

課題名 (タイトル) :

## タンパク質・核酸など生体高分子の分子シミュレーション

利用者氏名 : ○木寺 詔紀, 寺田 透, 福田 育夫, 藤崎 弘士,  
今井 隆志, 古田 忠臣, 森次 圭, 松永 康佑

所属 : 社会知創成事業 次世代計算科学研究開発プログラム・分子スケール研究開発チーム

## 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

本研究チームでは、次世代生命体統合シミュレーション研究開発プログラムの分子スケール研究の一環として、生体分子（タンパク質等）のシミュレーション法とそのソフトウェアの開発研究、特に、全原子シミュレーション法と疎視化モデルとの新規の連成の方法論の開発を行っている。この研究の目的は以下の 2 点である：

- 1) 次世代スーパーコンピュータの全計算機資源を用いて高効率で計算することができる
- 2) それによってこれまでの分子シミュレーションの方法ではできなかったレベルの計算をすることができる

生命活動をタンパク質や核酸などの生体分子のレベルからシミュレーションによって解こうという分野における問題は、その巨大な系の大きさと生命現象の時間スケールの大きさである。その大きさのために、全原子シミュレーション法には巨大な計算機資源を用いても多くの場合、生命現象の解明が可能な系の大きさと計算時間の長さを実現するシミュレーションは不可能である。そこで不可避免的に疎視化モデルの利用が求められるが、そこには精度の制約が生まれる。従って、その両者の利点を併せ持つ連成計算（全原子シミュレーション法の精度と疎視化モデルの効率）が必要となる。また、数十万コアという並列計算を実現するためには、不可避免的に弱連成のアルゴリズムであることが要請される。これらを可能とする新規アルゴリズムとして、経路探索法（On-the-fly ストリング法）と MultiScale Enhanced Sampling (MSES 法)の研究開発を行った。On-the-fly ストリング法では、タンパク質の構造変化パス上に配置し、それらマルチコピーを弱連成で並列に計算することで、最適経路をサンプルする。MSES 法では、全原子モデルの運動空間を連成

した粗視化モデルでドライブすることにより全原子モデル巨大系の高速なサンプリングを実現する。

これらふたつの方法は、系が巨大タンパク質系であればそれだけで多くのコアによる計算を必要とし、さらにマルチコピーを必要とするところから、それらの掛け算の巨大な計算機資源を用いる必要がある。また、コピー間は弱連成であるところから次世代スーパーコンピュータの全計算機資源を用いて高効率で計算し得る可能性を持つ。これらの方法によってこれまでに全原子モデルシミュレーションでは不可能であった、長時間現象を再現することが可能となった。また、疎視化モデルでは不可能であった、全原子モデルの精度を与えることができる。

## 2. 具体的な利用内容、計算方法

松永、森次が開発した新規アルゴリズム（On-the-fly ストリング法、MSES 法）については、寺田が開発を統括しているマルチコピー・マルチスケール分子動力学シミュレーションソフトウェアに実装した。マルチコピーシミュレーションでは、異なるパラメータを与えた数十の系のコピー（レプリカ）を発生させ、それらの間の相互作用を考慮しながら並行してシミュレーションを行う。各コピーについて数十のコア、合計数百のコアを用いた並列計算を flat MPI、または OpenMP と組み合わせたハイブリッド並列により実行した。本年度は下記のようなアプリケーションに適用し、これらの方法の妥当性と物理化学的意味づけを評価した。

網羅的シミュレーションにおいては、分子動力学シミュレーションソフトウェア AMBER を用いた。計算対象に含まれる低分子化合物（リガンド）の力場パラメータの計算には、量子化学計算ソフトウェア Gaussian および GAMESS を適用した。計算対象のタンパク質系の周囲に十分な数の水分子を配置したのち、定温・定圧で 50 ns の分子動力学シミュレーションを行った。

今井が実施した 3D RISM 計算では、自然科学研究機構分子科学研究所で開発されたプログラムを用いた。本計算では、生体高分子を含む空間を細かいメッシュに切ることによって溶媒分子の分布を表現するため、大規模容量メモリ計算機により計算を実行した。

