

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: ○森次 圭

所属: 情報基盤センター・計算工学応用開発ユニット

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

本研究チームは平成 24 年度までの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・分子スケール研究において、タンパク質を中心とした生体分子のシミュレーション法とそのソフトウェアの研究開発 (コード名: $\mu^2\text{lib}$, <http://www.mu2lib.org> のサイトで公開) を行った。特に、全原子シミュレーションと疎視化モデルシミュレーションとの連成手法を新規開発することを目指した。プロジェクトでの研究目的は以下の 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、一つの応用研究として、利用者らは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、 $\mu^2\text{lib}$ への MSES 法の実装および高度化は完了している。また応用研究として、ミニタンパク質シニョリンのフォールディング過程だけではなく、天然変性タンパク質 sortase の disorder-to-order 転移の構造探索や barnase-barster 複合体、HPr-EIN 複合体の自由エネルギー地形計算による複合体形成過程の理解、また、グルタミン結合タンパク質へのリガンド結合過程の解析といった、従来の拡張アンサンブル法では難しかった大規模系 (溶媒を陽に含んだタンパク質系という意味で) への適用も進めてきた。

2. 具体的な利用内容、計算方法

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 ($\mu^2\text{lib}$) の開発を継続しており、今年度は断熱分離の適用により MSES 法を拡張した手法の開発を進め $\mu^2\text{lib}$ に実装した。 $\mu^2\text{lib}$ ではマルチコピーシミュレーションを行っており、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実行する。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 OpenMP と組み合わせたハイブリッド並列により行なう。今年度は下記のような応用研究を行い、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。

3. 結果

(1) 断熱分離の近似を用いた MSES 拡張法の開発

MSES 法を巨大タンパク質のような広範な構造探索を必要とする系に適用するにあたり、全原子モデル(MM)の運動を効率的に高速化できるかという問題が生じる。今年度の研究では、断熱分離 (adiabatic separation) の考え方に基づき粗視化モデルを運動エネルギー的に全原子モデルから切り離すことで、粗視化モデルに高い運動エネルギーを与えることで、全原子モデルのサンプリング効率を最大化する近似手法を導出した。

この新規拡張法を溶媒中シニョリンのフォールディング過程に適用し、手法の妥当性を示した。水素結合距離に関する 2 次元エネルギーランドスケープや折れ畳み構造からの RMSD 分布を見るとほぼ同じアンサンブルを生成することから、この近似手法でも十分な精度で正しい構造分布が得られることを示した (図 1)。また、CG 質量が小さいような近似適用範囲外の場合には、得られる分布も正しくないことを実証した。

以上の結果について論文にまとめ投稿した。この拡張法では、Gibbs サンプリングに基づき CG シミュレーションを MM によらず単独で実行するため、従来法では時間を要した CG パラメタ決定が容易になる点においても、大規模系への応用に適していると考えられる。以下 2 つの応用研究では、この拡張法を適用して全原子構造サンプリングを実行した。

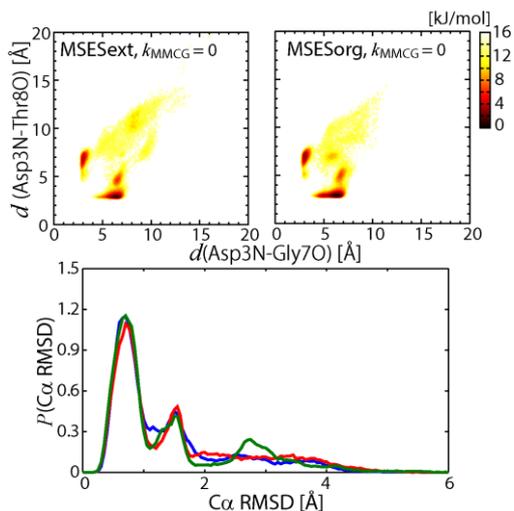


図 1 : 拡張法 (上図左) と従来法 (上図右) でのシニョリンフォールディングシミュレーションの自由エネルギーランドスケープ。下図は折れ畳み構造からの RMSD 分布。従来法 (青) と拡張法 (赤、 $m_{CG} = 10000$)。緑は拡張法で $m_{CG} = 1000$ の場合。

