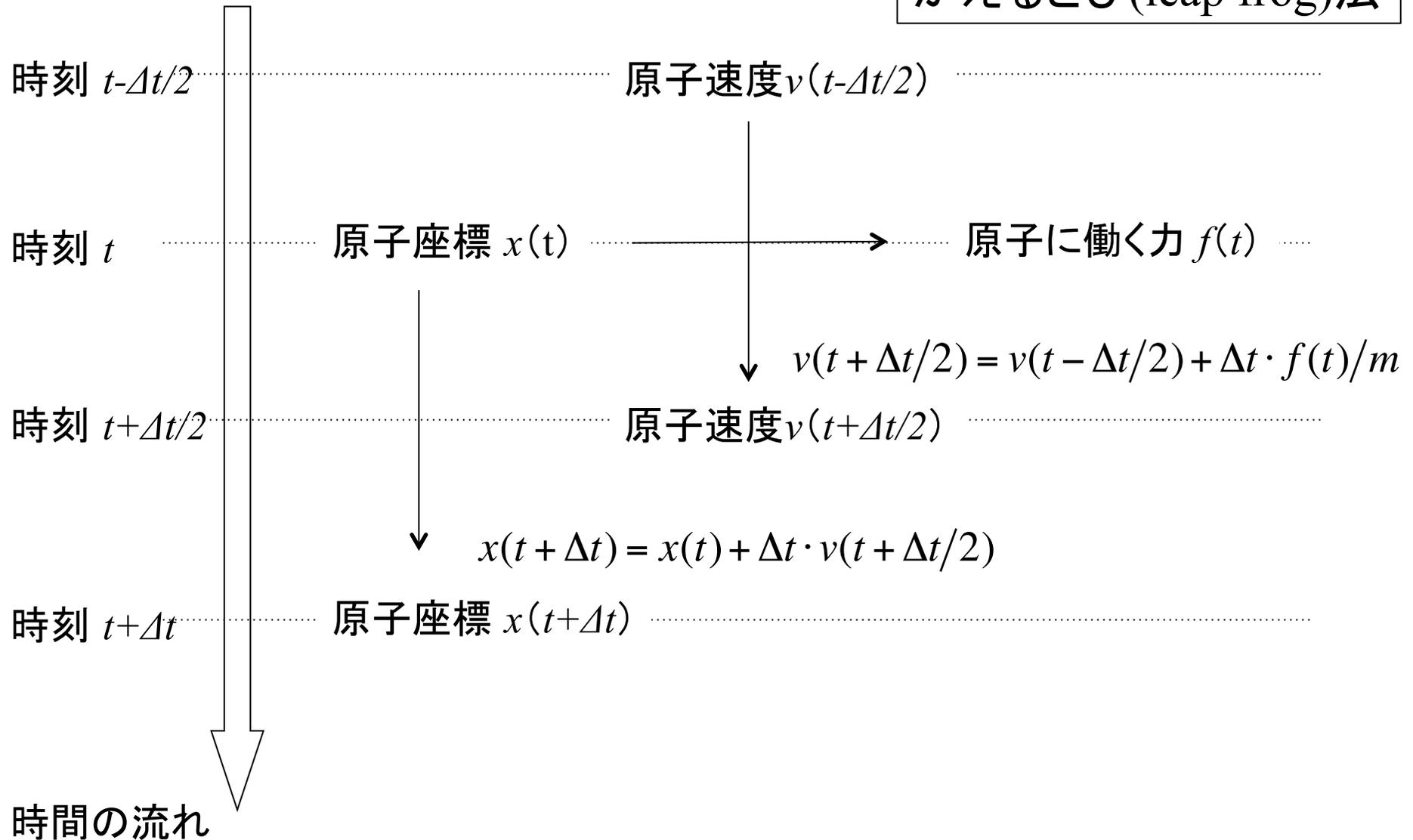


- ✓ 全原子MDでは、刻み幅は
1~2 fs (10^{-15} s) 程度
- ✓ 1ステップ1秒、計算にかかるとすると、
一日に、86,400ステップ=86 ps
しか計算できない！
- ✓ 1ステップ1ミリ秒で計算できると、
一日に、86,400,000ステップ=86 ns
- ✓ この速度でも1マイクロ秒のMD計算に
12日くらいかかる

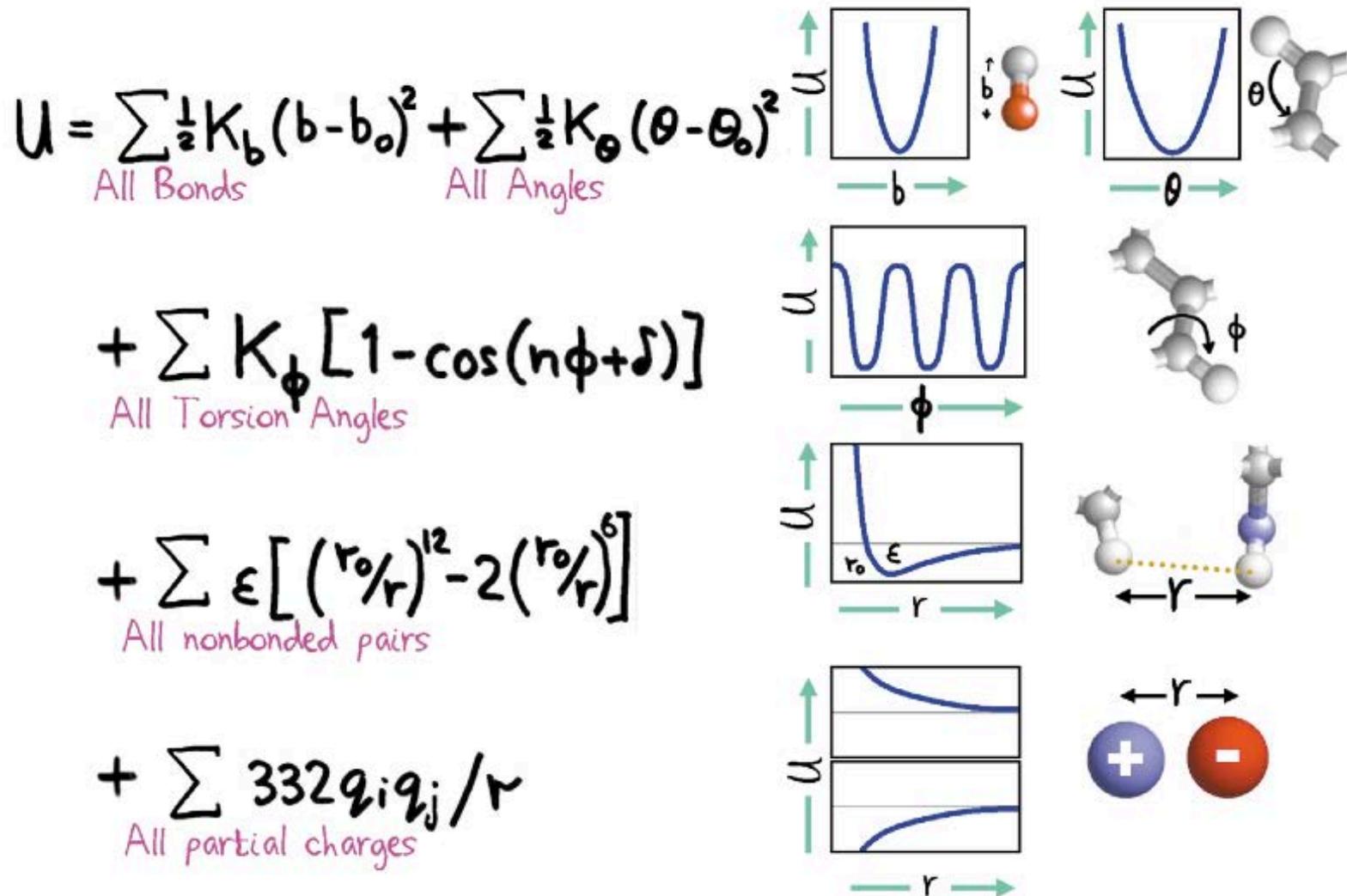
生体分子の動きを扱うには、高速な計算が必須！

分子動力学: 力による座標位置の更新(時間発展)

かえるとび(Leap frog)法



力場 (Force Field) とは



1969-1977くらいに、誕生。シーケンス解析も同じ頃に生まれる

M. Levitt, The birth of computational structural biology, Nature Struct. Biol. 8, 392, (2001)

力場 と MDソフトウェア

- 力場(Force Field)とは、ポテンシャル関数のパラメータ集
- それぞれの力場には、作成法のポリシーがある
- それぞれの力場は、それを開発したMDソフトウェアとセットであったが、最近はMDソフトウェアだけ開発し、力場は従来のものである例も多い
- 基本は、固定電荷の場合が多いが、誘導電荷を含める力場もある

相互作用(原子に働く力)の計算法

タンパク質、核酸用の力場 (force field)

AMBER (Kollman), CHARMM (Karplus), GROMOS (Berendsen), OPLS (Jorgensen), etc.

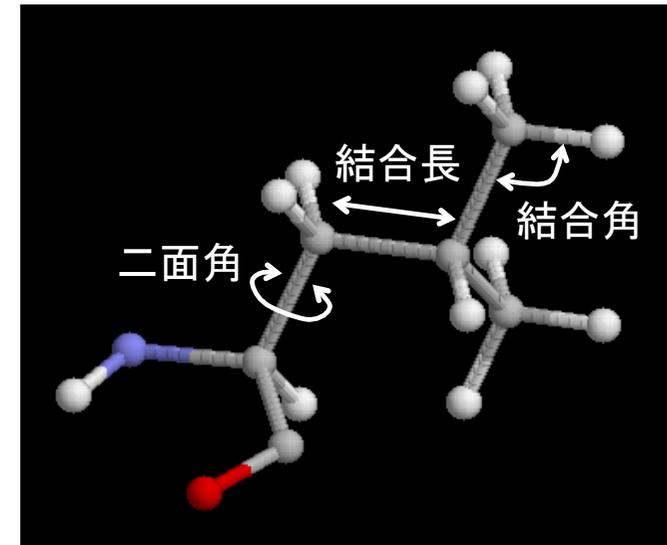
共有結合に関するエネルギー項

結合長, 結合角, 二面角(ねじれ角)

非共有結合に関するエネルギー項

静電相互作用 (水素結合)

ファンデルワールス相互作用



パラメータは、低分子の分子軌道法、実験データに基づいて決定

AMBER (Assisted Model Building with Energy Refinement)

<http://ambermd.org>

- Amber (有償) + AmberTools (無償)からなる
- 最新版: Amber16 AmberTools17
- GPU版MDコードが高速
- AMBER力場は、多くの派生版が存在
 - AMBER99SB-ILDNなどが有名
- 低分子化合物用力場: GAFF
- 力場作成のポリシー
 - 量子計算後、表面電場を再現するように電荷を配置 (RESP電荷)

CHARMM (Chemistry at HARvard Macromolecular Mechanics)

<https://www.charmm.org/>

- Karplusを中心に多くのグループにより、開発
- ソフトウェアには、多数の機能がある
- GPUなどの高速版(有償) 基本版(アカデミック無償)がある。最新版:42
- 力場:c36_Jul17が最新。Mackerell (U. Maryland)によって管理。リガンド力場はCGENFF。
- 力場のポリシー:
 - 相互作用をベース。水分子を水素結合のモデルとして、水分子との相互作用を量子計算を比較
 - CMAPというアミノ酸主鎖(ϕ 、 ψ)用相関ポテンシャルを搭載
 - 生体膜に強い

NAMD (Scalable Molecular Dynamics)

<http://www.ks.uiuc.edu/Research/namd/>

- イリノイ大学で開発
- 超並列計算が得意。Charm++ 並列プログラミングモデルを活用。C++で記述。無償。
- 巨大な系の計算もできる。
- GPU版もある。
- VMD(Visual Molecular Dynamics)というグラフィックスソフトも開発。
 - VMDには、MDのセットアップや解析のスクリプトなどが多数用意されている
- 力場は、専用のものではなく、CHARMM等を利用

GENESIS (GENeralized-Ensemble Simulation System)

<http://www.aics.riken.jp/labs/cbrt/>

- 理化学研究所の杉田グループで開発
- 「京」、ポスト「京」コンピュータに最適化
- GPU版もある。
- 無償。オープンソース。
- 超巨大な系の計算も可能。(細胞質計算など)
- レプリカ交換などのレプリカ計算も得意
 - String法などのレプリカパス法
- 力場は、従来のものを利用。
 - CHARMM, AMBER等

GROMACS (Groningen Machine for Chemical Simulations)

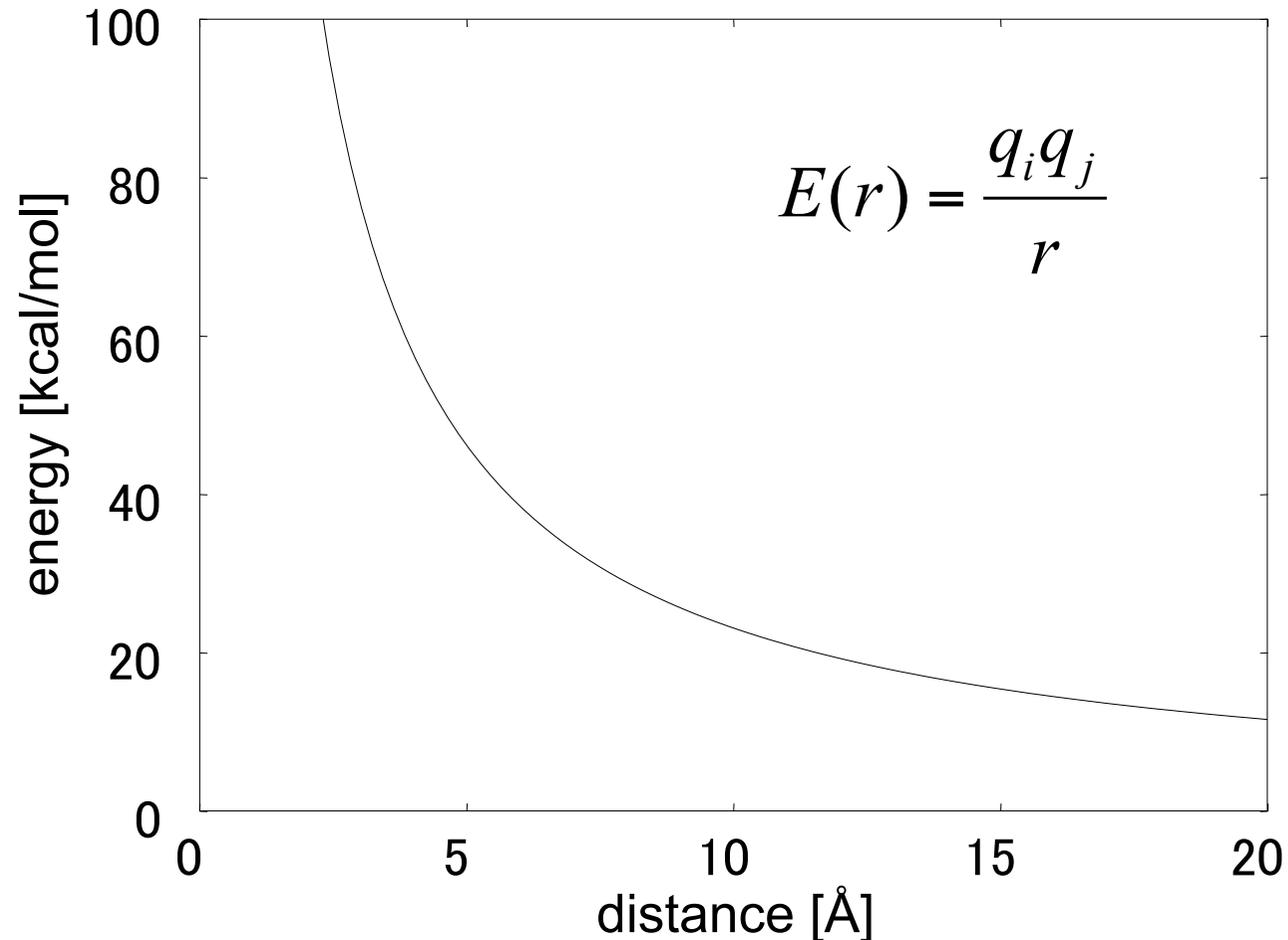
<http://www.gromacs.org>

- Lindahlらによって、ヨーロッパで開発。
- 計算機(CPU, GPU)に最適化したコードで、非常に高速
- CPU版とGPU版がある
- GPLでオープンソース。無償。
- MDコードだけでなく、多くのセッパアップや解析用機能も搭載している
- 多くの力場を利用可能。AMBERやCHARMM、OPLSも利用可能