### 3. 結果

#### (1) マルチコピー・マルチスケール分子動力学シミュレーションソフトウェアの開発

次世代計算科学研究開発プログラム分子スケール研究開発チームでは、次世代スーパーコンピュータの計算能力を最大限に活用して、生体高分子の立体構造形成・機能発現メカニズムの解明を高精度かつ効率的に行うことのできる、新規アルゴリズムの開発を行っている。寺田、森次、松永は、これら新規アルゴリズムを実装するためのプラットフォームとなる、マルチコピー・マルチスケール分子動力学シミュレーションソフトウェアの開発を行っている。本ソフトウェアは MPI を用いて並列化されていたが、今年度は、並列計算効率の更なる向上を目指して、OpenMP と組み合わせたハイブリッド並列化を行った。本ソフトウェアでは、各コピー内の相互作用の計算を分割することにより、並列化を行っているため、相互作用の計算が完了するごとに、すべてのプロセスが計算結果を共有する必要がある。ハイブリッド並列化により、MPI のプロセス数を 1 プロセス/コアから、1 プロセス/ノードに減らすことができる (RICC の場合プロセス数は 1/8 になる) ため、計算結果の共有にかかる時間を大幅に削減することができた。RICC では、スレッド生成・消滅にかかるオーバーヘッドが大きいため、計算速度全体の向上には至っていないが、次世代スーパーコンピュータではオーバーヘッドが小さくなることが期待されており、通信時間の大幅な減少は、計算速度や並列計算性能の向上につながると期待される。

#### (2) スtring法の adenylylate kinase への応用

松永は多数のレプリカをタンパク質の構造変化パス上に配置し、それらレプリカを並列に計算することで構造変化パス空間をサンプルする手法 (string法) の開発・実装を行い、adenylylate kinase (溶媒を含めて 65,000 原子) という構造変化と機能が密接に関わるタンパク質系へ応用した。まず、string法による構

造変化パス探索を行い、得られたパス上のアンブレラサンプリングによりパス上の自由エネルギー地形を評価した。その結果、自由エネルギーの低い物理的に尤もな構造変化パスを捉えることに成功した (図 1)。また遷移状態の厳密な評価のための committer test を行った。これは、遷移状態候補周りを初期条件として多数の分子動力学計算を行い、トラジェクトリが 2 つの状態に別れるか否かをチェックするテストである。その結果、自由エネルギー地形上の凸部分で確かに軌道が分かれることが確かめられた。詳細な構造解析の結果、この原因は adenylylate kinase のヒンジ領域の原子とリガンドの協奏的な相互作用によることが分かった (図 2)。