(3) 上皮成長因子受容体(EGFR)キナーゼの全原子構造サンプリング

プロテインキナーゼは基質タンパク質をリン酸化しシグナル伝達や代謝の調節因子として働く。その作用機構は、キナーゼの活性型・不活性型間の構造変化をキナーゼ自己リン酸化や基質結合により高度に制御する分子機序に基づいている。そのようなダイナミックな過程を理解するため、本研究では実験・計算などで研究が進んでいる上皮成長因子受容体(EGFR)キナーゼを研究対象として取り上げ、ATP 結合・非結合の 2 状態について、活性型・不活性型間の広範な構造変化を MSES 法によりシミュレートした。

全原子モデルは、結晶構造から溶媒水を陽に含んだシミュレーション系を構築し、粗視化モデルとしてはキナーゼタンパク質の $C\alpha$ 原子を選択した。全原子力場としては、Amber ff96SBildn を用いた。粗視化力場としては、結晶構造が解かれている EGFR の活性型・不活性型構造 (PDB: 2GS2, および 4HJO) をつないだ double-well 弾性ネットワークモデルを適用した。ハミルトニアンレプリカ交換 MSES は 16 個のレプリカを用い、ATP 結合では 100 ns、非結合では 200 ns のプロダクトランを実行した。

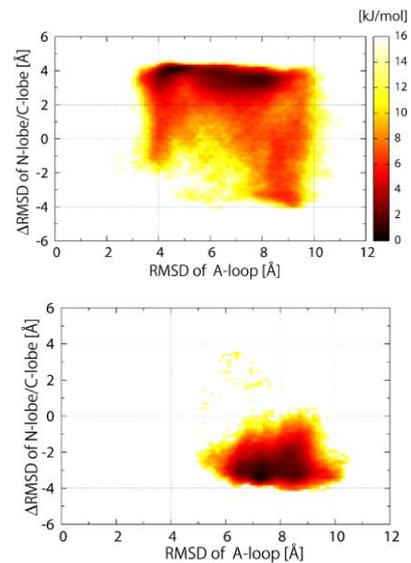


図 2 : activation loop と lobe 間運動を自由度にした 2 次元自由エネルギーランドスケープ。ATP 非結合 (左) と結合 (右) 状態。

EGFR の活性型・不活性型結晶構造では、2 つの N-/C-lobe 間の相対配置とともに、活性に重要な役割があると考えられる activation loop と αC ヘリックスが大きく変化する。このため、それらの部位の運動についての自由エネルギーランドスケープを計算した。まず

activation loop と lobe 間運動について見たところ、ATP 結合状態では activation loop が開いた構造で安定し基質との結合が可能であるのに対して非結合状態では activation loop が閉じた構造が安定であるが大きく動きうること、また、それに応じて、ATP 結合にともない lobe 間運動が固定化されるが ATP 非結合状態では揺らぎが大きいことが分かった (図 2)。また、 α C ヘリックスに注目すると、activation loop がほどけることで始めて活性型の倒れた構造になることができることが示された (図 3)。また、キナーゼの活性型構造において観測されるリシンと α C ヘリックス上のグルタミン酸の結合や R-spine と呼ばれるリン酸転移反応部位を周囲の溶媒環境から切り離す役割を持つ疎水性残基コンタクトに注目した解析を行った結果、 α C ヘリックスがねじれてリシン-グルタミン酸結合が生成されることが R-spine 形成の driving force であり、活性・不活性構造間における遷移状態であることが示唆された。

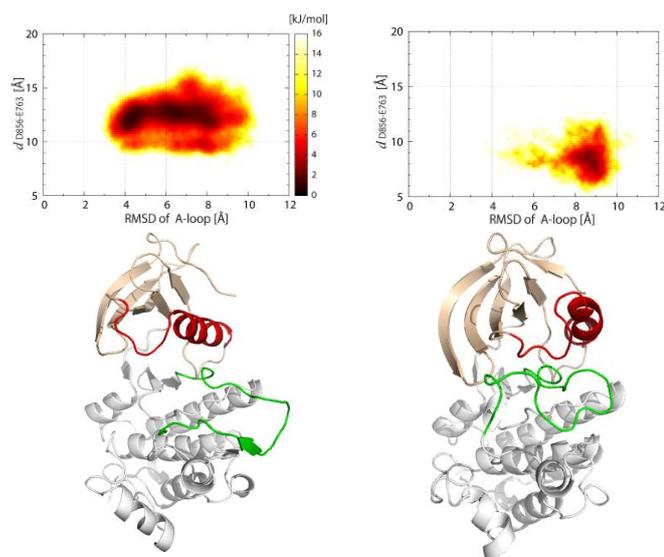


図 3 : activation loop と α C ヘリックスを自由度にした 2 次元自由エネルギーランドスケープ。lobe 間相対配置が不活性 (上図左) と活性 (上図右) 状態のとき。下図はそれらの代表構造。