OPLS (Optimized Potentials for Liquid Simulations)

- Jorgensenらによってモンテカルロ法により開発
- 液体の密度や気化熱の実験値に最適化
- 多くの力場に影響を与えた。
 - 水のTIP3PやTIP4Pは有名。
- OPLS-UA (united atom), OPLS-AA (all-atom)等
- 多くのMDパッケージでも利用可能 (GROMACS等)
- 最近は、Schrodinger社が開発継続。
OPLS3がリリース(商用)
 - DE Shaw ResearchのMDコード: DESMONDで利用可能

力場計算の注意点： 静電相互作用



水の酸素間のクーロンポテンシャル

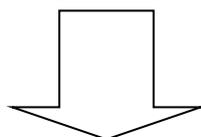
クーロンポテンシャルは、長距離まで減衰しないので途中で省略できない。

力場計算の注意点：静電相互作用

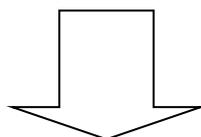
- 静電相互作用(クーロンポテンシャル)は、長距離まで減衰しないので、途中で省略できない
- 特殊な計算アルゴリズムを用いる必要あり
- Ewald法
 - 近距離と遠距離の相互作用に分割
 - 近距離：直接計算、遠距離：フーリエ変換により計算
 - FFTを用いるParticle Mesh Ewald (PME)が有名
- 多重極子展開法(FMM)
 - 近距離：直接計算、遠距離：多重極子展開により計算
 - 高速多重極子法(Fast Multipole Method;FMM)が有名
 - 名大で開発されているMODYLASに搭載

分子シミュレーション研究の流れ

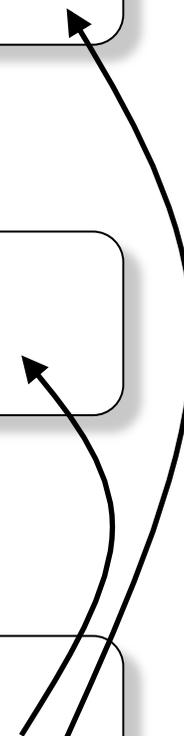
システムのセットアップ
(欠損部位のモデリング, 水素の追加, 水・イオン・膜の配置)



分子シミュレーションの実行
(スーパーコンピュータ/並列コンピュータの利用)

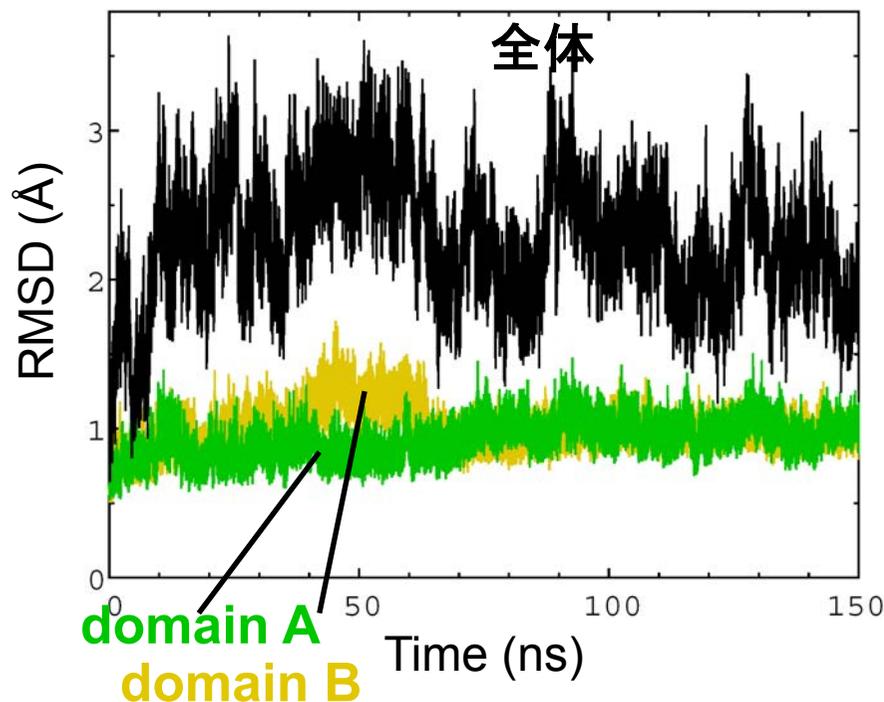


計算結果の解析
(目的に応じ, 様々な解析法がある)



MDシミュレーションの解析法: RMSD

(RMSD from crystal structure)

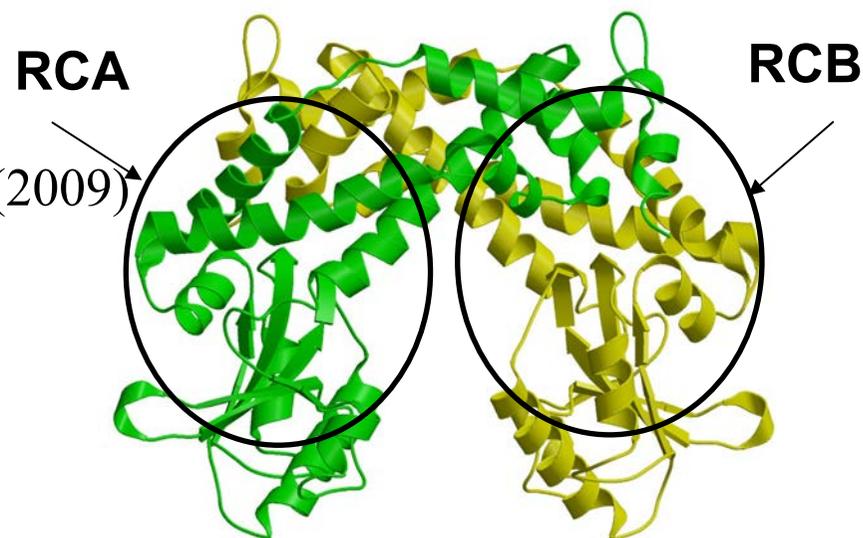


Root Mean Square Deviation (RMSD)

$$\text{RMSD} = \sqrt{\frac{(\sum_i \mathbf{x}_i - \mathbf{x}_{\text{ref}i})^2}{N}}$$

- ✓ 参照構造からのずれを観測
- ✓ MDの収束具合を見るのにも利用
- ✓ ドメイン運動している場合:
全体のRMSDは比較的大きくなる
各ドメインRMSDは小さい

Oroguchi, MI et al., Biophys. J., 96, 2808 (2009)



MDシミュレーションの解析法: RMSF, 相関行列

Root Mean Square Fluctuation (RMSF)

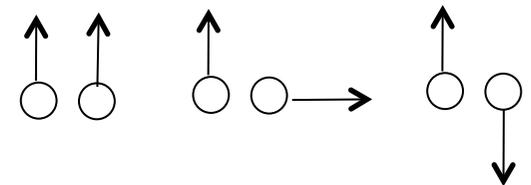
- ✓ 平均構造からのずれ(標準偏差)
- ✓ 揺らいでいる箇所、固い箇所を判別
- ✓ 2次構造部位は揺らぎが小さい
- ✓ ループ部位は揺らぎが大きい

$$R_i = \sqrt{\langle \Delta \mathbf{r}_i^2 \rangle}$$

$$C_{ij} = \frac{\langle \Delta \mathbf{r}_i \cdot \Delta \mathbf{r}_j \rangle}{\sqrt{\langle \Delta \mathbf{r}_i^2 \rangle \langle \Delta \mathbf{r}_j^2 \rangle}}$$

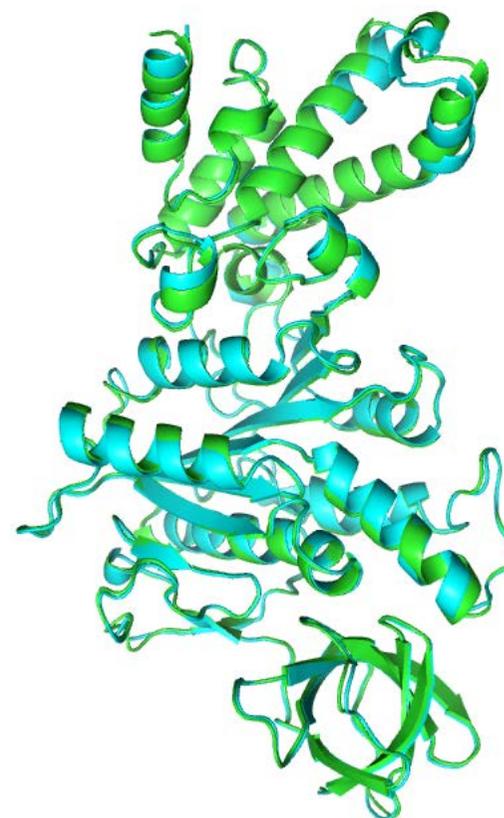
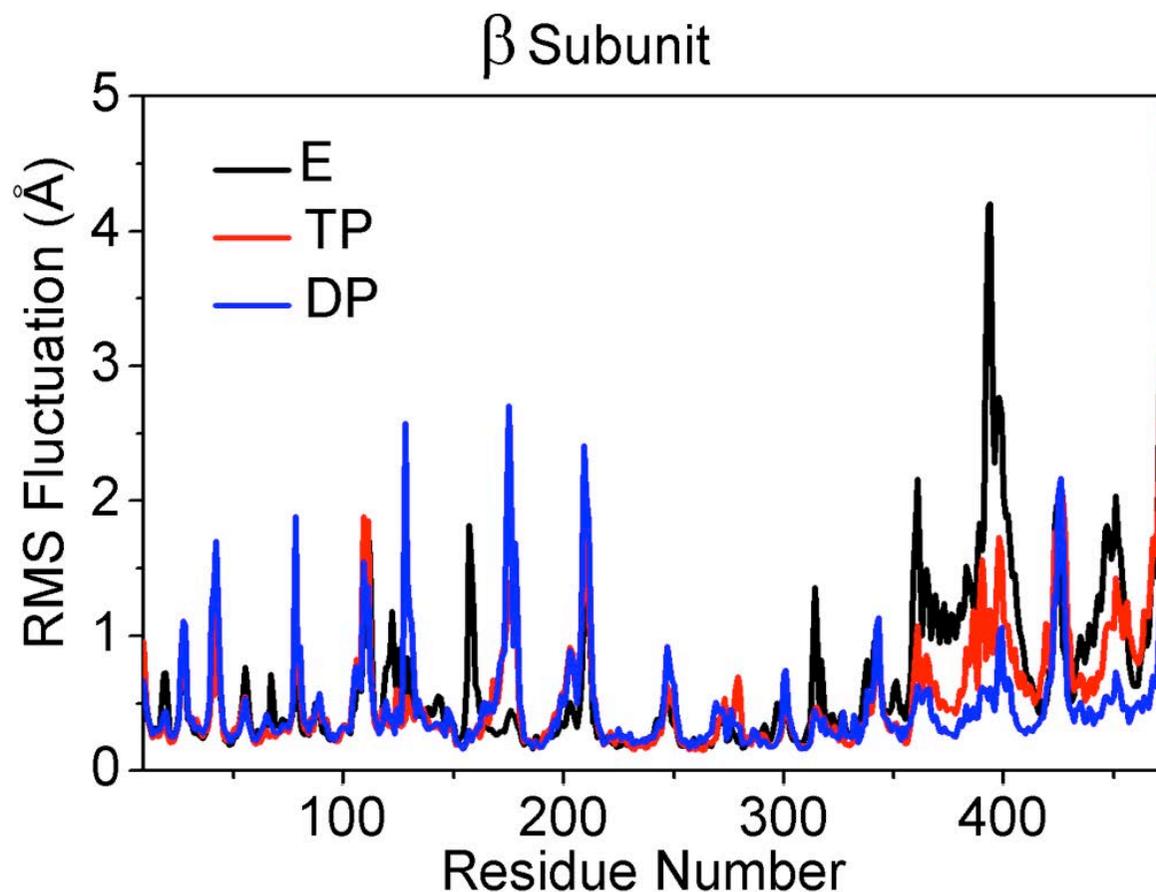
Correlation Matrix (相関行列)

- ✓ 運動相関の度合
- ✓ 一緒に動いている箇所を判別
- ✓ 動的ドメインがわかる



$$C_{ij} = 1 \quad C_{ij} = 0 \quad C_{ij} = -1$$

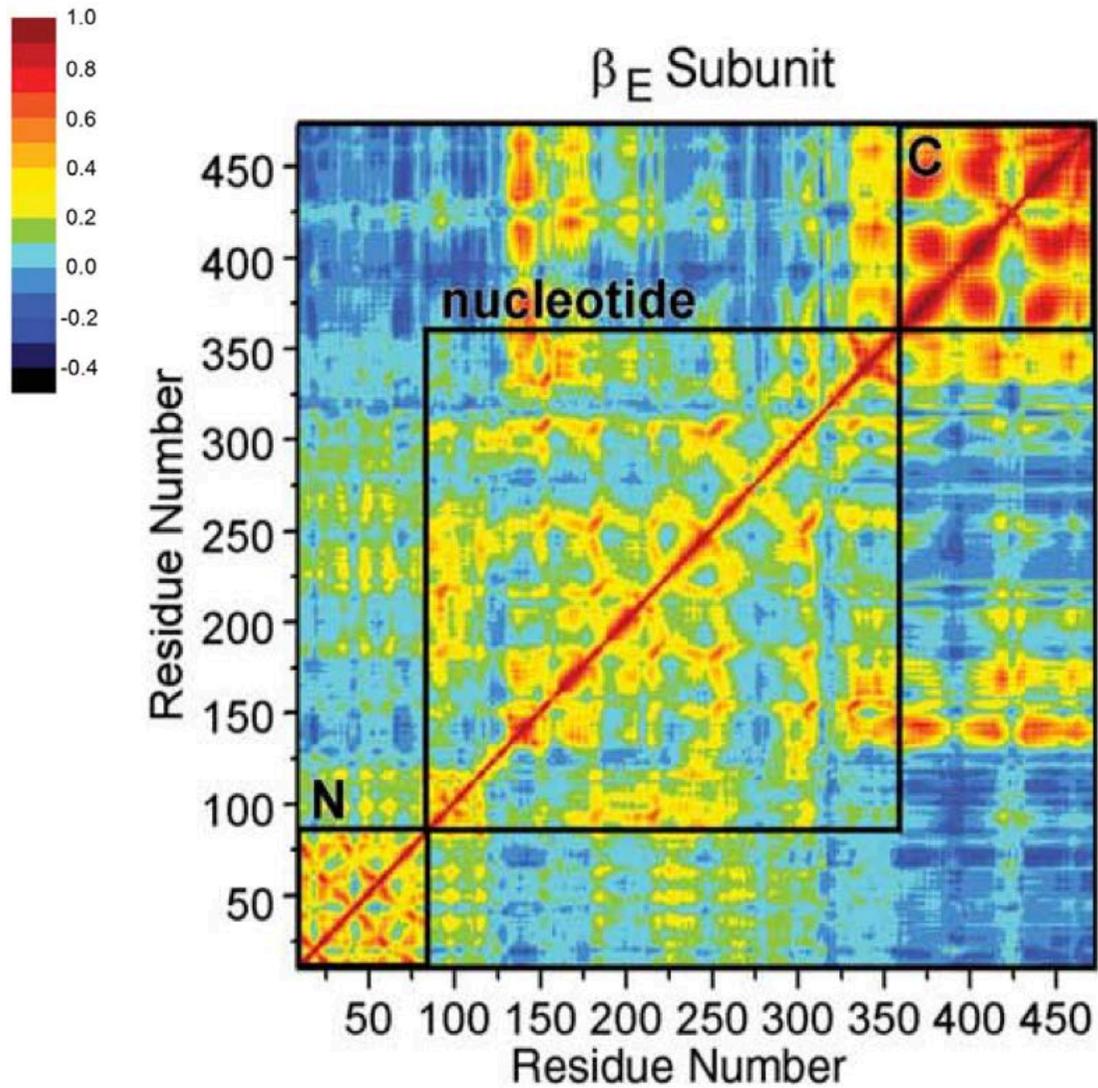
RMSF: β_E , β_{TP} , β_{DP} in complex



- ✓ Fluctuation of C-terminal domain: $\beta_E > \beta_{TP} > \beta_{DP}$
- ✓ Despite fairly similar conformations of β_{TP} and β_{DP} , fluctuations of C-terminal domain are different.

