

MSES 法によるタンパク質複合体シミュレーション
Multiscale enhanced sampling simulation of protein complexes

森次圭
Kei moritsugu

理化学研究所・次世代計算科学研究開発プログラム
CSRP, RIKEN

新規構造サンプリング手法である MSES 法を barnase-barster 複合体と脂肪酸ベータ酸化マルチエンザイム複合体し、タンパク質複合体形成過程のシミュレーションを行った。MSES 法により barnase-barster 複合体近傍の構造アンサンブルを計算し、複合体形成における脱水過程や側鎖の緩和過程、また、複合体形成における水素結合の役割を明らかにした。脂肪酸ベータ酸化マルチエンザイム複合体についても MSES 法を適用し、マルチスケールを用いた新規手法が巨大タンパク質にも適用可能であることを示した。

A novel conformational sampling simulation technique, MSES method, was applied to barnase-barster complex and multienzyme complex to simulate the protein complex formations. All-atom conformations of the barnase-barster complex sampled by the MSES revealed the desolvation and side-chain arrangement mechanisms as well as the role of hydrogen bonds formed during the complex formation. The conformational sampling of the multienzyme complex demonstrated the applicability of the MSES to biomacromolecules.

Keywords: protein interaction, all-atom conformational sampling, MSES, barnase-barster, multienzyme complex

背景と目的

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：

- ・次世代スーパーコンピュータ「京」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

- ・それによって従来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ため、新規アルゴリズムである MultiScale Enhanced Sampling (MSES)法を開発した。

MSES 法は、全原子モデルと低自由度の疎視化モデルによる連成シミュレーションである。疎視化モデルのポテンシャルが規定する運動空間に全原子モデルをドライブし、全原子モデルと疎視化モデルとを接続するバネ強度を 0 に外挿することで、全原子モデルの空間での分布関数を得ることができる。バネ強度の 0 への外挿は、バネ強度を変数としたハミルトニアンのレプリカ交換法によって行う。従って、MSES 法はバネ強度の異なる多数のコピーを用い

(様式第 20) 成果報告書

た弱連成のシミュレーションであり、高度の並列計算が可能である。レプリカ交換が疎視化モデルの自由度により決まることから、通常の温度レプリカ交換法と異なり全原子モデルの自由度の制限なく巨大系のサンプリングが実行可能となる画期的な方法である。昨年度までの研究において MSES 法の実装は完了し、テストケースとして小タンパク質のフォールディング過程に適用することで方法論の妥当性を示した。また、従来の拡張アンサンブル法では難しかった、天然変性タンパク質である sortase の（溶媒水を取り入れた）生理学的環境下での構造アンサンブル計算を実現し、sortase 変性領域の折れたたみ過程やシグナルペプチド結合におけるカルシウムイオンの働きといった生物学的な知見を得ることができた。

概要

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、タンパク質の全原子モデルと粗視化モデルを連成した新規アルゴリズムである Multiscale enhanced sampling (MSES) 法の開発を行っている。昨年度からの継続課題である本利用課題では、新たな応用研究として、barnase-barster 複合体と脂肪酸ベータ酸化マルチエンザイム複合体の複合体形成過程シミュレーションを行った。barnase-barster 複合体はタンパク質複合体のモデルタンパク質として数多くの実験・計算が行われており、MSES 法により複合体近傍の構造アンサンブルを計算し、特に複合体形成における脱水和過程や側鎖の緩和過程を調べた。脂肪酸ベータ酸化マルチエンザイム複合体は脂肪酸の代謝を担う巨大タンパク質であり、ダイナミックな立体構造変化とカップルして結合する脂肪酸がいかにして複数の活性部位に効率的に運ばれるかを解明するため、構造変化パス近傍での全原子モデル構造サンプリングを目指した。

