

課題名 (タイトル) :

大規模生体分子系におけるマルチレゾリューション手法の開発

利用者氏名 :

○松永 康佑*
 岩橋—小林 千草*
 Jaewoon Jung**
 神谷 基司*
 田村 康一*

理研での所属研究室名 :

*計算科学研究機構 粒子系生物物理研究チーム

**杉田理論分子科学研究室

カルシウムイオンポンプの反応機構の解析 (担当:岩橋—小林)

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

近年、結晶構造解析などから、生体分子の詳細な構造情報の取得が可能となり、生体分子機能に対する理解が進んでいる。詳細な構造情報を用いた分子動力学法(Molecular Dynamics; MD)等の計算機シミュレーションもまたその一端を担っている。シミュレーションは、生体分子が機能するときの詳細な機能ダイナミクス情報を得られることが強みであり、分子機能を物理化学的に解明するために不可欠のアプローチとなってきた。我々のチームでは、京や RICC で高効率に動作する超並列 MD プログラム「GENESIS」を開発・公開してきた。本課題では、GENESIS の高速 MD エンジンを中心に、HOKUSAI において大規模生体分子系に対するマルチレゾリューション連成計算法を開発・応用する。GENESIS の高速 MD エンジンと、マルチレゾリューション計算を組み合わせることで、これまで計算不可能であった生体分子機能をシミュレートすることを目的とする。具体的には、以下の 5 つの研究を行った。

- 1) カルシウムイオンポンプの反応機構の解析 (担当:岩橋—小林)
- 2) 分子動力学計算によるヘム輸送体のホモロジーモデルの高精度化 (担当:田村)
- 3) Optimization of GENESIS MD software with multi-kernel (担当:Jung)

4) 新しいレプリカ交換法を用いたタンパク質構造探索 (担当:神谷)

5) Protein G フォールディングのデータ同化解析 (担当:松永)

2. 具体的な利用内容、計算方法

1) カルシウムイオンポンプの反応機構の解析

本年度は、反応経路解析手法を求める手法の開発を主に行った。String 法はイメージと呼ばれるシミュレーションシステムの複数のコピーを弱く相互作用させながらサンプリング計算を行うことで広い構造空間を探索し、最も可能性の高い反応経路を探索する方法である。更に string 法で得られた反応経路上で umbrella sampling を行うことで自由エネルギープロファイルを得る事ができる。これらの一連の方法は生体高分子の複雑な構造変化とカップルした反応経路の解析に使われている。しかし、イメージ数や拘束ポテンシャルの力の定数など様々なパラメータに結果に依存することが現在までの研究で分かってきた。効果的なプロトコルを作成すべく、これらのパラメータを変更し、その自由エネルギー経路がどの様に影響されるのかを解析を行った。カルシウムイオンポンプや大きなドメイン運動がみられる Adenylate Kinase を用いて解析を行った。モデルは粗視化モデル (DoME モデル) を用い、プログラムは GENESIS を用いた。

2) 分子動力学計算によるヘム輸送体のホモロジーモデルの高精度化

MD によって、MODELLER で生成したヘム輸送体のホモロジーモデルを高精度化した。具体的には、モデルの構築と、その後の全原子 MD による不安定部位の同定、のサイクルを複数回繰り返すことによって、脂質二重膜中での長時間のシミュレーションに耐える安定なモデルを生成した。タンパク質のパラメータは CHARMM36 を用いた。MD 計算は GENESIS を使用した。

3) Optimization of GENESIS MD software with multi-kernel

We developed GENESIS MD software according to Greatwave and new BigWaterfall machines. Because each machine has different hardware architecture, we need optimize GENESIS such that it has the maximum performance for all machines. To do this, we applied multi-kernel scheme where the kernel programs of most time-consuming non-bonded interactions are prepared according to each hardware architecture. In particular, we prepared the kernel program for the Big Waterfall machine.