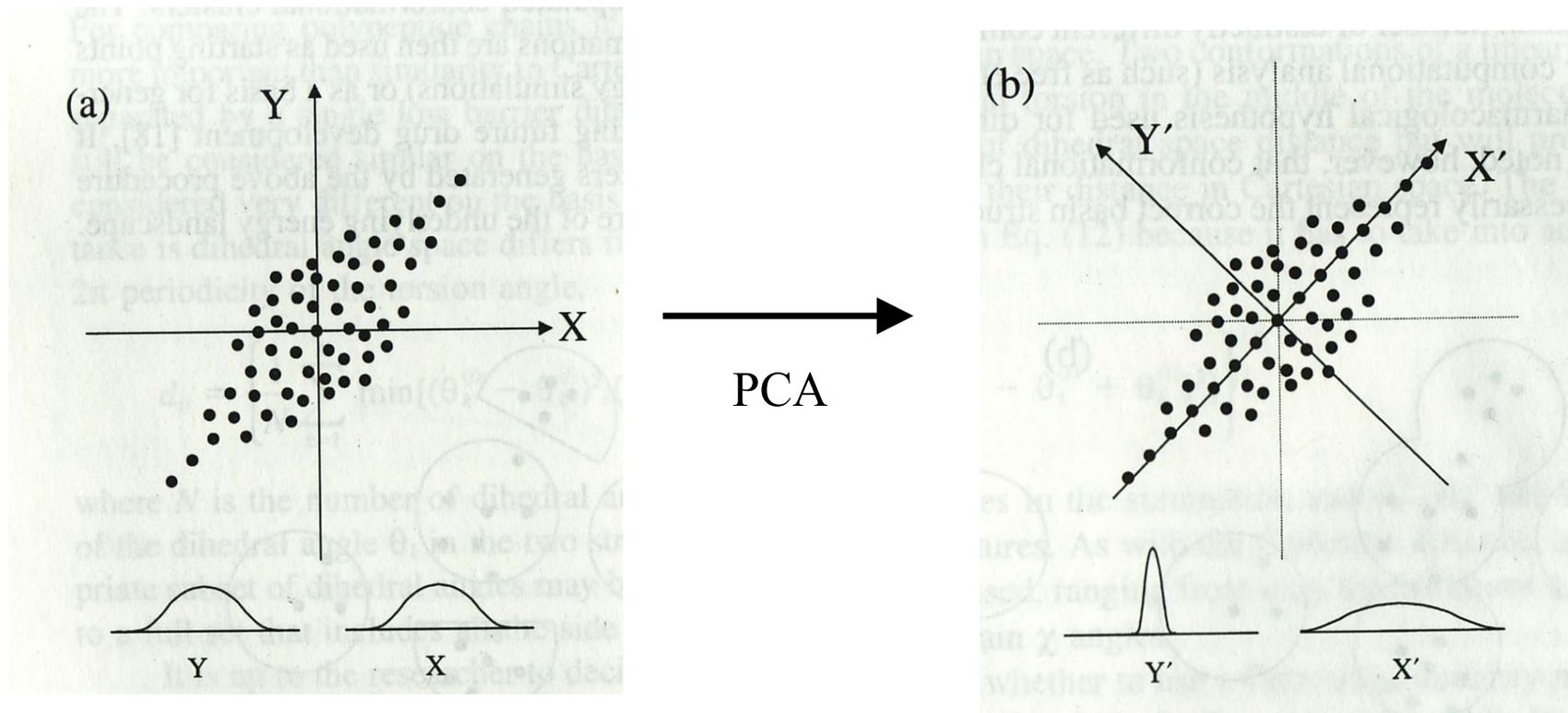
β_{TP} : cyan
 β_{DP} : green
 C_α -RMSD: 0.7 Å

Correlation Matrix



MDシミュレーションの解析法: PCA

主成分解析 Principal Component Analysis (PCA)

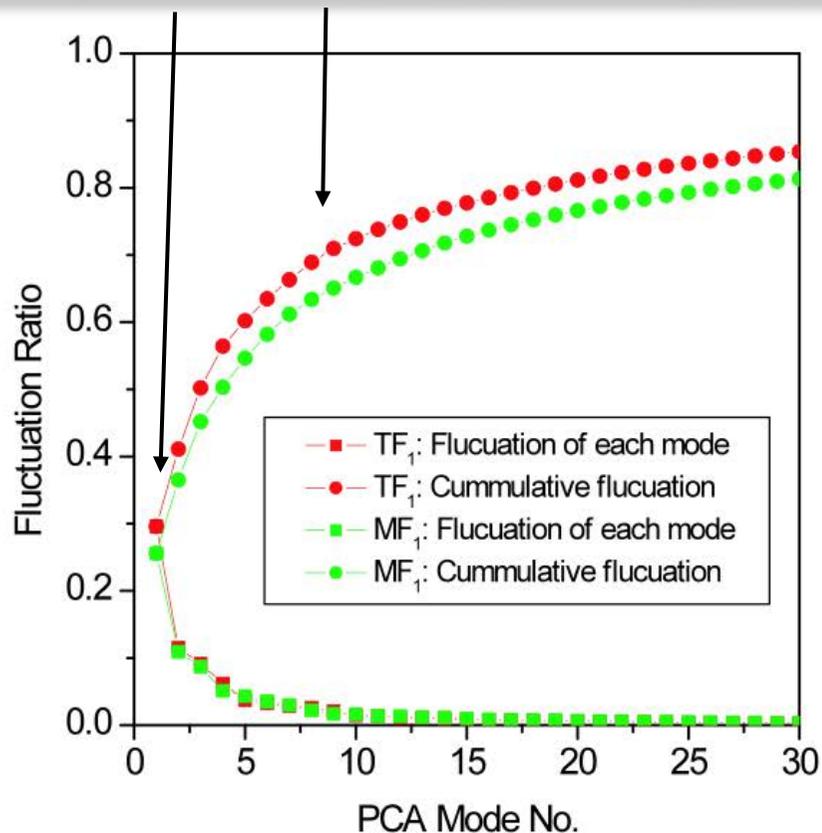


- ✓ PCAでは、大きな揺らぎ方向を抽出
- ✓ 原子の集団運動を得ることができる。

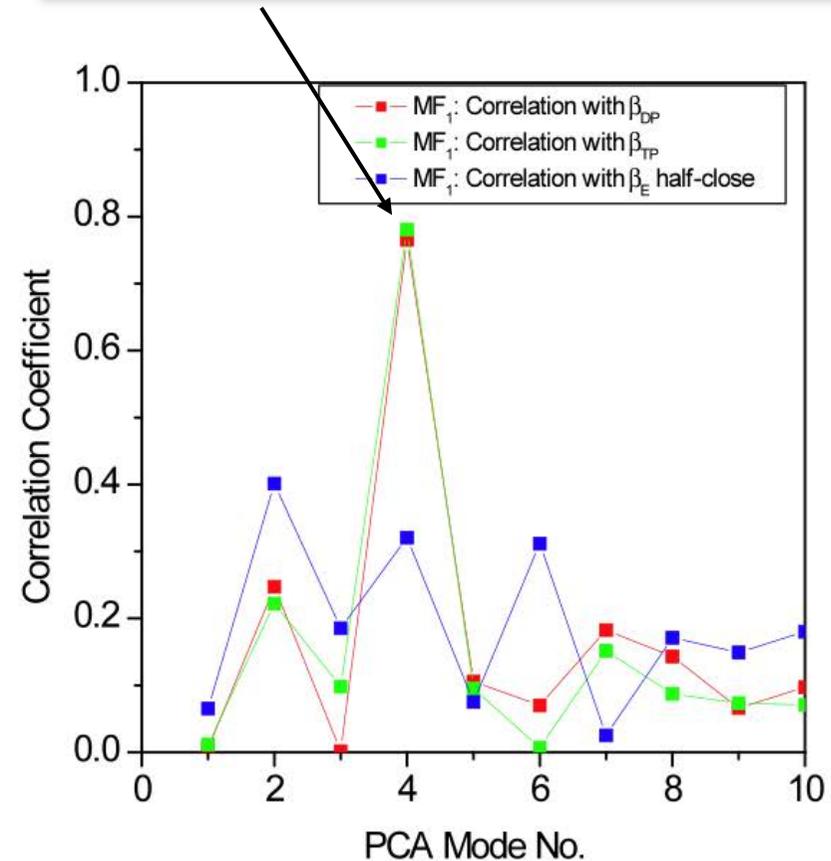
Computational Biochemistry and Biophysics, Becker et al. eds., 2001.

PCA: F_1 -ATPase β サブユニット(単体)

- ✓ Top mode は、全体揺らぎの 30 % を占める
- ✓ Top ten modes は、全体揺らぎの 70 % を占める



- ✓ 4番目のモードは ATP結合時の構造変化と良い相関を示す

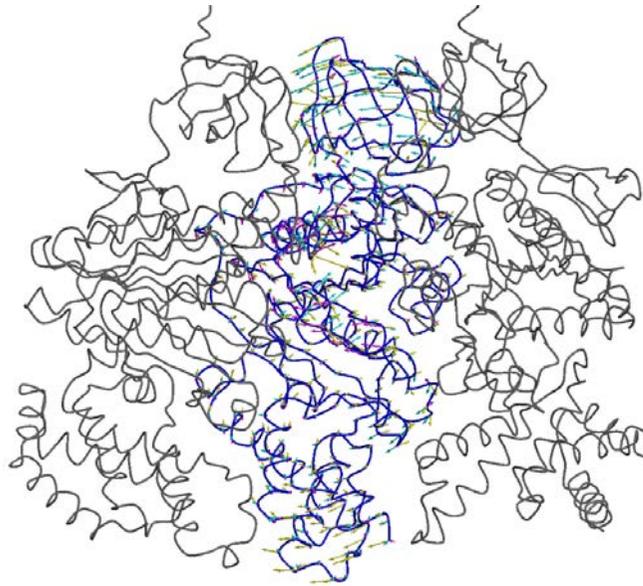


蛋白質は、非常に大きな自由度を持つが、揺らぎで見ると、異方性が強く、限られた方向にのみ、揺らいでいる傾向がある。

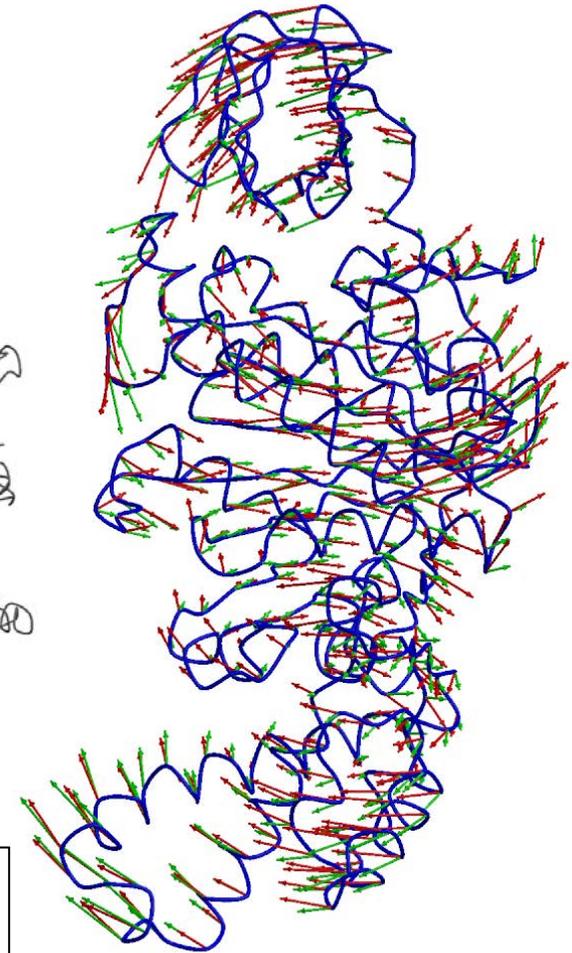
Principal Component Analysis: Mode 1-4

✓ Relaxation from interactions with adjacent subunits

Mode 1: magenta
Mode 2: yellow
Mode 3: cyan



Mode 4: green
 $\beta_E \rightarrow \beta_{DP}$: red



✓ Intrinsic flexibility is well correlated with structural transition

MDとNMRの連携法

方法1(拘束MD)

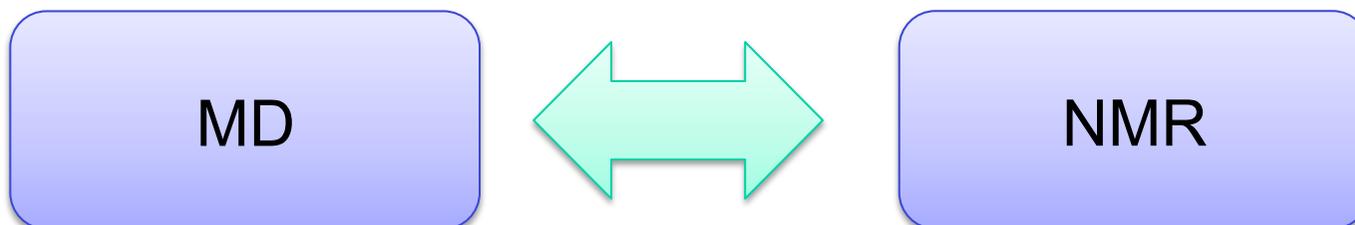
拘束条件(NOE, RDC等)を
入れて計算



- 構造決定のためのツール
- 利点: 実験情報を構造上に再現
- 欠点: 純粋な物理化学的シミュレーションではない

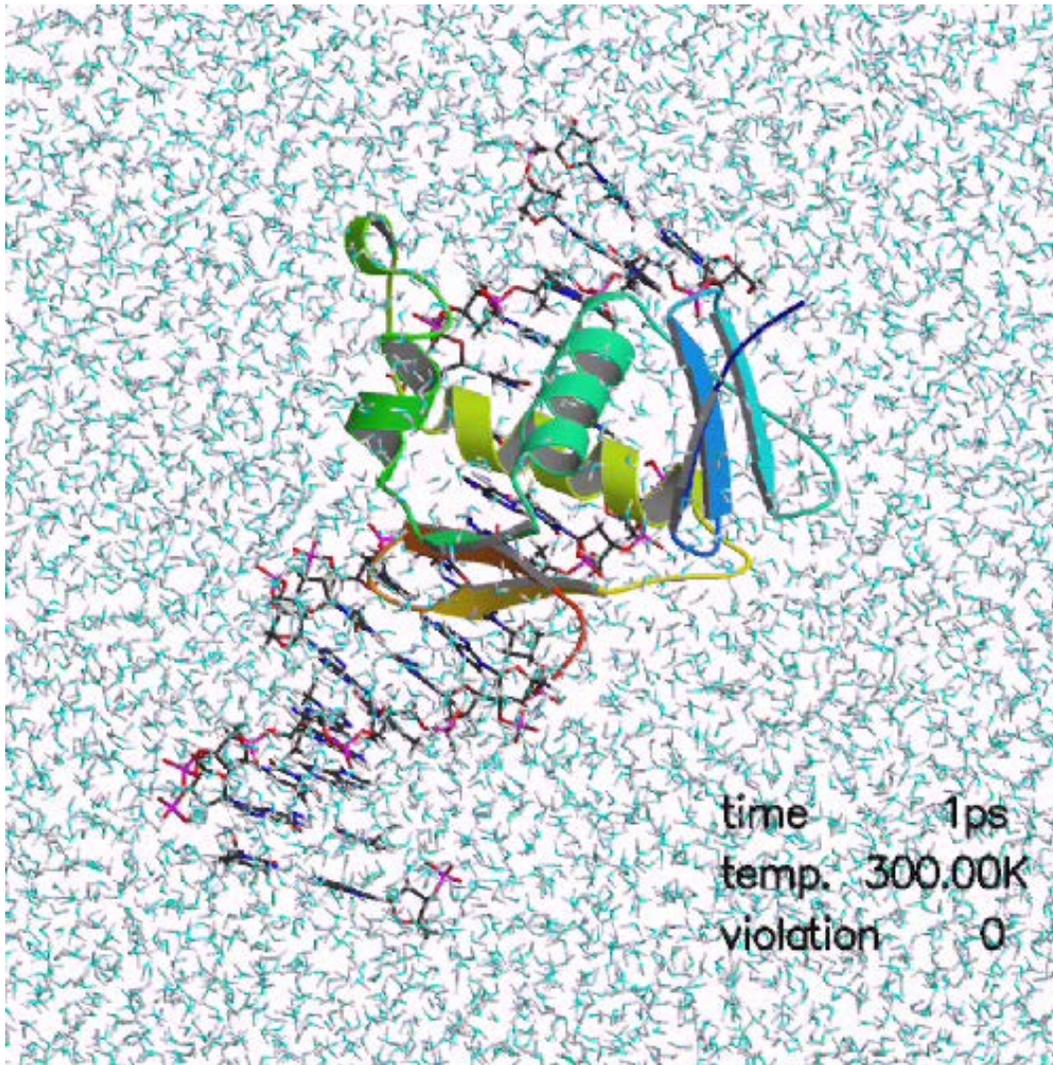
方法2(物理化学MD)