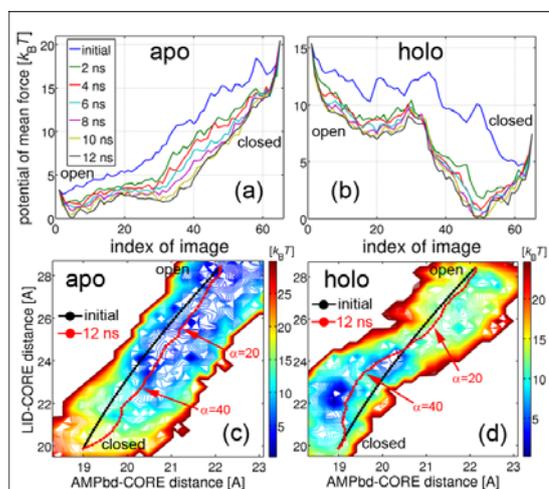


図 1: adenylylate kinase 構造変化パスの自由エネルギー地形

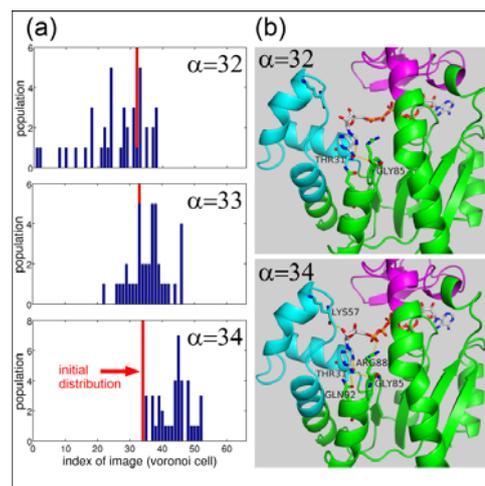


図 2: adenylylate kinase 構造変化パス上の committer probability と遷移状態構造。

#### (3) MSSES 法による蛋白質相互作用の解析

森次は天然変性タンパク質である sortase に MSSES 法を適用し、天然変性タンパク質の order-disorder 転移過程について調べた。

天然変性タンパク質は本質的に構造が決まっていな  
い柔軟に動きうる部分 (disorder 領域) を持っており、  
その柔軟さがターゲットの認識や結合といった分子機  
能に関わると考えられている。本研究では、生理学的  
環境下における sortase の disorder ループ構造アンサ  
ンプルを原子レベルの解像度で計算した。シグナルペプ  
チドあるなし、カルシウムイオンあるなしの 4 つの状  
態での構造アンサンプルをシミュレーションにより求  
め、4 状態について自由エネルギー地形を計算した。そ  
の結果、disorder ループがフリーの状態ではきわめて柔  
軟に動きうること、また、カルシウムイオンの結合状  
態では disorder ループの柔軟さが失われることにより、  
ペプチドとの結合が (エントロピー的に) より有利に  
なることを明らかにした (図 3)。

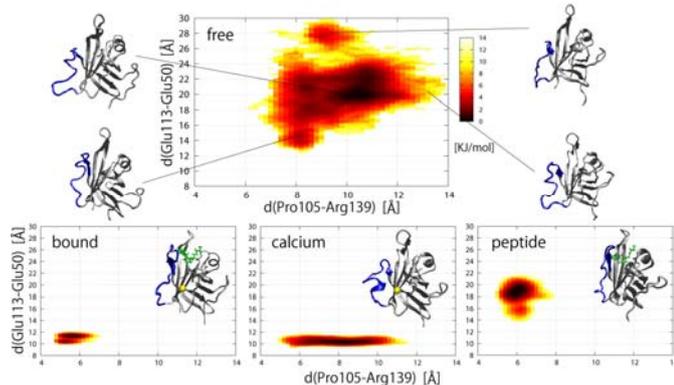


図 3: 天然変性タンパク質 sortase の disorder ループ自  
由エネルギー地形。

#### (4) 拡張アンサンプル Onsager-Machlup 法の開発・ 実装

藤崎は、パスサンプリング法の一つである  
Onsager-Machlup 法と、本研究チームの森次によって開  
発された Multi-scale essential sampling (MSES) を組み合  
わせることで、パスサンプリングをより強力に行う方  
法論を開発した。それを現在はモデルポリマーに応用  
しており、パス空間をより広くサンプリングできるこ  
とを検証中である (図 4)。