4) 新しいレプリカ交換法を用いたタンパク質構造探索

レプリカ交換アンブレラサンプリング(REUS)法やレプリカ交換溶質焼きなまし(REST)法やその発展形である gREST 法を Trp-cage や tri-alanine 系に適用した。計算には GENESIS を用い、2-10 程度のレプリカを同時に実行させた。

5) Protein G フォールディングのデータ同化解析

Protein G フォールディングダイナミクスを 1 分子 FRET 計測データとデータ同化してモデリングするために構造サンプリングを行った(図 1)。多数の MD シミュレーションを独立にながして、網羅的に構造サンプリングを行った。構造空間を効率良くサンプリングするために、 Q 座標 (天然構造と同じ残基ペアを形成している割合)上で、シミュレーション同士が互いに異なる領域をとるようなサンプリング手法(Adaptive sampling)を用いた。

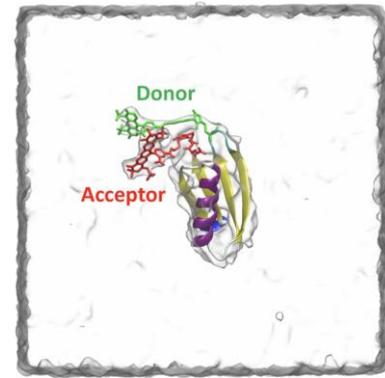


図 1 : Protein G の全原子モデル。FRET 計測に使う蛍光色素分子も explicit にいれている。

3. 結果

1) カルシウムイオンポンプの反応機構の解析

イメージ数や拘束ポテンシャルの力の定数を変更し、得られた経路や自由エネルギープロファイルの比較・解析を行った。主成分座標の拘束ポテンシャルを用いた際に、レプリカ交換 umbrella sampling 法にて広いサンプリングが可能となる力の定数を決める方法を構築した。(論文 2) 自由エネルギープロファイルの取得において効果が得られることも分かった。

2) 分子動力学計算によるヘム輸送体のホモロジーモデルの高精度化

MD 計算による不安定部位の同定に要した典型的なシミュレーション時間は <100 ns 程度であった。一例として、 α ヘリックスの崩壊を挙げる(図 2)。この部位は、テンプレート(ビタミン輸送体)とターゲット構造(ヘム輸送体)の間の配列相同性が著しく低く、そのために MODELLER によって無理なモデル生成が行われたと考えられる。この部位のモデルを改善するために、別の生物種の輸送体の対応するヘリックスを切り貼りすることで新しいモデルを生成した。この部位の不安定性はこの手続きによって解消された。全てのサイクルの後に 1,500 ns のシミュレーションを行い、構造的に脆い部分のない安定なモデルであることを確認した。

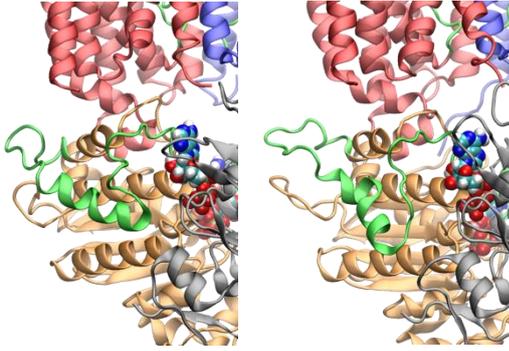


図 2: MODELLER が生成したモデル (左) と、MD 計算の途中で崩壊したヘリックス (右)。注目するヘリックスを緑で示した。

3) Optimization of GENESIS MD software with multi-kernel

Performance result of GENESIS MD software on Big Waterfall machine

Due to the optimization scheme considered in Big Waterfall machine, we could obtain very good performance result with good parallelization efficiency (Figure 3).

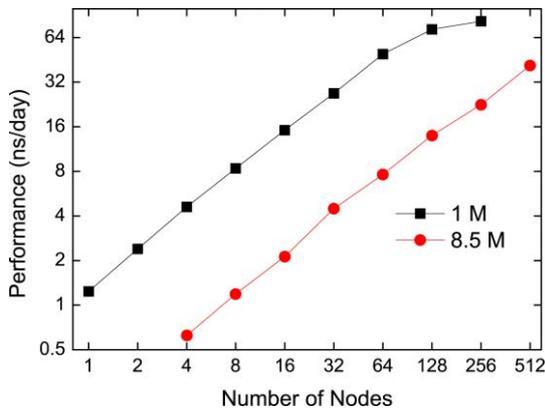


Figure 3 : Performance of 1 million and 8.5 million atom systems

4) 新しいレプリカ交換法を用いたタンパク質構造探索

リガンド結合計算を含む様々な応用計算を行うにあたり、REST 法の各レプリカに与えるパラメータ設定は非常に重要である。しかし現状、これらの方法における標準的なパラメータ決定の方法は示されていない。そこで、今年度はそのパラメータ決定を自動化する方式の検証を行った。REST 法や gREST 法は非常に柔軟な方法であるが故、理論的にきちんとしたパラメータ決定の枠組みを与えることは難しい。そこで、目的とする交換確率からのズレに応じてパラメータの変動量を調整す