MDの結果の検証 (S²など)



- 物理化学的MDの計算結果とNMR実験結果を比較
- 利点: 物理化学的現象の理解. MDの精度向上.
- 欠点: 実験情報を完全に再現できるかどうか不明

拘束MDの改良1： 水を入れる



西村先生との共同研究
PhoB(転写因子)の
DNA結合ドメイン

真空中の構造決定
CNS, NMR情報
クーロン力なし

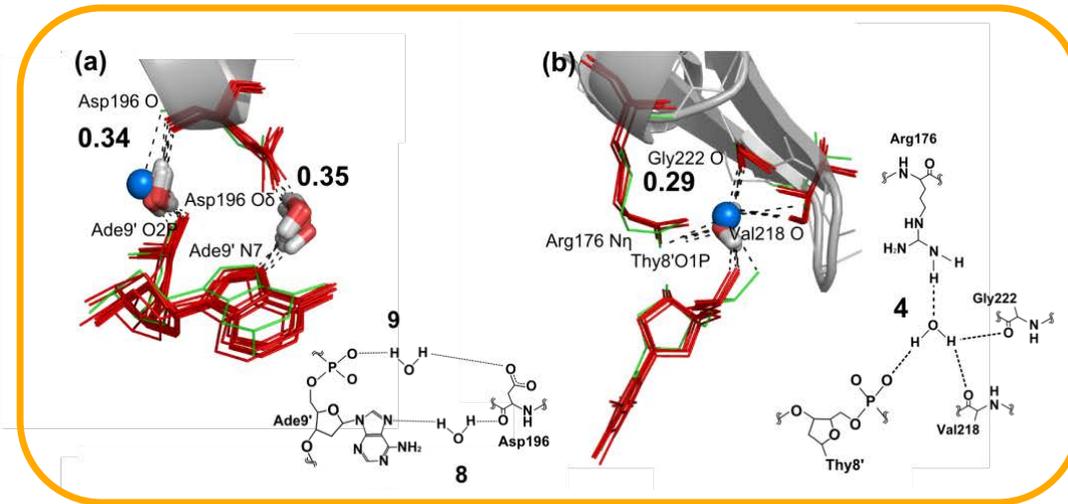


水中の構造精密化
MARBLE, NMR情報
CHARMM27/CMAP
クーロン力もカットなし
(PME)

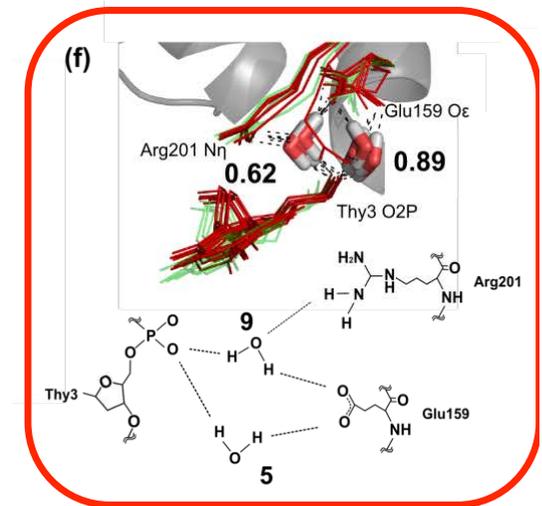
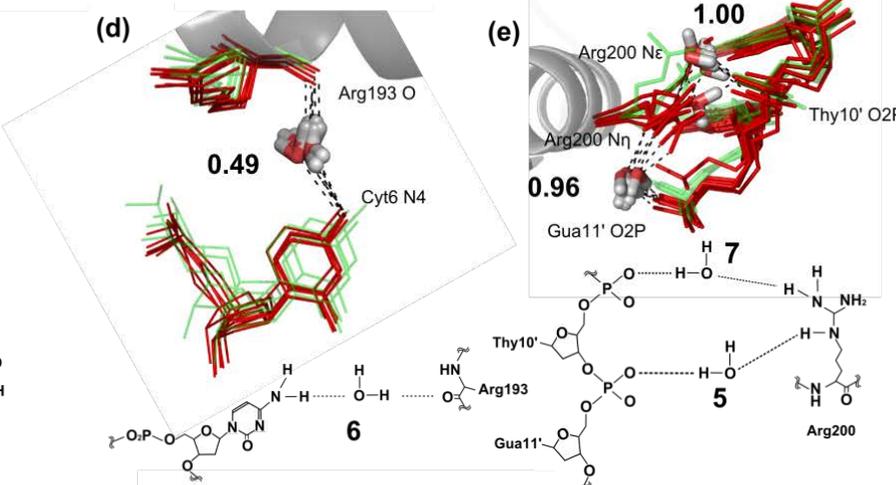
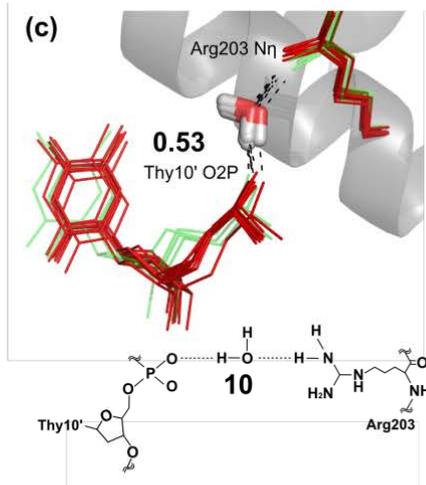
300K→800K→0K

Water-mediated interactions between PhoB and DNA

the same water position in NMR and crystal structures



Glu159



red lines: NMR structures
green lines: crystal structures

control DNA binding activity

MDとNMRの連携法

方法1(拘束MD)

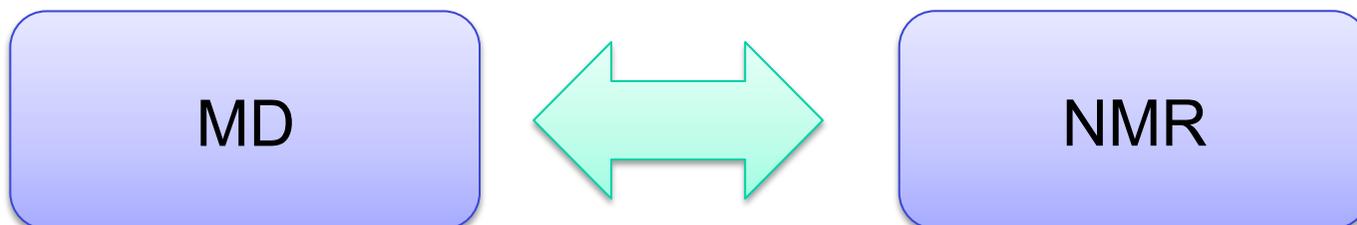
拘束条件(NOE, RDC等)を
入れて計算



- 構造決定のためのツール
- 利点: 実験情報を構造上に再現
- 欠点: 純粋な物理化学的シミュレーションではない

方法2(物理化学MD)

MDの結果の検証 (S²など)



- 物理化学的MDの計算結果とNMR実験結果を比較
- 利点: 物理化学的現象の理解. MDの精度向上.
- 欠点: 実験情報を完全に再現できるかどうか不明

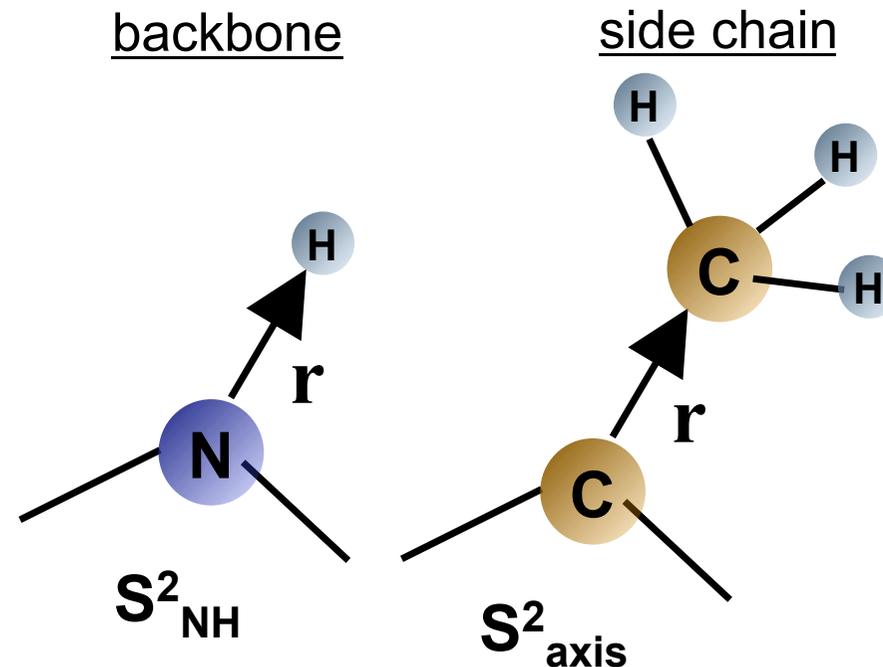
NMR: Order Parameter

NMR provides information on dynamics of both backbone and side chains.

Order parameters (S^2) represent the amplitude of fluctuation of bond vectors.

MD simulations without any restraints were conducted for 10 ns x 20 structures.

Calculated S^2_{NH} and S^2_{axis} were compared with experimental data.



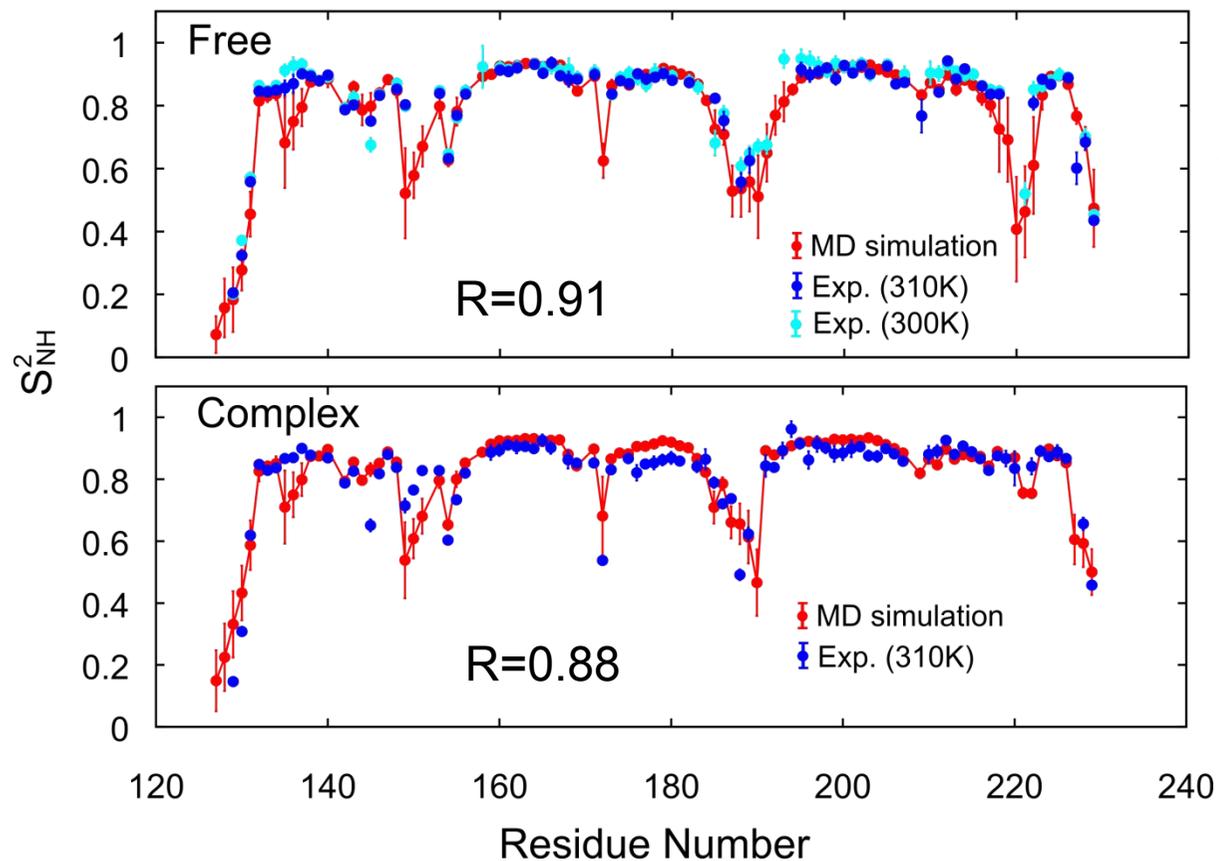
$$\Phi_{ij} = \frac{r_i r_j}{r^2}$$

$$S^2 = \frac{3}{2} \text{tr} \langle \Phi \rangle^2 - \frac{1}{2} (\text{tr} \langle \Phi \rangle)^2$$

$$i, j = x, y, z$$

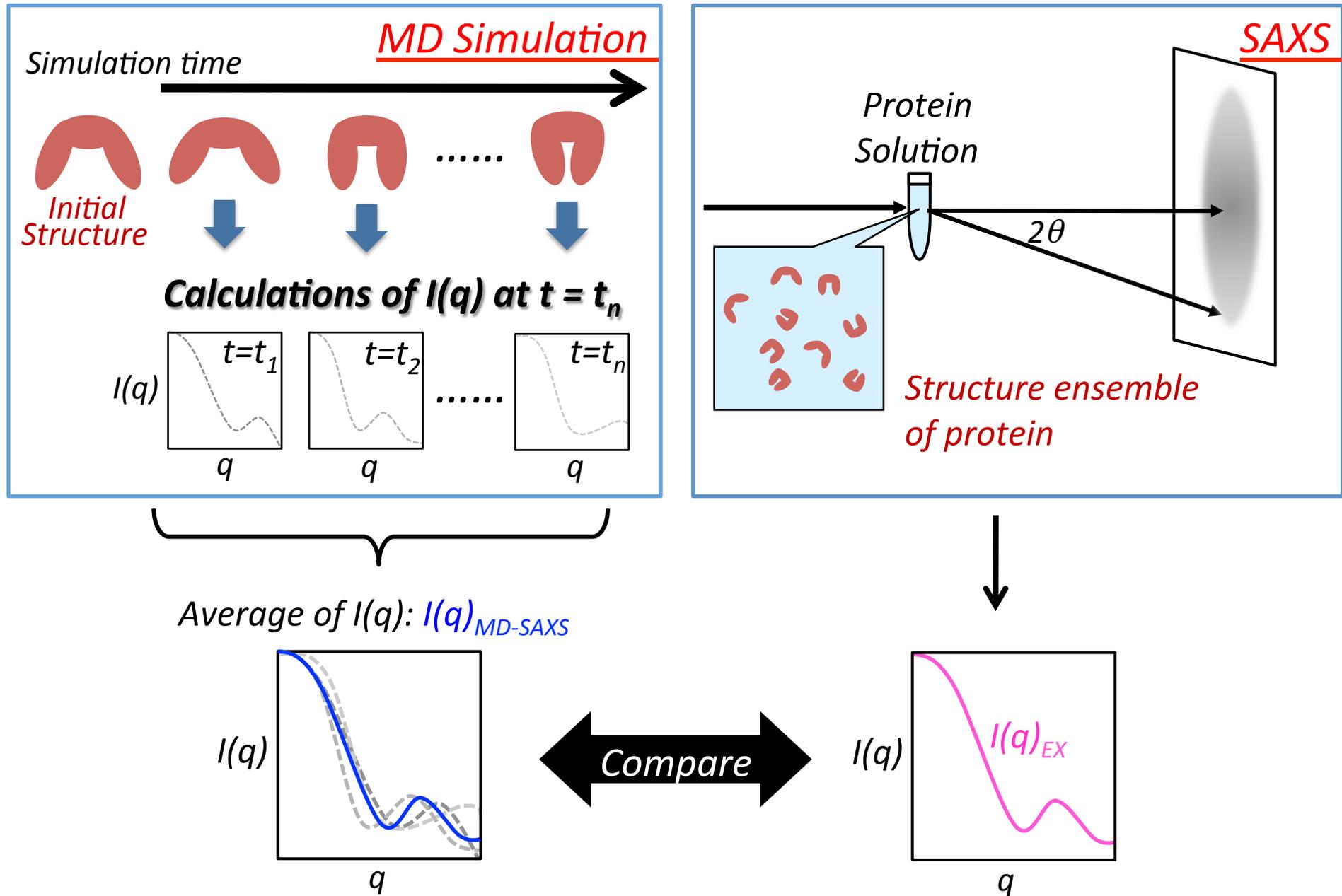
Best, Clarke, Karplus, J. Mol. Biol. 349, 189, (2005)

