

課題名 (タイトル) :

## 蛋白質機能解析のための自由エネルギー地形計算手法の開発

利用者氏名 : ○原田 隆平

所属 : 計算科学研究機構 プログラミング環境研究チーム

## 1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

蛋白質は自らも熱揺らぎを効率的に利用し、柔軟に構造変化することで機能を発現する。従って、「蛋白質の柔らかなダイナミクス」を調べることは、機能を理解する上で重要である。さらに、それらのダイナミクスを規定している自由エネルギー地形を探索することは必要不可欠である。蛋白質は多自由度複雑系であるので、地形は多数の準安定状態が存在する起伏の激しい多谷構造をもつ。このため、有限時間の分子動力学シミュレーション (Molecular Dynamics simulation: 以下 MD) による構造探索では、準安定構造の1つにトラップされてしまう可能性が高く、効率的な地形計算が困難である。また、機能に関係する「タンパク質の柔らかなダイナミクス」は、通常の MD では到達することが難しい長時間領域(ミリ秒～秒)で観測される「遅い運動」である場合が多い。例えば、フォールディング過程や、機能に関係するドメイン運動などである。最近では、蛋白質に特化した MD 専用計算機を設計・構築することにより、ミリ秒スケールの MD が実現し始めているが、専用計算機を使用できる環境は限られるため、現実的に使用しやすい PC クラスタを用いて、「タンパク質の柔らかなダイナミクス」を実用的な実行時間内で観察することは難しい。さらに、これらの運動は確率的に非常に起こりにくい「レアイベント」であるため、たとえ長時間の MD を実行できたとしても、抽出できる保証はない。以上の様な問題点から、蛋白質機能に関係する重要な構造変化を抽出し、ダイナミクスを規定している自由エネルギー地形を効率的に計算可能な手法を構築することが本研究課題の目的である。

な計算コストで機能に関係する重要な構造変化を抽出するためには、一般的に外部摂動などを用いて強制的に構造変化させることが多く、パラメーター設定には経験を要する。また、不適切な外部摂動を設定してしまった場合、アーティファクトが含まれる可能性が高くなる。このような理由から、第 1 に、出来るだけパラメーター設定に煩わされず、かつアーティファクトを含まない自然な形で構造変化を抽出する手法として「超並列カスケード型 MD シミュレーション」である、Parallel CaScade Molecular Dynamics MD (以下、PaCS-MD)を開発した。PaCS-MD (図1)は、始構造と終構造の構造類似性を表す物理的指標を定義し、指標をもとに始構造が終構造へ収束していくように独立かつ多数の短時間 MD を繰り返すという思想のもとで開発した。適当な指標の一つとして、構造間の構造類似性を表す平均自乗距離(Root Mean Square Deviation: RMSD)などが考えられる。もちろん、様々な物理量を指標として定義することができ、複数の指標を定義しても構わない。PaCS-MD では、MD における初期速度再配分によって再度構造サンプリングを行うことにより、外部摂動に頼ることなく構造変化経路を抽出することが可能である。下記に PaCS-MD の手順を記す。(i) 終構造に対する構造類似性等の指標をもとに、MD のスナップショットをランキングする。(ii) 構造類似性が高い上位の複数のスナップショット(以下、ランキング上位 10 構造)を初期構造とし、初期速度再配分のあと短時間 MD (以下では 0.1 ns) を実行。(i), (ii) のサイクルを、スナップショットが十分終構造に近くなるまで繰り返す。サイクルを繰り返すごとに、終構造に類似したスナップショットが選択され、構造変化経路を探索していく。