結果および考察

(1) タンパク質複合体形成シミュレーション

タンパク質複合体のモデルとして数多くの実験・計算

をとおして研究がすすんでいる barnase-barster 複合体に MSES 法を適用した。2つのタンパク質合わせて 194 残基、系としては 3 万原子程度であるが、タンパク質側鎖を含めた全原子モデルにより溶媒水を含めた生理学的環境下でのタンパク質間相互作用過程をシミュレートすることは MSES 法によって初めて可能になった試みである。

粗視化モデル(CG)の自由度としてはアミノ酸の $C\alpha$ 原子を考え、粗視化モデル力場としては複合体構造を安定状態とするような弾性ネットワークモデルを用いた。全原子モデル(MM)と CG のばね強度を変えたテスト計算により、12 個のレプリカ計算により十分な構造サンプリングができることを確認した(図 1)。1 レプリカに 16 コアを用い、 $12 \times 16 = 192$ コアでのシミュレーションを 100 ns にかけて実行した。

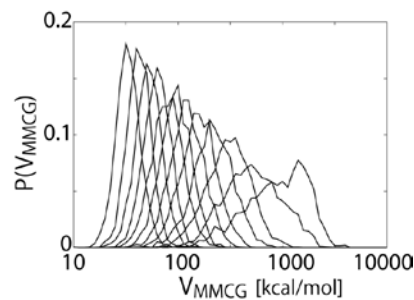


図 1: barnase-barster 複合体ハミルトニアンレプリカ交換 MSES 法での各レプリカの $P(V_{MMCG})$ 。

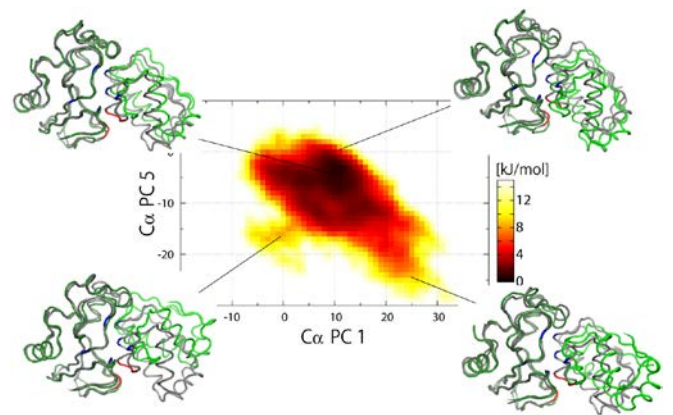


図 2: barnase-barster 複合体ハミルトニアンレプリカ交換 MSES 法で得られた自由エネルギー地形。

得られた構造アンサンブルから自由エネルギー地形を計算した結果(図 2)、2つのタンパク質が離れた構造から複合体近傍に向けて特定のパスを経ることなく downhill なエネルギー面を滑り降りるように近づく複合体形成過程が確かめられた (Approaching)。

Approaching でのタンパク質間の接触原子ペア数と水和水の数を計算したところ、複合体に近づくにつれて原子間相互作用が形成され、それに従い水和水が離れる脱水和が起こることがわかった。

しかしながら、MSES 法で得られる複合体近傍構造と X 線の複合体構造とは構造に小さな差異がみられた。その違いをより明らかにするため、X 線の複合体構造を初期構造としてシミュレーションを行い、両方の構造アンサンブルを比較した。タンパク質間の水素結合に注目した結果、Approaching による複合体近傍で形成される水素結合とそこからさらなる側鎖のパッキングや脱水和 (Arrangement) を経て形成される水素結合に分類できることを見出し、Approaching と Arrangement 過程において水素結合が形成される順序を決定した。また、Arrangement 過程での遷移状態における水素結合状態を特定した。この水素結合には比較的小数のアミノ酸残基が関わっており、これらの残基は 1 残基置換実験で $\Delta\Delta G$ が大きい、つまり、複合体形成における相互作用形成に必須であることが実験で示されている。