Model-free order parameter S^2_{NH} -backbone dynamics-



- ★ Agreement with experiments
- ★ Small changes between free and complex except for $\beta 6$ - $\beta 7$ loop
- ★ Flexible regions:
 - ✓ transactivation loop ($\alpha 2$ - $\alpha 3$ loop)
 - ✓ Arg172 in $\beta 5$
 - ✓ $\beta 3$ - $\beta 4$ loop

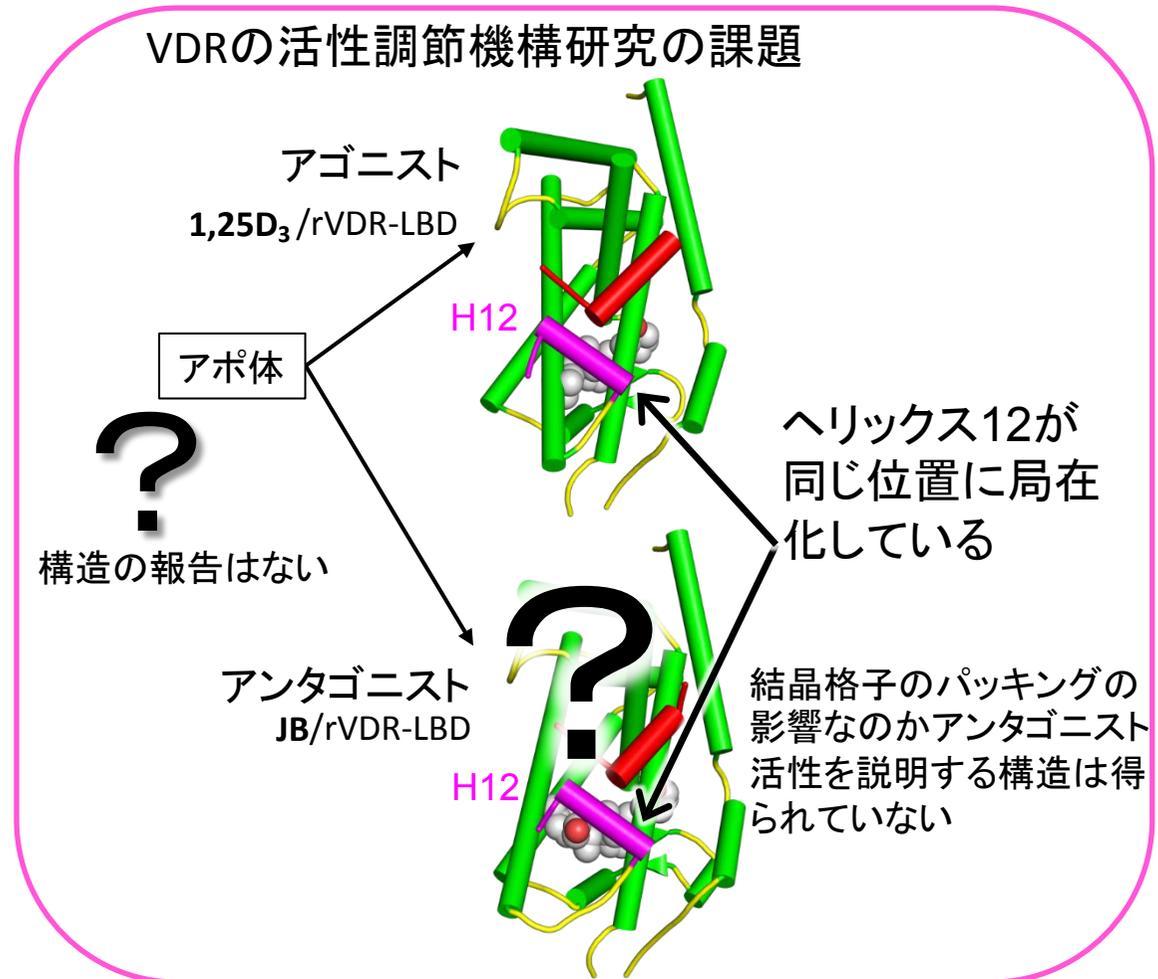
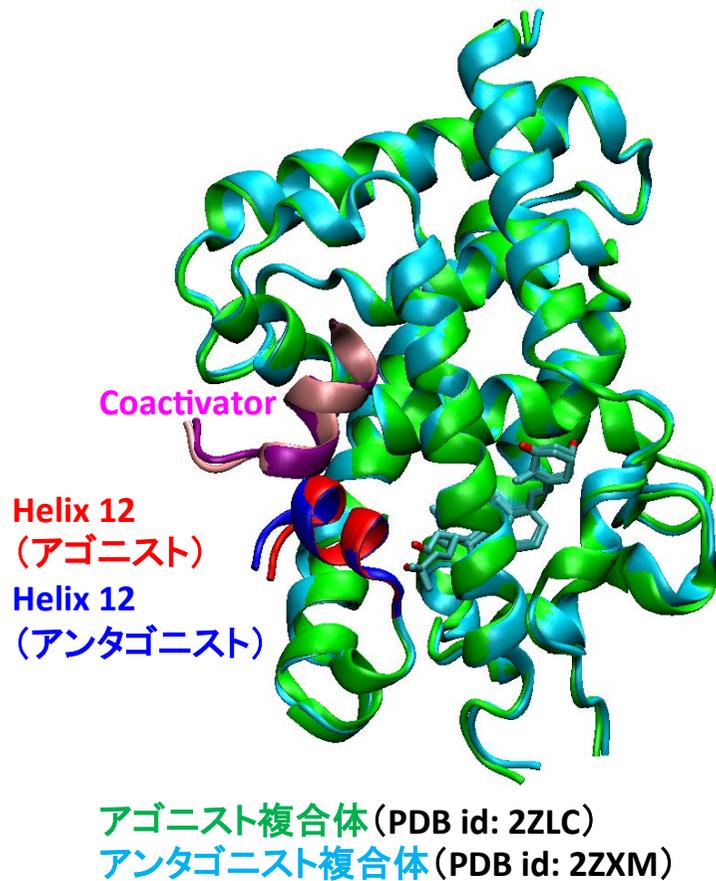
Principle of MD-SAXS



○ビタミンD受容体リガンド結合ドメインの結晶構造

VDRの活性調節機構を、構造変化に基づいて明らかにする事が求められていた。

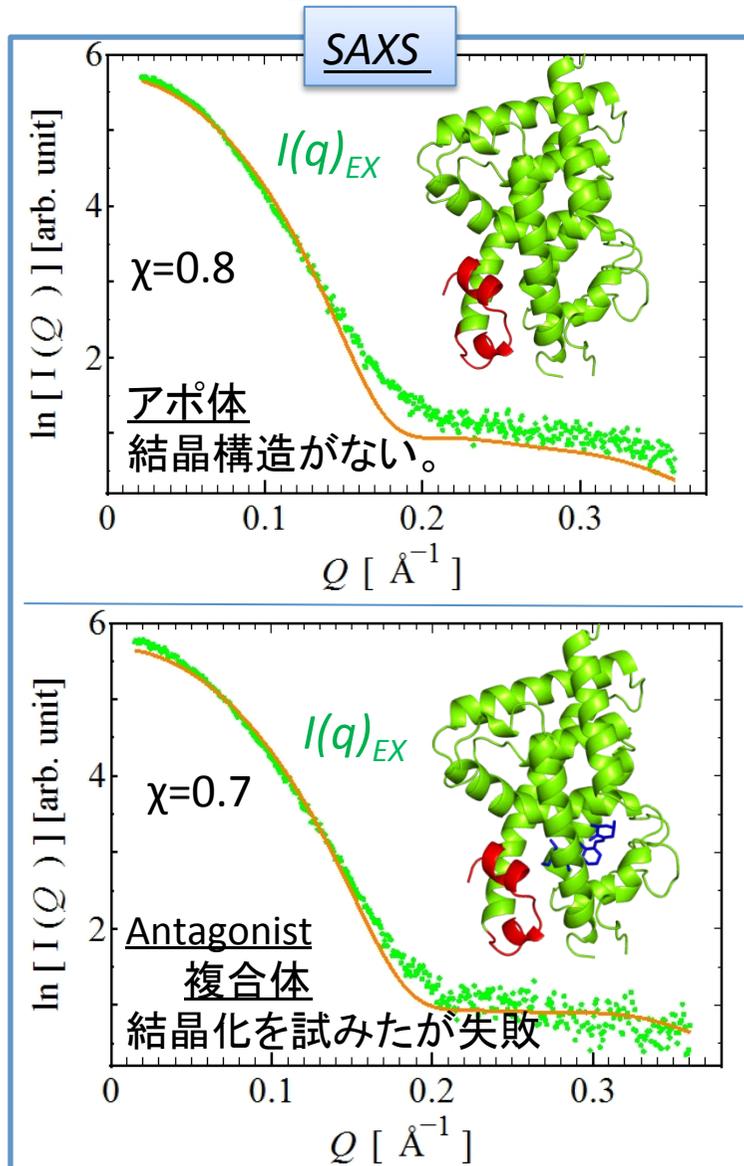
- ・アポ体の結晶構造は報告されておらず、ヘリックス12形成について知見が乏しい。
- ・アゴニスト/アンタゴニスト複合体の結晶構造は報告されているが、両者の構造に違いがない。



3. ビタミンD受容体リガンド結合ドメインの溶液構造探索

Y. Anami, N. Shimizu, T. Ekimoto, D. Egawa, T. Itoh, M. Ikeguchi, and K. Yamamoto; *J. Med. Chem.* **59**, 7888 (2016).

結晶化によるクリスタルパッキングの影響がない、生体内に近い溶液中で溶液構造解析を行った。アポ、Antagonist複合体の単分散状態の散乱データを得たが、それらの溶液構造の解釈には、**低分解能に起因する困難**があった。

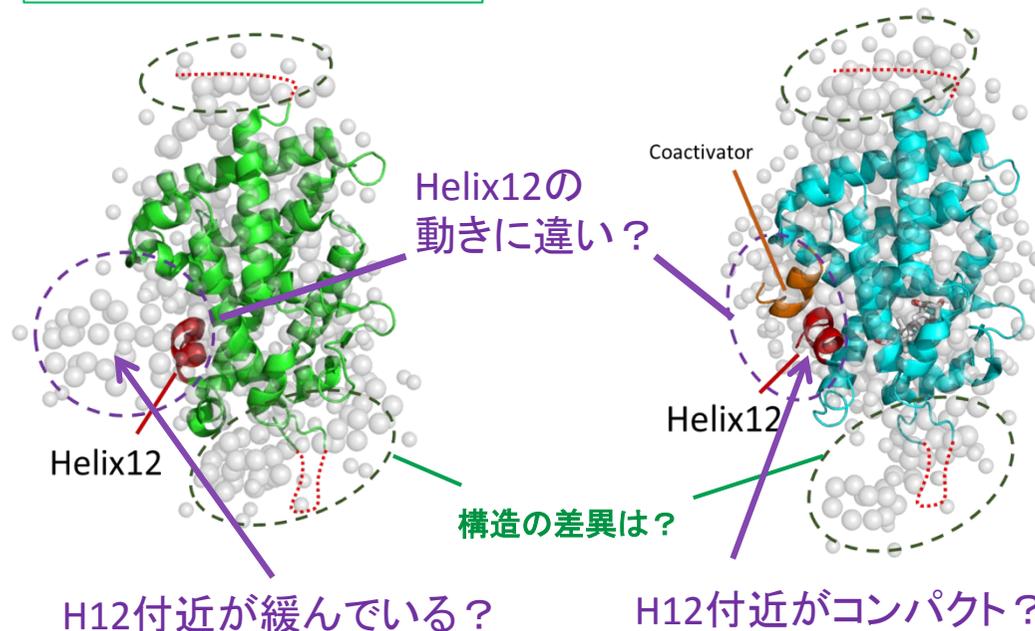


Ab-initio構造解析

OGASBORで計算したAb-initioモデルと結晶構造の比較

・アポ体の溶液構造と
Agonist複合体結晶構造
(PDB id: 2ZLC)との比較

・Antagonist複合体の
溶液構造と結晶構造
(PDB id: 2ZXM)との比較



(欠損部位・構造変化の両方の意味で)
結晶構造で見えていない領域に、特徴的な違いがある。

(Ab-initio構造解析の図は昭和薬科大学 穴見康昭さんよりご提供いただいた。)

3. ビタミンD受容体リガンド結合ドメインの溶液構造探索

* MD-SAXS関連構造解析の実行

①モデリング



②MDの実行



③理論散乱曲線を計算し
実験値と比較



④溶液構造モデル
の提案



⑤MD構造の詳細
を議論

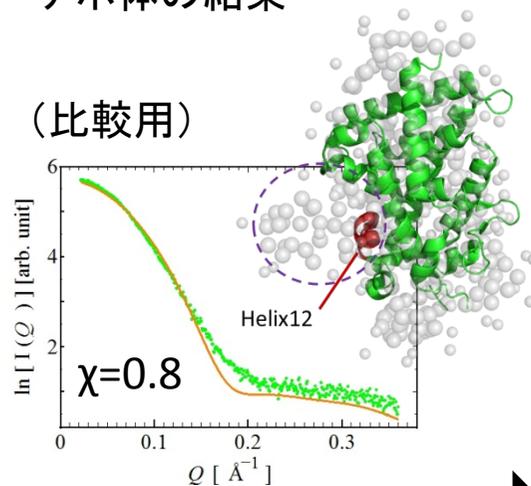


(⑥フレキシブルな
構造の解析)

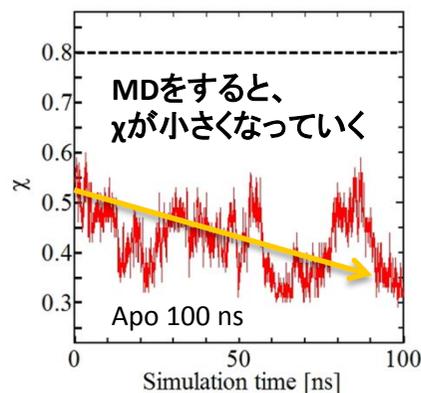
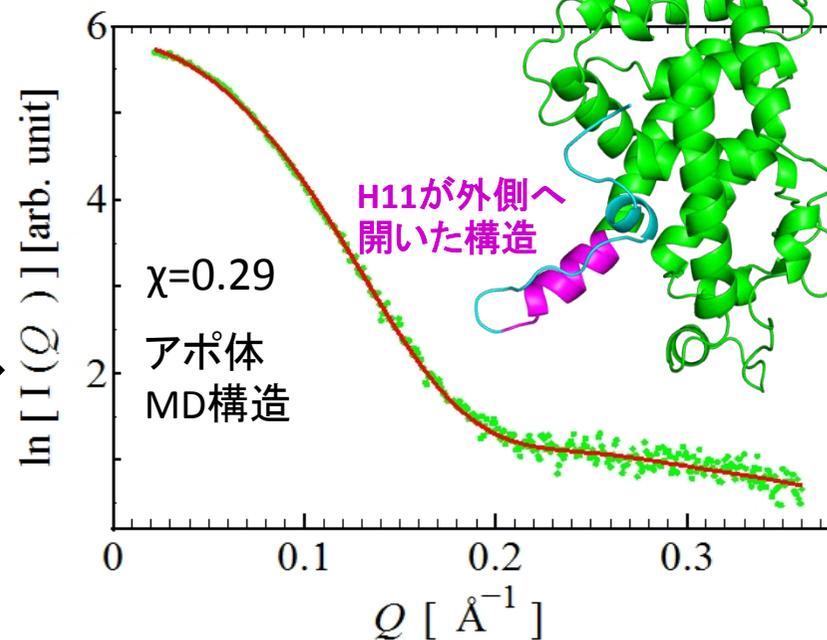
○SAXS実験結果と一致した溶液構造モデル

・アポ体の結果

(比較用)



100 ns MD中で、
最も χ の低い構造のプロファイル



実験溶液中では、上記の構造を平均状態とする構造分布があると考えられる。

なぜ、結晶構造は合わなかったのか...