## 2. 具体的な利用内容、計算方法

本年度は、研究目的の 1 つである蛋白質機能に関係する重要な構造変化抽出法を開発した。現実的

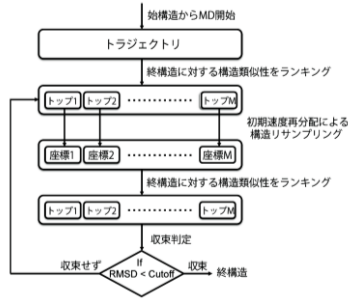


図 1. PaCS-MD のフローチャート

3. 結果

PaCS-MD では、外部摂動を付加することなく、初期構造選択と初期速度再分配により、繰り返し構造サンプリングを実行することにより、構造変化を抽出していく。PaCS-MD のアプリケーションとして、溶媒を含んだ全原子モデルを用いた人工タンパク質である chignolin の folding 過程と加水分解酵素 T4 lysozyme の open/closed 構造変化に適用した結果を示す。Chignolin では始構造として、完全に伸びきったアミノ酸構造をモデリングし、天然構造を終構造(図2左)として PaCS-MD を実行した。T4 lysozyme は、始構造として野生型の open 構造を始構造、closed 構造を終構造(図2右)とした。

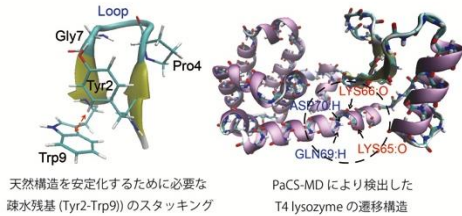


図 2. Chignolin (左) と T4 Lysozyme (右)

Chignolin と T4 lysozyme それぞれについて、5 本の異なる独立な PaCS-MD を実行した。図3に、PaCS-MD のサイクルごとの RMSD の変化を示す。終構造に対する RMSD を指標として、値が小さい(終構造に近い)上位 10 構造を初期構造に選択し、初期速度再分配による PaCS-MD のサイクルを繰り返した。図 3 から分かるように、どちらのシステムも、20 サイクル程度で RMSD が 1.0 Å 以内に収束しており、始構造から終構造への構造変化経路を抽出できたことがわかる。

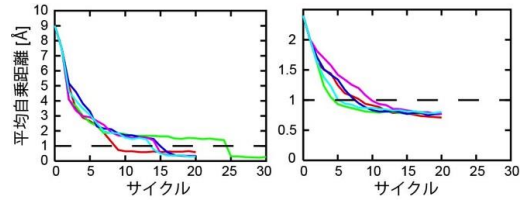


図 3. 終構造に対する平均自乗距離の変化

Chignolin (左) と T4 lysozyme (右)

図 4 は、Chignolin/T4 lysozyme の構造変化において、それぞれのサイクルで、どのように初期構造が選択されていくかを示した「カスケード」である。図 4 において、赤い矢印がサイクルごとに選択された初期構造に対応している。この赤い矢印に対応するトラジェクトリをつなぎ合わせることで、始構造から終構造へいたる仮想的な構造変化経路が生成される(リアクティブトラジェクトリと呼ぶ)。リアクティブトラジェクトリの物理的解釈としては、現実的に可能な構造変化経路候補の1つと考えることができ、様々な用途に利用することができる。例えば、構造変化経路のリファインメントとして、Transition Path Sampling や、String 法の初期経路パスとして利用することができる。また、アンブレラサンプリングを用いて、リアクティブトラジェクトリ周辺の自由エネルギー地形を計算することができる。

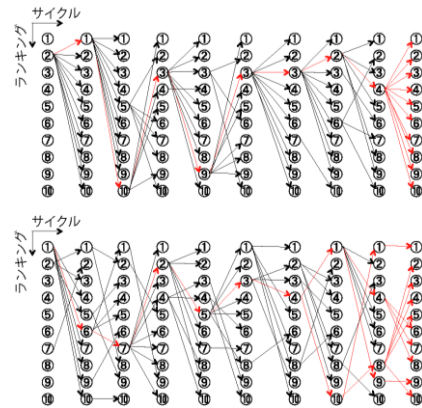


図 4. 構造選択を表すカスケード

(Chignolin/T4 lysozyme, 10 サイクル)