(2) 脂肪酸ベータ酸化マルチエンザイム複合体の構造空間探索

脂肪酸ベータ酸化マルチエンザイム複合体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におけるターゲット分子の一つで、数万原子の 8 つのドメインからなる巨大分子である。ダイナミックな構造変化とカップルして結合する脂肪酸を複数の活性部位に移動させることにより効率的に脂肪酸の代謝を行うと考えられており、その構造変化パスとパスに沿った脂肪酸の結合様式を見ることが研究の目的である。

X 線結晶解析により解かれた上下 8 つのドメインが対称 (Form 1) と非対称 (Form 2) の二つの構造間の構造変化 (図 3) をシミュレートするため、二構造を安定状態としてその間をなめらかに遷移するような粗視化モデル (Mixed elastic network model) 力場を構築した。全原子モデルを粗視化モデルによりドライブするには MM と CG 間の拘束が必要であるが、拘束エネルギーが大きすぎるとレプリカ交換 MSES 法において多くのレプリカが必要になり、サンプリングの収束に時間がかかる。そのため、MM/CG 拘束は α サブユニットと β サブユニットの 6 つの各ドメイン間の $C\alpha$ 原子ペアのうちで原子間距

離が小さなもののみにかけた。全原子モデル (MM) と CG のばね強度を変えたテスト計算により、28 個のレプリカ計算により十分な構造サンプリングができることを確認した (図 4)。今後、MSES 法による構造サンプリングを十分な時間スケールで行い、Form 1 か Form 2 への構造変化パスを調べる予定である。

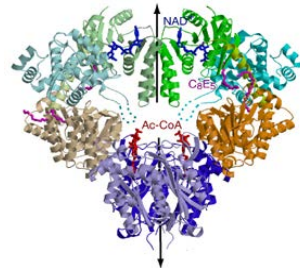


図 3: 脂肪酸ベータ酸化マルチエンザイム複合体の立体構造。対称構造 (Form 1) と非対称構造 (Form 2)。

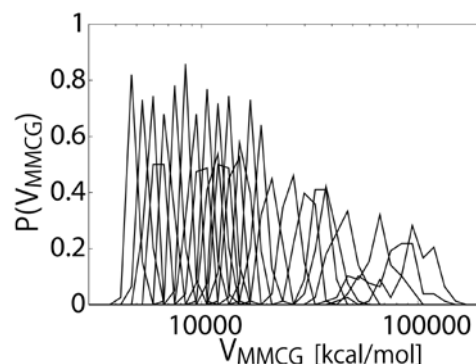


図 4: 脂肪酸ベータ酸化マルチエンザイム複合体ハミルトニアンレプリカ交換 MSES 法での各レプリカの $P(V_{MMCG})$ 。

まとめ、今後の課題

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一環としてマルチコピー・マルチスケール分子動力学シミュレーションのためのクラスライブラリ ($\mu 2lib$) の開発を行い、そのアプリケーションとして MSES 法をタンパク質複合体と巨大タンパク質分子に適用した。系の自由度に応じて構造探索に必要なシミュレーション時間は指数関数的に増大するが、新規アルゴリズムとマルチコピー・マルチスケール手法の組み合わせにより初めて可能になったシミュレーション成果であるといえる。

barnase-barster 複合体の形成過程については、これまでの研究では粗視化シミュレーションや 2 つのタンパク質の重心間距離のみを自由度にした自由エネルギー計算が行われてきた。本研究でははじめて、複合体構造まわりの広範囲の全原子サンプリングを溶媒水を含めて行い、側鎖のパッキングや脱水和、水素結合形成

(様式第 20) 成果報告書

といった原子情報を用いた詳細な解析により複合体形成の原子レベルでの描像を明らかにした。

MSES 法のアプリケーションとしては、引き続き脂肪酸ベータ酸化マルチエンザイム複合体の構造サンプリングを行い、2 構造での平衡 MD の結果を参照しながらその構造変化パスとパスに沿った脂肪酸の結合様式を観測し、脂肪酸を複数の活性部位に移動させることにより効率的に脂肪酸を代謝する過程を原子レベルで明らかにする。