3. ビタミンD受容体リガンド結合ドメインの溶液構造探索

* MD-SAXS関連構造解析の実行

①モデリング



②MDの実行



③理論散乱曲線を計算し
実験値と比較



④溶液構造モデルの提案



⑤MD構造の詳細を議論

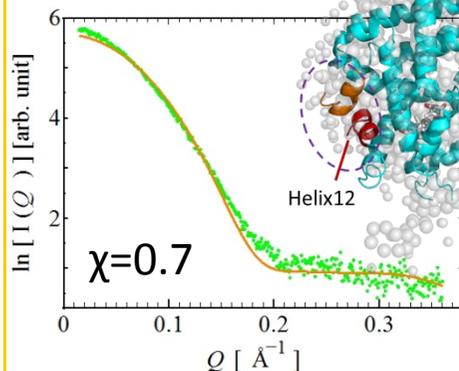


⑥フレキシブルな構造の解析

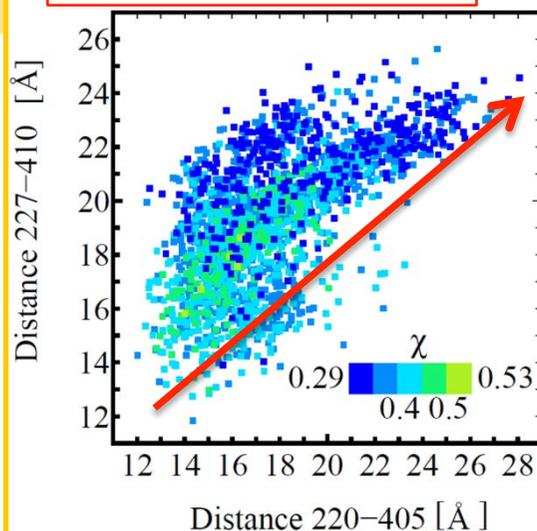
○SAXS実験結果と一致した溶液構造モデル

・Antagonist複合体の結果

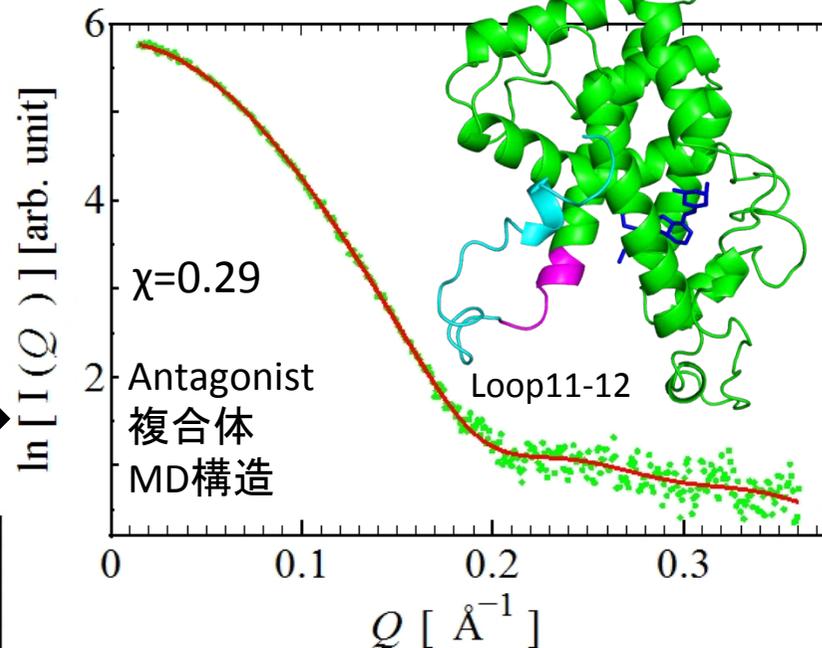
(比較用)



入口が広がると、 χ が低くなる



100 ns MD中で、
最も χ の低い構造のプロファイル

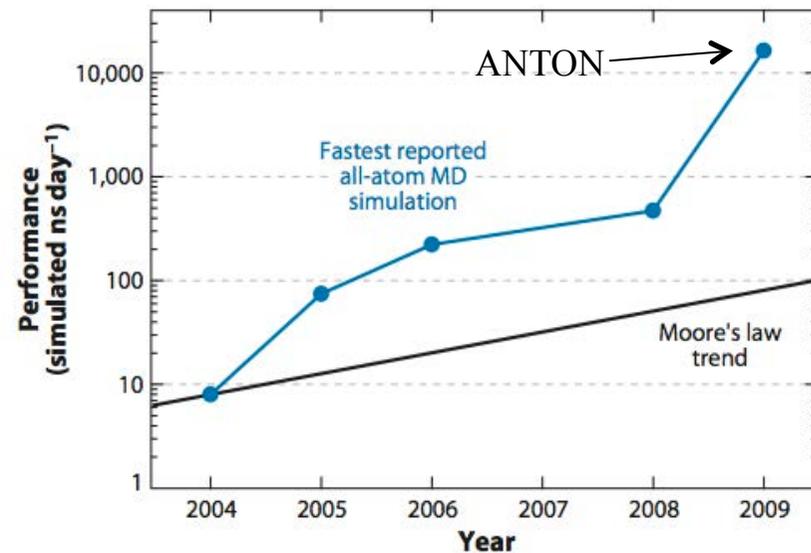
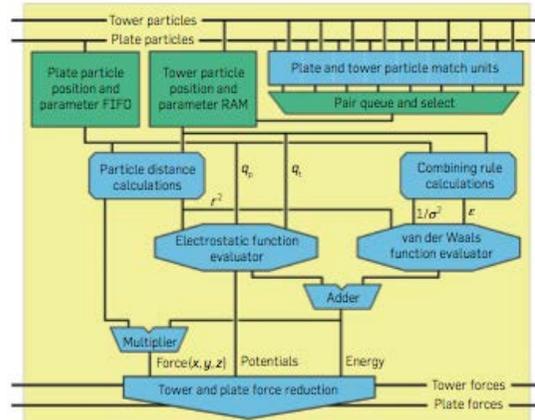


アポ体同様に、ポケット入口が開いた構造をとっていた。

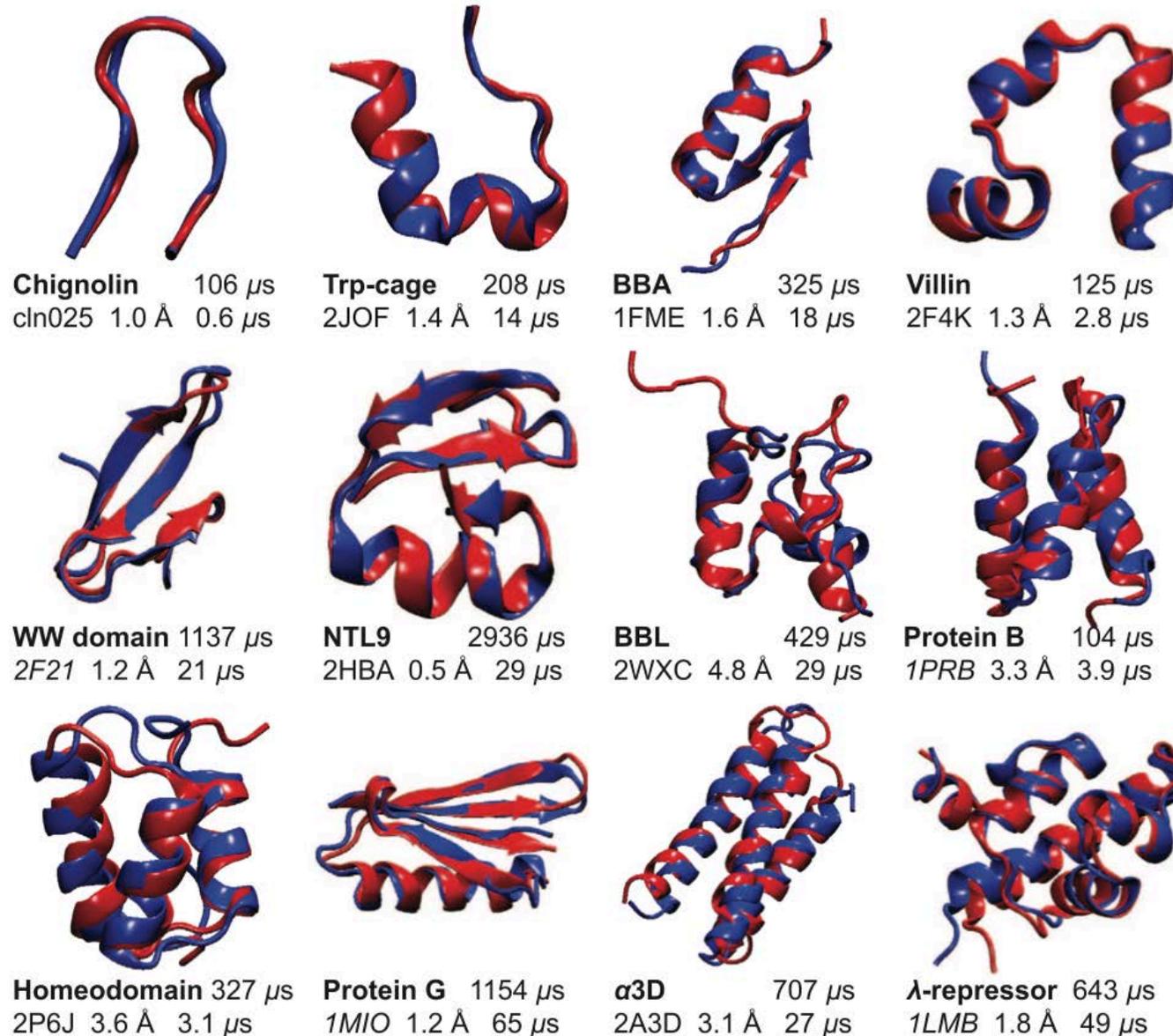
ANTON: MD専用スーパーコンピュータ

ANTON 2009 Gordon Bell Award
ANTON2 2014 Gordon Bell Award

Shaw et al. Comm. ACM (2008)
Shaw et al. Proc. SC'14 (2014)



ANTON: MDによるフォールディング



赤: 実験による
立体構造

青: MDによる
立体構造

12個のタンパク質でフォールディング

Shaw et al. Science, 2011

汎用スパコンでの動的機能解析

数十 $\mu\text{s}/\text{day}$ 計算可能な専用計算機（ANTON2等）と違って、我々が使用可能なのは汎用スパコンである

一つの計算では、汎用計算機だと、現状からの飛躍的向上は、非常に困難である

しかし、汎用スパコンでは、ノード or CPUコアは沢山ある



レプリカを沢山用意して、統計的に計算を行うことで、実質的に、長時間ダイナミクス情報を得る

汎用スパコンでの動的機能解析

レプリカ系のMD計算

方法① 熱力学的方法

自由エネルギー分布を得る

(取りやすい複数の状態と頻度比を得る)

e.g. レプリカ交換、アンブレラサンプリング等

方法② 速度論的方法

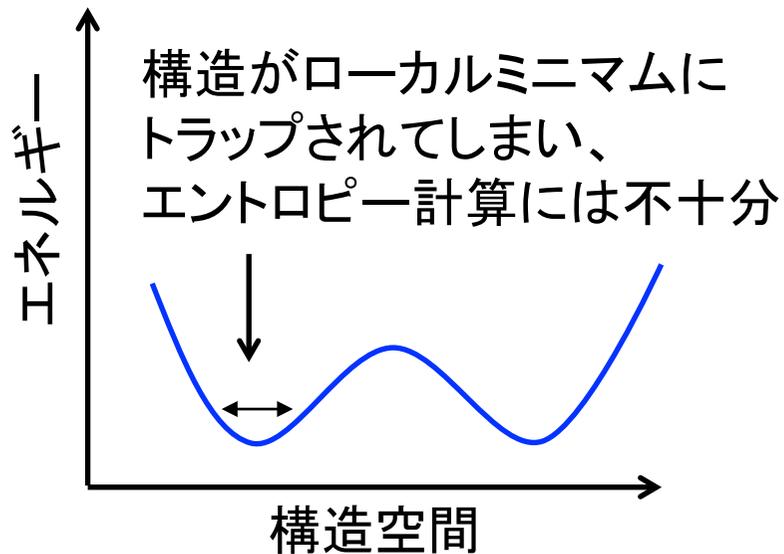
キネティックス (速度論) を得る

(取りやすい複数の状態間遷移を得る)

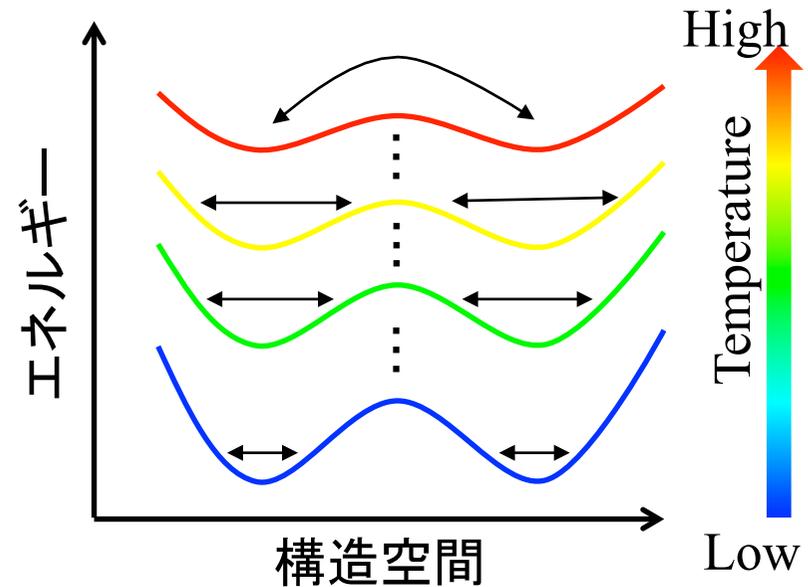
e.g. マルコフ状態遷移モデル (MSM)

温度レプリカ交換MD法

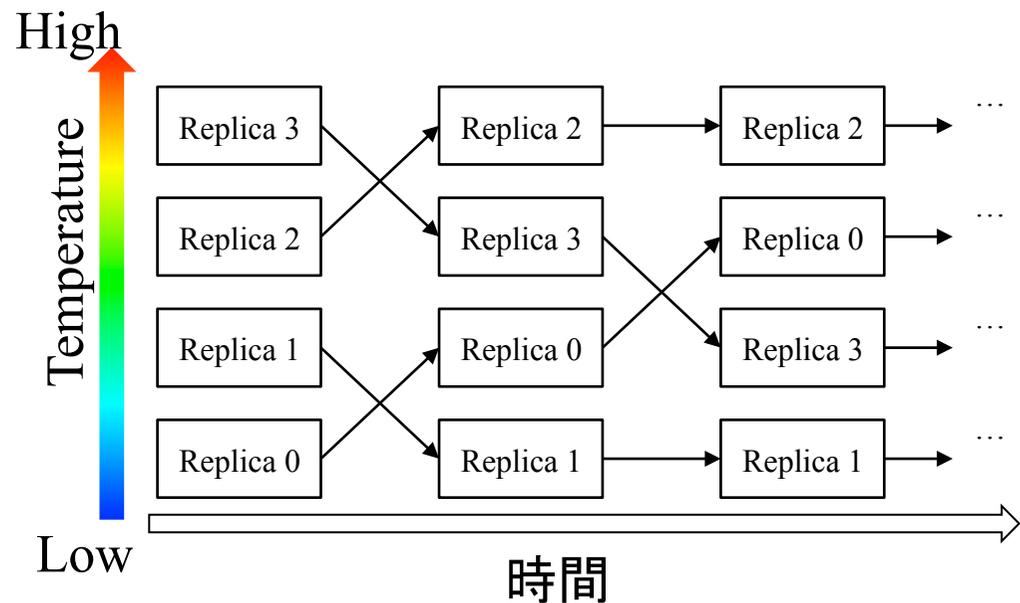
通常のMD法



温度レプリカ交換MD法

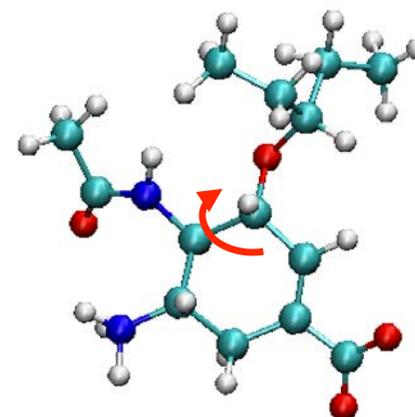
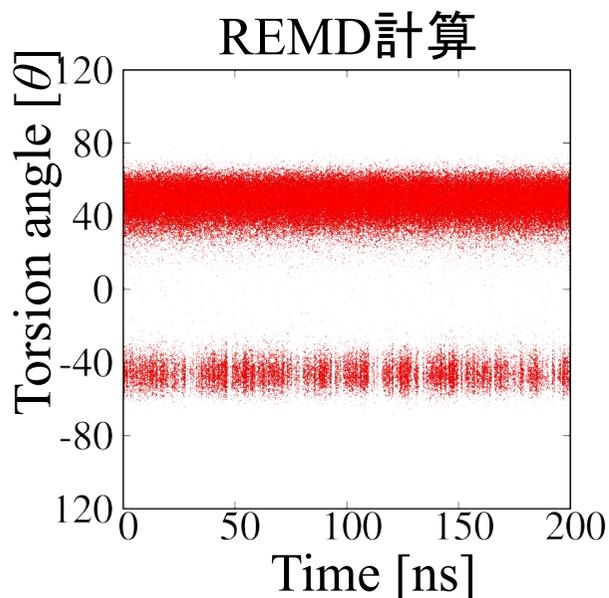
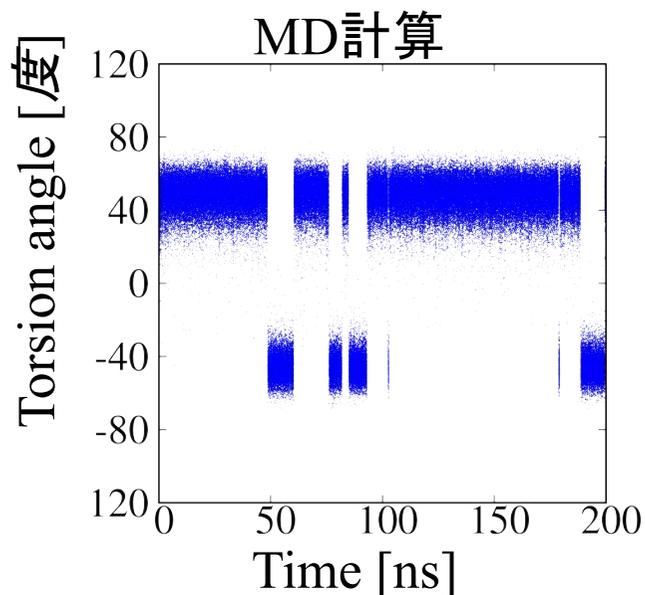