図 5 は、PaCS-MD により抽出したリアクティブトラジェクトリに沿って参照構造を配置させ、その構造まわりのアンブレラサンプリングを実行することにより計算した自由エネルギー地形を示す。

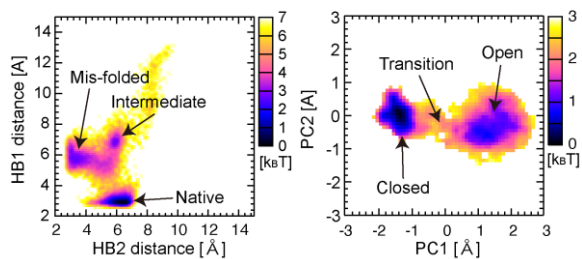


図 5. 自由エネルギー地形. Chignolin (左:主鎖水素結合距離 HB1/HB2 に射影した地形), T4 lysozyme (右:主成分分析により定義した主成分空間 PC1/PC2 に射影した地形).

このようにして, PaCS-MD により抽出した始構造から終構造への構造変化経路を, 自由エネルギー地形計算により定量的に評価することも可能である.

#### 4. まとめ

PaCS-MD が効率的に構造をサンプリングできる理由の 1 つは, 次サイクルの初期構造としてある特定の方向に起こるまれなゆらぎを選択することによって, 機能に関係している大規模構造変化をより起こりやすくしていることにある. PaCS-MD は当初考えていたように終構造が明確である必要性はなく, 探索の方向性を示すよいパラメーターをユーザーが指定できればよい (例えば主成分分析の第 1 主成分やドメイン間距離など). これに伴い, 現在, PaCS-MD を 1 構造からの蛋白質構造変化予測法へ拡張する研究も進めている. また, 探索に指定した方向以外には特にサンプリング効率が向上するわけではないので, 多数回シミュレーションを行って, 経路を多数生成することが望ましい. 他のサンプリング法と比較した優位性の 1 つとしては, 実際に実行するのが通常の MD 計算であるので, ユーザーが使いなれたシミュレーションソフトですぐに実行できる点があげられる. また, サイクルの前後である程度探索範囲が重なっていればよく, トラジェクトリの前後に特に厳密な関係が成り立っている必要がない. 従って, 超並列計算中にしばしば起こりうる, 計算機の一部が計算途中でダウンしたような場合にも計算を継続し, 生成できたトラジェクトリのみから計算を継続すればよい. つまり超並列計算に適した計算法であるといえる.

#### 5. 今後の計画・展望

本年度は, 蛋白質構造変化抽出法を確立したので, 今後は様々なシステム (基質結合も含む) に適用し, 機能解析を進めていく予定である. とくに, PaCS-MD から得られる構造変化経路のリファインメントをストリング法やパスサンプリングなどを用いて行い, リアクティブトラジェクトリの物理的解釈を考察したい. また従来のアンブレラサンプリングなどと組み合わせた自由エネルギー地形計算を行い, 動的静的両面からの機能解析を進めていきたいと考えている.

平成 25 年度 RICC 利用研究成果リスト

【論文発表】

1. Parallel cascade selection molecular dynamics (PaCS-MD) to generate conformational transition pathway, Ryuhei Harada and Akio Kitao, Journal of Chemical Physics, **139**, 035103, 2013.

【国際会議におけるポスター発表】

1. Fluctuation Flooding Method (FFM) for enhancing conformational sampling of proteins, Ryuhei Harada, Yu Takano, and Yasuteru Shigeta, 3rd International Conference on Molecular Simulation, Kobe, November 18-20, 2013
2. Fluctuation Flooding Method (FFM) for enhancing conformational sampling of proteins, Ryuhei Harada, Yu Takano, and Yasuteru Shigeta, Biophysical Society 58<sup>th</sup> Annual Meeting, San Francisco, California, February 15-19, 2014.