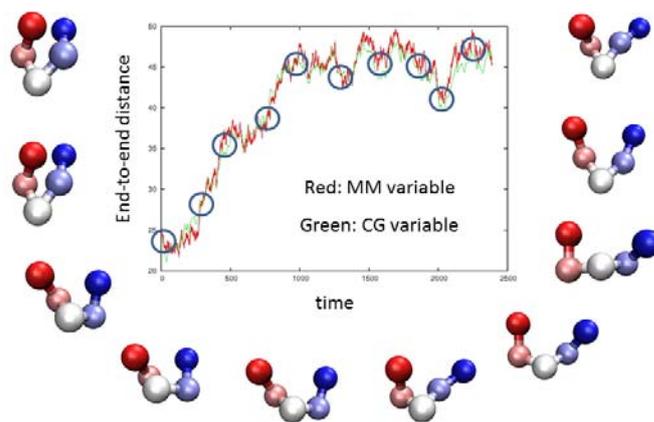


図 4: 拡張アンサンプル Onsager-Machlup 法により得ら  
れたモデルポリマーの構造変化パス。

#### (5) 網羅的シミュレーションシステムの研究開発

木寺、寺田、古田は、生体分子の運動に関する統一  
的な理解につながる知識を抽出することを目指して、  
多数の生体高分子に対する網羅的な分子動力学シミュ  
レーションを行うことのできる分子シミュレーション  
システムの研究開発を行っている。網羅的タンパク質  
シミュレーションシステムは、大きく 3 つの部分 (デ  
ータベース部、モデリング部、シミュレーション部)  
からなるが、モデリング部において、本年度は新しく、  
リガンド分子 (1 分子) がある場合の力場作成の実装を  
行った。本システムのリガンド結合タンパク質への適  
用 (50 ns の分子動力学シミュレーションの実行) を行  
い、数種のターゲットにおいて、リガンドを含まない  
Closed 構造を初期構造とするシミュレーションにおい  
て、数 10 ns 程度で構造変化 (Opening) が再現される  
という結果を得た (図 5)。比較として、(Closed 構  
造を初期構造とする) リガンドを含む系でのシミュレ  
ーションでは、大きな構造変化は観測されず安定である  
ことも確かめられた。

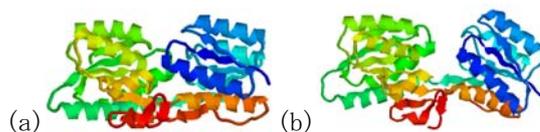


図 5: RBP の (a) closed 構造 (b) open 構造

#### (6) Zero-dipole 法の開発

福田らは既に、Wolf らが提案した静電相互作用エネ  
ルギーの計算法を改良し、正確な全エネルギー計算を  
行い、かつ、正しく導出された力の関数を用いてコン  
システントな MD 計算を実現する手法

(Force-switching [FSw-]Wolf 法)を開発した。しかしながら、より小さいカットオフ長を用いて、より精度の良い計算を行うためには、パラメータ調整などでは限界があり、新しい原理の構築が求められていた。そこで、点電荷集合の中性条件を拡張して、1次モーメントの中性条件を考え、これを二体相互作用で扱う新規な方法を開発した (Zero-dipole [ZD] 法)。本手法によれば、FSw-Wolf 法と同様に、全エネルギーがペア相互作用項と定数付加項のみで表されるので、インプリメントが簡単であり、高性能並列計算機においても十分な低コスト化と高スケーラビリティの達成が可能である。さらに、ZD 法が、全く異なる原理から構成された別のエネルギー計算法の拡張になっている事も発見した。強結合イオン系の場合に、本手法の精度が、FSw-Wolf 法に比べて、実用的なより短いカットオフ距離にて、大きく向上することを示した (図 6)。

$r_c$	9	10	11	12	13	14
$\alpha = 0.05$	8.749	49.039	5.124	29.109	9.446	15.593
	0.629	0.016	0.296	0.062	0.150	0.057
$\alpha = 0.1$	4.554	17.846	1.161	7.513	1.781	2.707
	0.569	0.029	0.232	0.042	0.106	0.022

図 6 : エネルギーエラー誤差 (%). パラメータ  $\alpha$  の各値について、上段が従来の FSw-Wolf 法, 下段が ZD 法.

#### (7) 3次元 RISM リガンドマッピング法を用いた応用研究

今井は、3次元 RISM リガンドマッピング法を用いて、多剤排出トランスポーター AcrB に対する一連の薬物フラグメントの化学官能性のマッピングを行った。その結果、AcrB の輸送ドメイン内空洞に少数の“多官能性”のリガンド結合サイトが存在することを発見した。それらを空間的に結ぶことにより、AcrB の薬剤排出経路のモデルを提案した。また、薬剤排出経路におけるリガンド-タンパク質の弱いフラストレートした相互作用が薬剤のスムーズな輸送のメカニズムであることを示唆した。その経路内における構造変化に誘導された拡散が排出の主なドライビングフォースと考えられる。