温度レプリカ交換MD法の導入により、幅広い構造サンプリングが可能になる



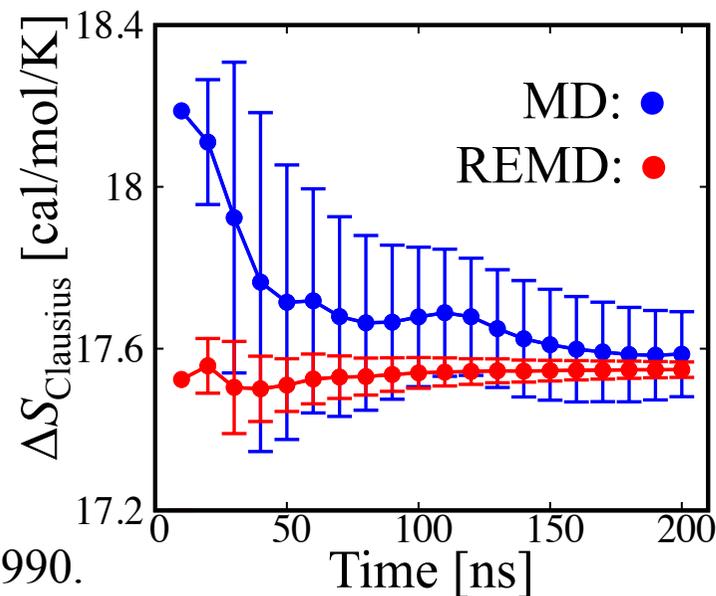
温度レプリカ交換MD(REMD)法により計算の収束性が向上

例: Oseltamivir シミュレーション中のねじれ角の分布

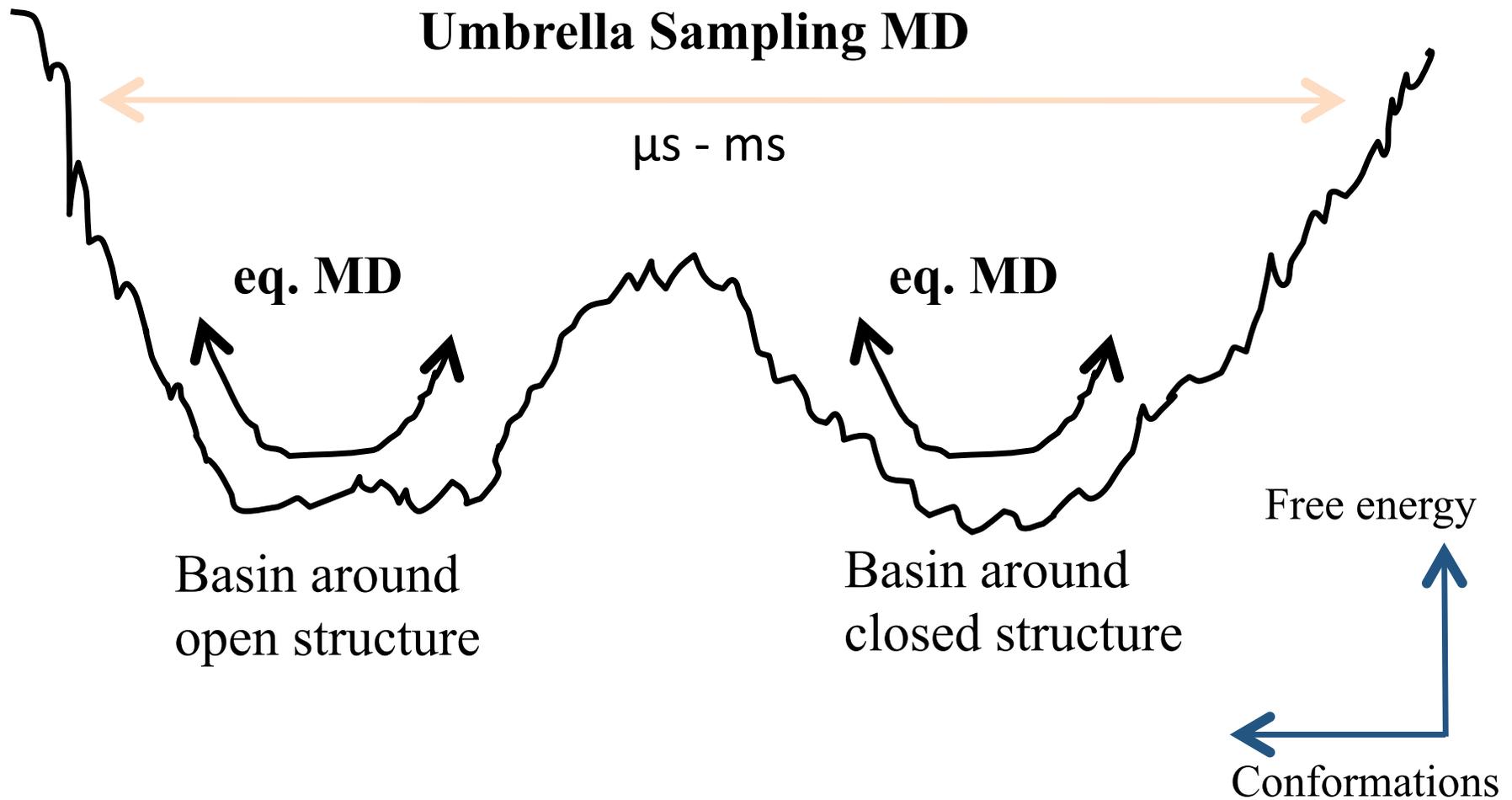


ねじれ角の位置

温度レプリカ交換MDの導入により、
構造分布の収束性が向上し、
構造エントロピー計算の収束性も向上



Method: umbrella sampling simulations

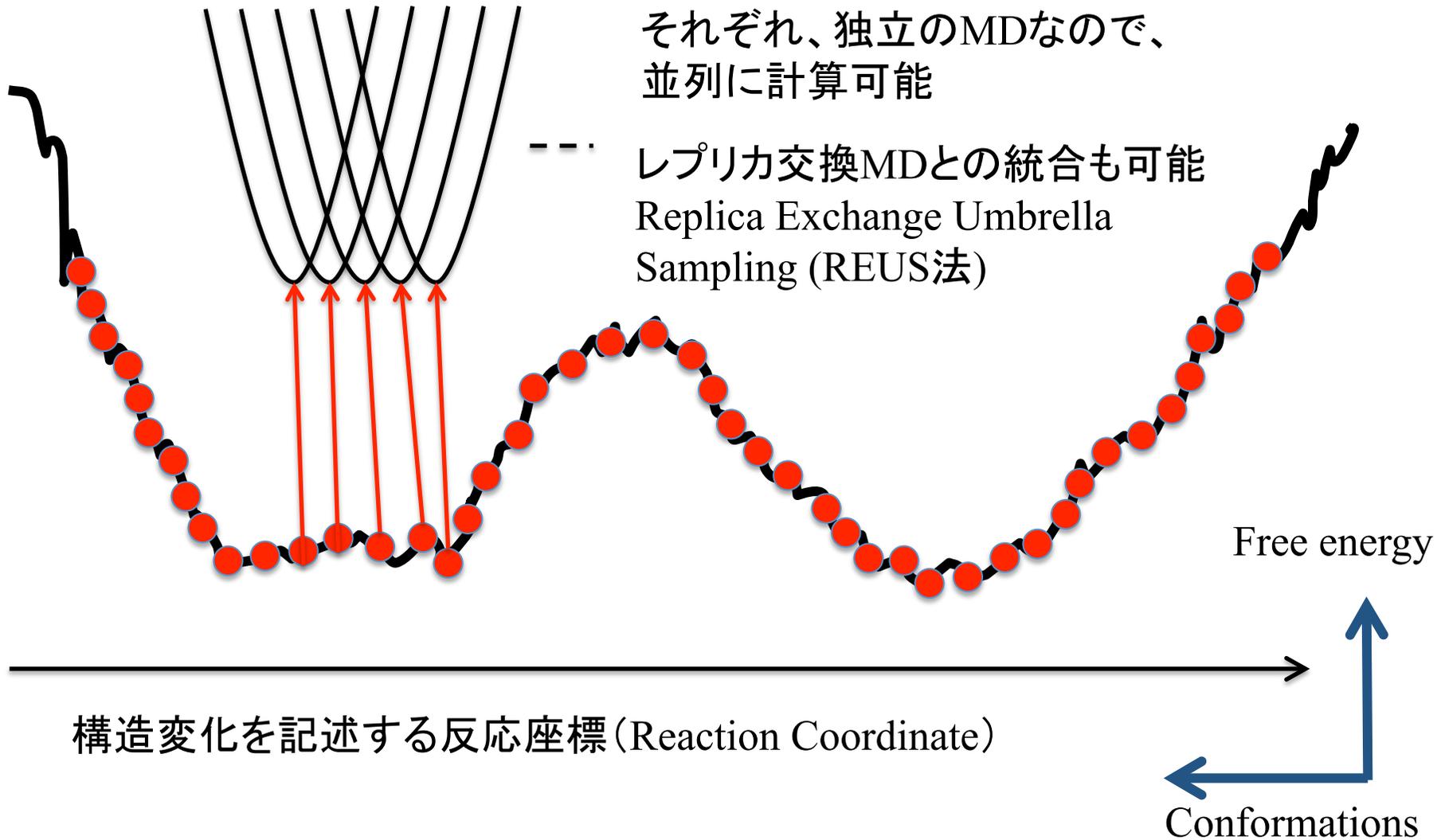


Method: umbrella sampling simulations

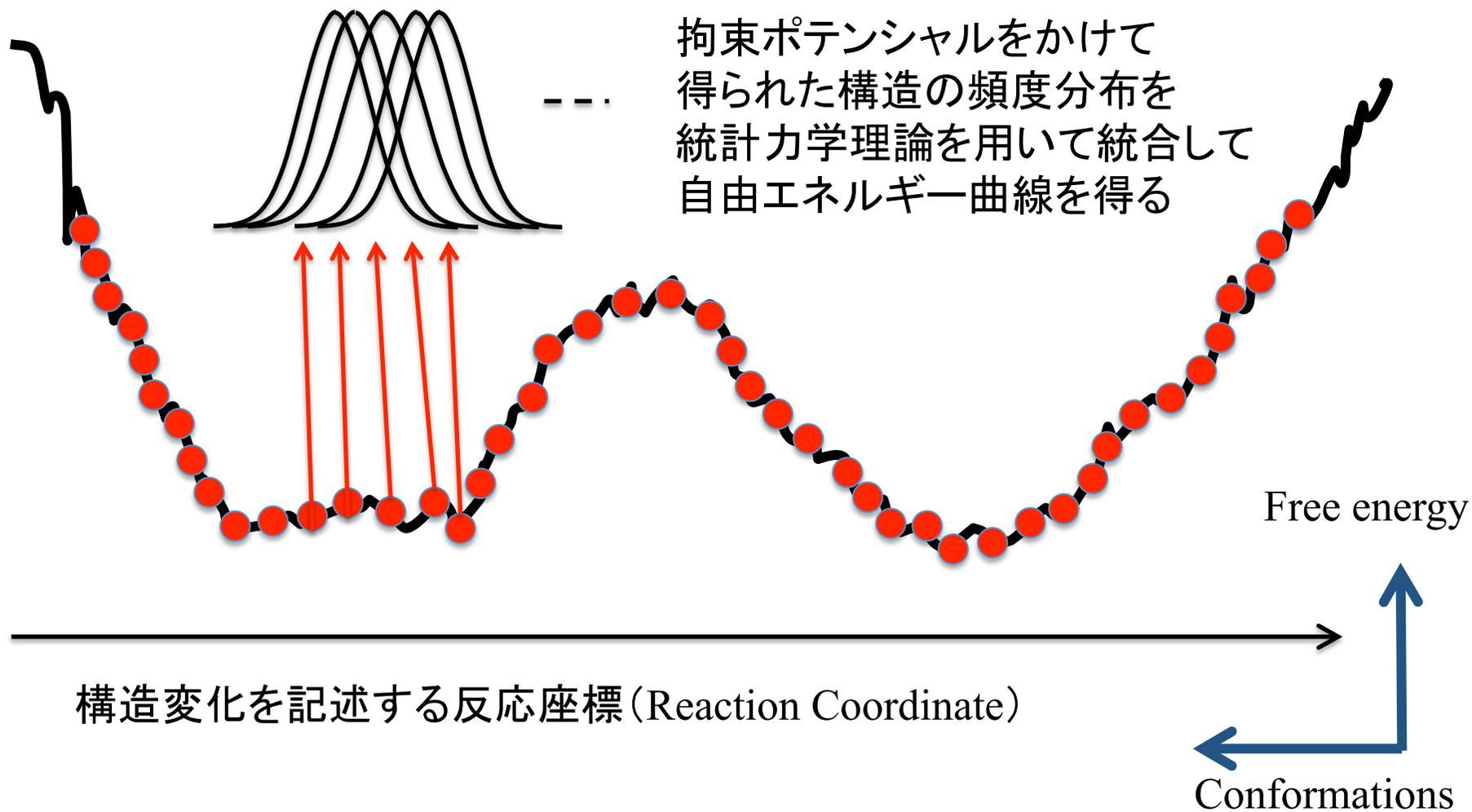
拘束ポテンシャルをかけて
構造サンプリング

それぞれ、独立のMDなので、
並列に計算可能

--- レプリカ交換MDとの統合も可能
Replica Exchange Umbrella
Sampling (REUS法)

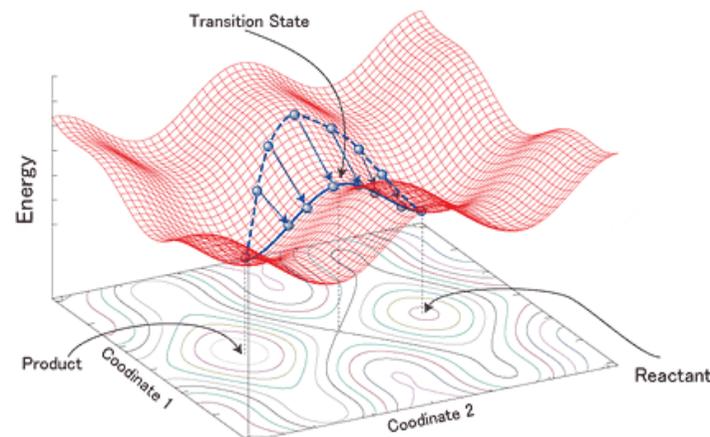
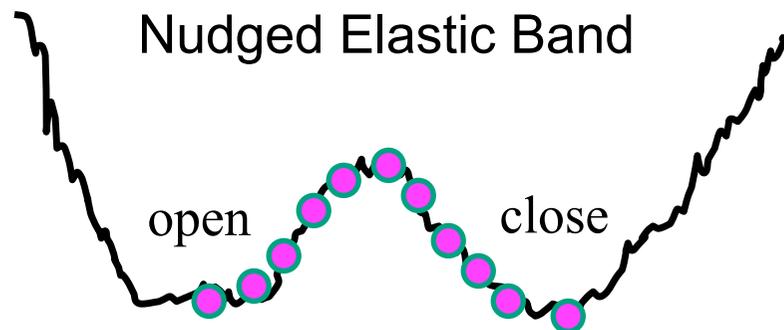


Method: umbrella sampling simulations



構造変化に対するFree Energy Simulations

1. Initial Path



2. Umbrella Sampling Simulations

For each intermediates along path, simulations with restraint,

$$w_j = K_{\text{rmsd}} (\Delta D_{\text{rmsd}} - \Delta D_{\text{min}})^2$$

$$\Delta D_{\text{rmsd}} = \text{rmsd}(X, X_{\text{open}}) - \text{rmsd}(X, X_{\text{closed}})$$

was carried out.