(3) イオンチャネル共役型グルタミン酸受容体 (iGluR) のリガンド結合過程の解析

iGluR はグルタミン酸結合を介してイオンチャネルの開閉を制御しており、中枢神経系における記憶や学習等の脳の高次機能に重要な役割を果たしている。グルタミン酸結合・非結合の X 線結晶構造が解かれており、グルタミン酸の結合と受容体タンパク質の大きなドメイン運動がカップルしている。本研究では、昨年度に引き続き、MSES 法による iGluR グルタミン酸複

合体の全原子構造サンプリングを実行し、300 ns のプロダクトランを完遂した。

得られた構造アンサンブルから自由エネルギーランドスケープを計算した結果、MSES 法により複合体構造近傍だけではなく、グルタミン酸の結合解離と iGluR のドメイン運動両方の自由度について、広範な構造サンプリングが実現できた (図 4)。また、複合体構造を最安定とするファネル状のランドスケープであることを示した。この結果は、iGluR グルタミン酸複合体のリガンド選択性の大きさを表していると言える。

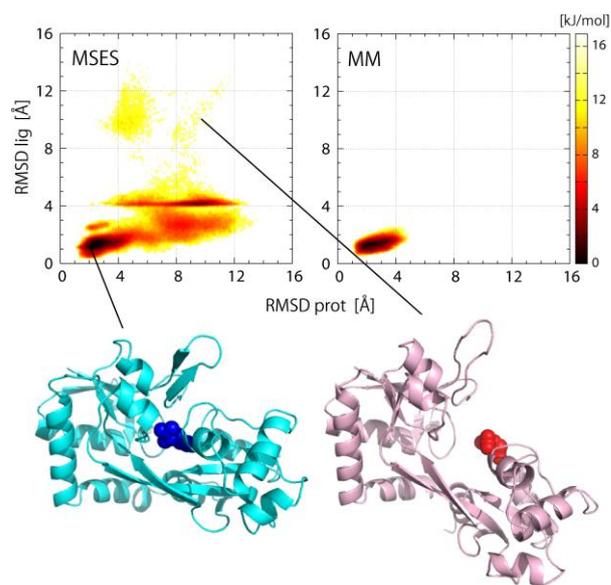


図 4 : iGluR グルタミン酸複合体の自由エネルギーランドスケープ。100 ns 全原子 MD からの結果を上図右に示した。下図は、複合体近傍とグルタミン酸解離/iGluR open 状態での代表構造。

4. まとめと今後の計画・展望

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 μ^2 lib の開発を進め、大規模系への応用に向けた MSES 法の新規拡張法を導出した。また、応用研究として、iGluR のグルタミン酸結合過程の解析、及びプロテインキナーゼ EGFR の活性化に関わる構造変化過程に適用した。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。2つの結果については、実験および創薬データベースとの関連性についても検討しつつ、結果

を論文にまとめていきたい。

来年度は、グアニンヌクレオチド交換因子 SOS と Ras との相互作用過程のシミュレーションを行う。MAPK シグナリングパスウェイでは、細胞質側へのシグナルの受容がグアニンヌクレオチド交換因子 SOS に伝えられることで、SOS が細胞膜上にリクルートされるが、水溶液中での (Ras 非結合) SOS の網羅的サンプリングシミュレーションを MSES 法により実行することで、SOS の自己阻害型から Ras 結合型に至る広範な領域で

のエネルギーランドスケープを計算し、そのような構造変化をシグナル伝達につなげる作用機序を理解する。また、グルコースへのリン酸化転移を行うグルコキナーゼの網羅的構造サンプリングを行い、グルコース濃度に対する cooperativity がドメイン運動によりどのように制御されるかを明らかにする。また、weighted ensemble 法の適用により、構造変化の時定数といった kinetic なパラメタの計算や遷移状態の同定も組み合わせて行う。

平成 28 年度 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, Tohru Terada and Akinori Kidera, "Multiscale enhanced sampling for protein systems: An extension via adiabatic separation", Chemical Physics Letters (2016) 661: 279-283.
2. Ikuo Fukuda and Kei Moritsugu, "Coupled Nosé-Hoover equations of motions without time scaling", Journal of Physics A (2016) 50: 015002.
3. Kei Moritsugu, Tohru Terada and Akinori Kidera, "Free-Energy Landscape of Protein-Ligand Interactions Coupled with Protein Structural Changes", Journal of Physical Chemistry B (2017) 121: 731-740.