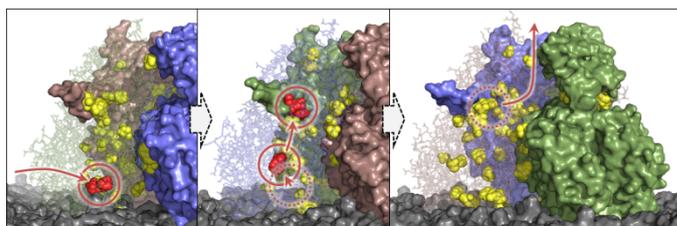


図 7 : AcrB の薬剤排出経路モデル

#### 4. 今後の計画・展望 (これまで利用した状況や、継続して利用する際に行う具体的な内容)

マルチコピー・マルチスケール分子動力学シミュレーションソフトウェアの開発においては、新規アルゴリズムの開発において追加される機能の追加、入出力などのクラスライブラリ高度化、また、超並列マシンである次世代スーパーコンピュータへの対応として OpenMP と組み合わせたハイブリッド並列化によるコードの高速化をさらに進めていく。

MSES 法の継続課題として、次世代生命体統合シミュレーション研究開発プログラムの中間評価において実施を約束したタンパク質間相互作用の分子シミュレーションを引き続き行う。ターゲットとして、タンパク質複合体のテストケースとして数多くの実験・計算が行われている barnase-barster 複合体を用いる。MSES 法により複合体近傍の構造アンサンブルを計算することで、特に複合体形成における脱水と過程や側鎖の緩和過程について明らかにする。また、新たなアプリケーションとして、次世代生命体統合シミュレーション研究開発プログラムの分子スケール研究のターゲットの一つである脂肪酸ベータ酸化マルチエンザイム複合体の構造変化パス周辺の構造アンサンブルを計算する。

Onsager-Machlup 法については、平成 21 年度に開発したパスサンプリングの方法論 (温度レプリカ交換法 + Onsager-Machlup 法) と平成 22 年度に開発した方法論 (Multi-scale essential sampling + OM 法) を現実の分子系に適用する。まずはわれわれが開発中の Platypus-MM/CG を使って、アラニンジペプチドの構造変化を動的に特徴づけることを行い、その後、シニョリンのような MSES が試されている小さなタンパク系のパスサンプリングを行う。これらの手法は、スト

## 平成 22 年度 RICC 利用報告書

リング法と同様に、離散化された経路（各レプリカ）に対して、超並列計算が容易に実装できるので、将来のペタコンで使えるアルゴリズムとして有望である。

ZD 法については、平成 22 年度に手法の開発を行ったが、平成 23 年度では ZD 法の実装に取り組み、溶液系、生体分子系などでの実証計算を行う。1-2 相互作用などの取り扱いに留意する必要がある。また、実証結果を踏まえ、さらなる高効率計算を目的に、ZD 法をさらに拡張した zero multiple moment 法(仮称)の開発も念頭に置く。

平成 22 年度 RICC 利用研究成果リスト

【論文、学会報告・雑誌などの論文発表】

1. Kei Moritsugu, T. Terada and Akinori Kidera, “Scalable Free Energy Calculation of Proteins via Multiscale Essential Sampling”, *J. Chem. Phys.* 133, 224105 (2010).

【国際会議などの予稿集、proceeding】

該当なし

【国際会議、学会などでの口頭発表】

1. 森次 圭

「MSES シミュレーションによる Sortase A 不規則領域の構造空間探索」

次世代生命体統合PJ 分子スケール研究会 2010 前期、 京都、2010 年 8 月

2. 松永 康佑、寺田 透、森次 圭、古田 忠臣、木寺 詔紀

「マルチコピーシミュレーションによるタンパク質構造サンプリング」

2010 年度理研シンポジウム ペタフロップス時代のセンターシステム、和光、2011 年 2 月

RICC カutting エッジアワード受賞

3. 木寺 詔紀

"Supercomputing of protein systems using innovative algorithms for multiscale simulations"

バイオスーパーコンピューティング研究会 講習会、大阪、2011 年 3 月

【その他】

該当なし