Restraint potential is applied to both main chains and side chains.

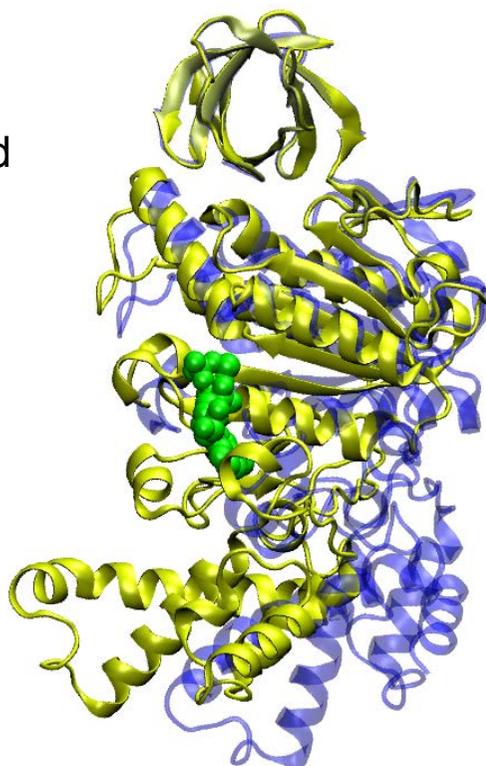
B. Roux *JACS* (2005) C.L. Brooks *PNAS* (2007)

3. WHAM

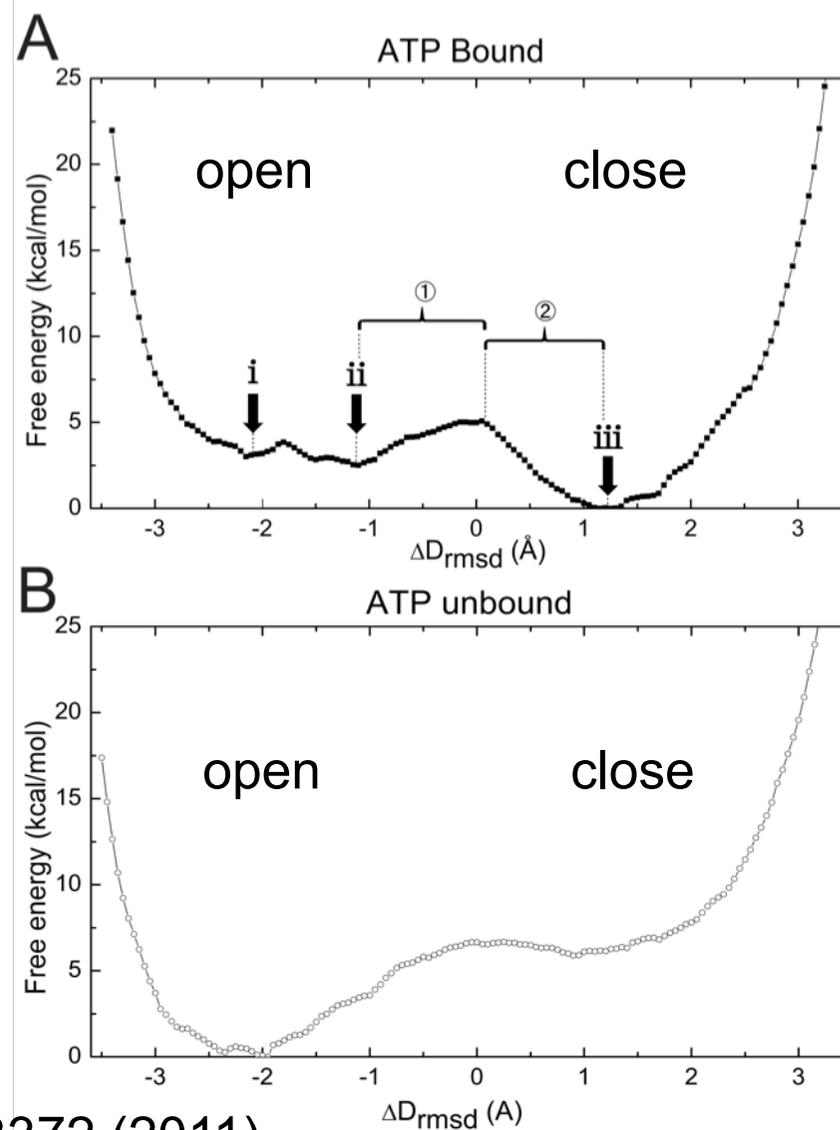
Weighted Histogram Analysis Method was used for removing restraints and calculating free energy profiles along structural transition between open and closed conformations.

Free Energy Profiles for Open-Close Transition

blue: open
yellow: closed
green: ATP

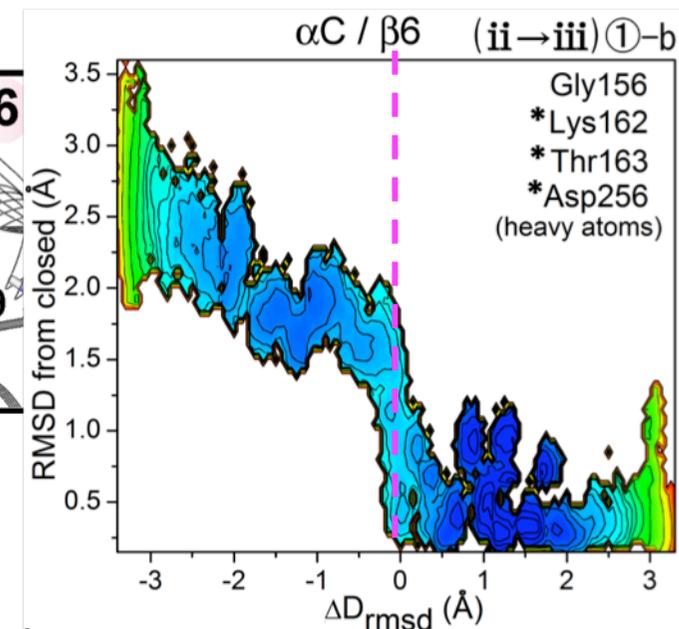
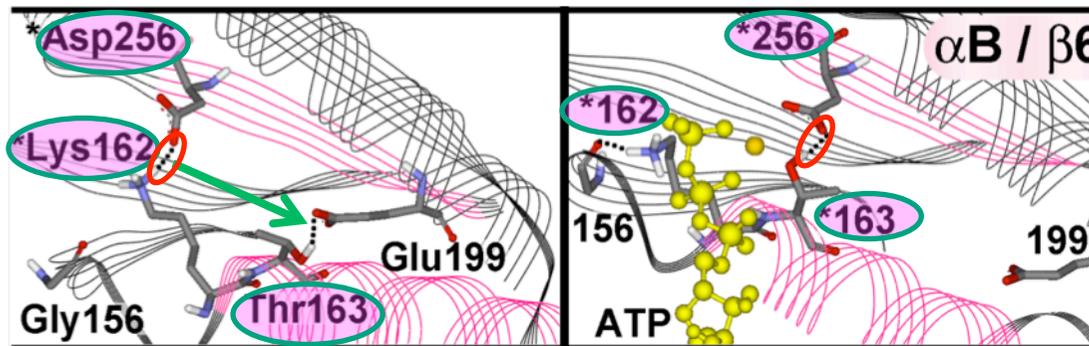
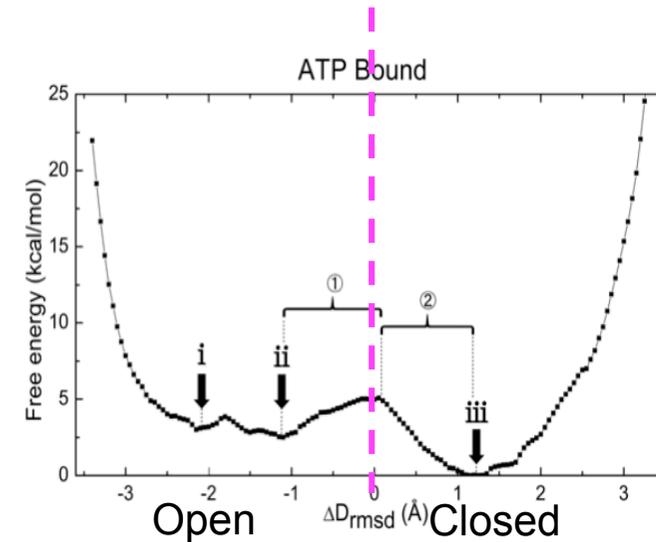
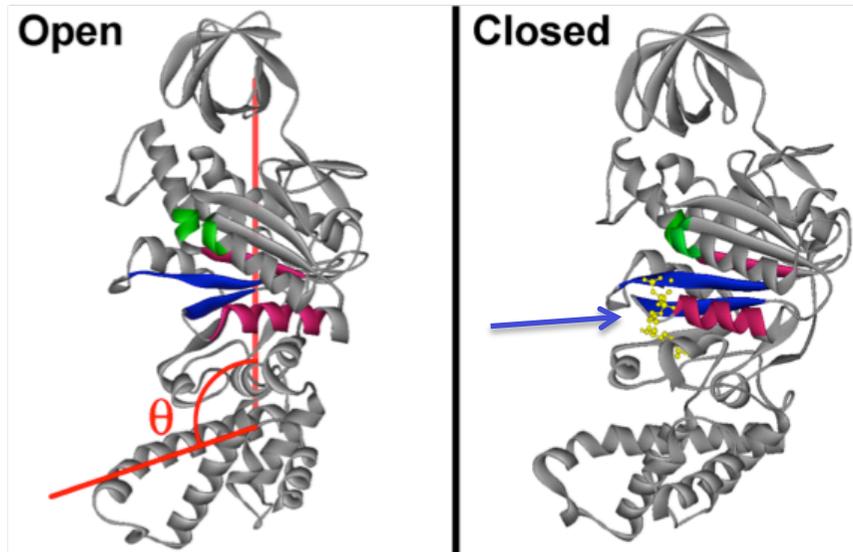


- ✓ Meta-stable state is found:
ATP-bound open conformation.
→ encounter complex



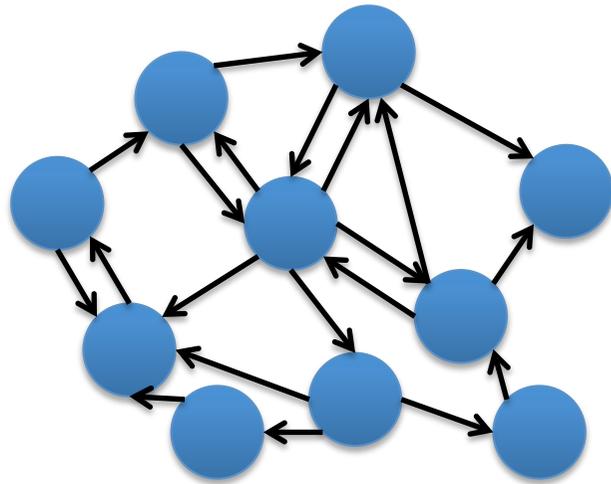
Ito, Oroguchi, Ikeguchi, JACS, 133, 3372 (2011)

$ii \rightarrow iii$ ① transition: H-bond in P-loop



H-bond partner of Asp256 is changed from Lys162 to Thr163 (P-loop). Mutation of Asp256, Lys162, or Thr163 results in remaining open conformation even with ATP bound.

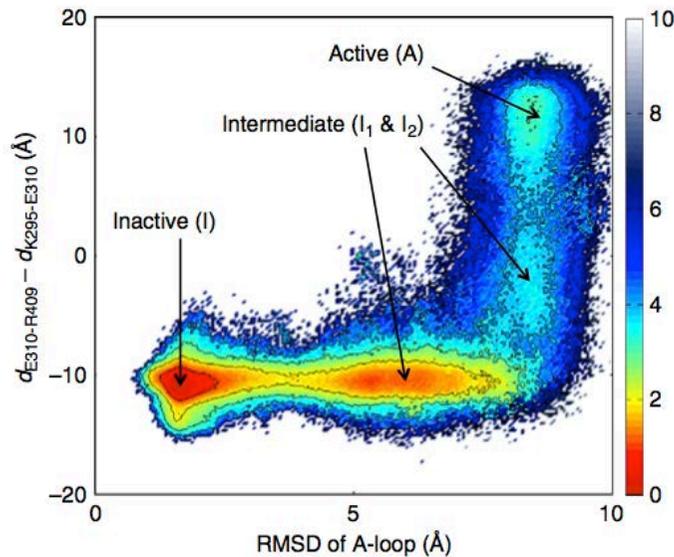
マルコフ状態遷移モデルによるダイナミクス解析



各、短いシミュレーションにより、状態間遷移確率を求める

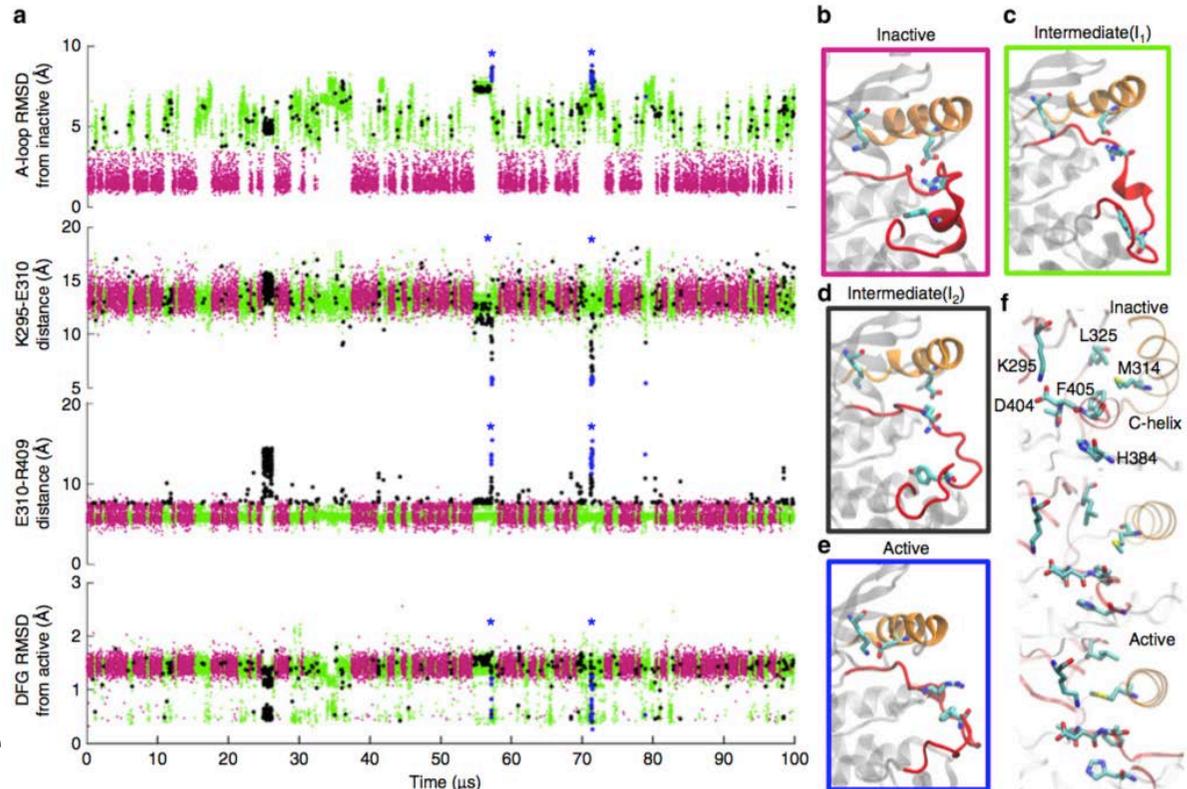


状態間時間発展として、長時間ダイナミクスを記述する



Pande et al. Nature Comm. (2014)

12,000 runs, total 150 μ s;
~12.5 ns/run etc. Folding@home



まとめ

- 生体分子の動きが機能に重要
- 分子動力学シミュレーションでは、様々な力場が提案され、今も継続して改良中である
- 分子動力学シミュレーションの解析は、その機能ごとに様々である
 - 例：RMSD, RMSF, 相関行列, PCA等
- 分子動力学シミュレーションは、様々な応用されている。
 - 実験との連携も重要な点
 - 超並列計算への展開：レプリカ系の計算