るといふ、非常に直観的な方式で試行錯誤を行った。その結果、平衡化を短くし、交換の試行頻度を大きく上げた非常に粗い計算を用いることで高速に良好なパラメータを得られることがわかった。

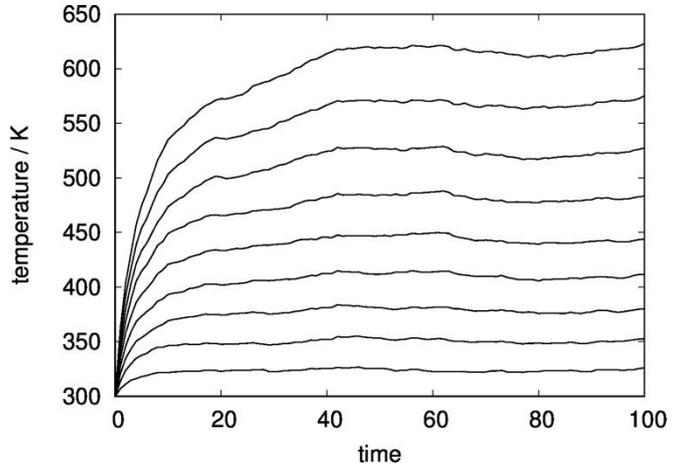


図 4 : パラメータ自動調整の様子

5) Protein G フォールディングのデータ同化解析

Q 座標上で Adaptive sampling を行った結果、Q 空間をくまなく埋め尽くすサンプリングを行うことができた。ただし、これだけでは全ての構造空間を埋め尽くしているかはわからないので、Protein G にラベルしている蛍光色素分子間の距離(ドナーアクセプター間距離)の軸にも投影してチェックした。その結果、Q 空間だけでなく他の軸でも構造を埋め尽くしていることが確認できた。

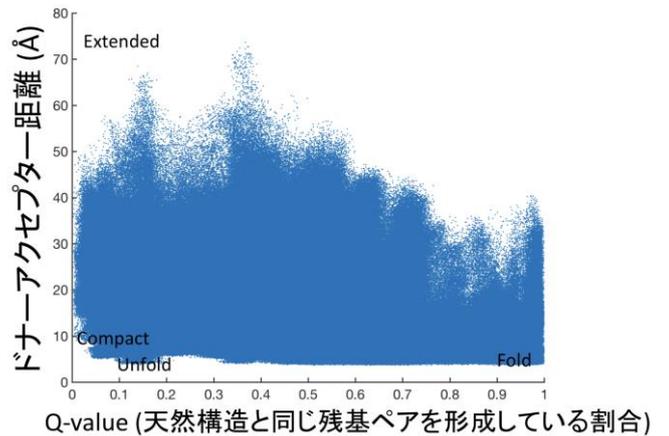


図 5 : サンプル点の二次元への投影

4. まとめ

1) カルシウムイオンポンプの反応機構の解析

本年度は、反応経路解析手法の効果的な利用プロトコルの開発を行い、より効率的にサンプリングを行える手法を構築した。その結果、複雑な構造変化を含む反応経路の解析が可能となった。本年

度では、網羅的な計算を行うために粗視化モデルを用いて解析を行ったが、今回の研究で得られたプロトコルは全原子モデルでの詳細な計算でも対応可能であると考えられる。

2) 分子動力学計算によるヘム輸送体のホモロジーモデルの高精度化

MD 計算によってホモロジーモデルの高精度化を行い、脂質二重膜中での長時間シミュレーションに耐える安定なモデルを生成した。今後はこのモデルを基に、ヘムが輸送される過程のシミュレーションを行う。

3) Optimization of GENESIS MD software with multi-kernel

In this year, we developed GENESIS MD software such that it is optimized for various hardware architectures including Big Waterfall machine, by preparing multi-kernel programs.

4) 新しいレプリカ交換法を用いたタンパク質構造探索

パラメータの自動決定法の開発により、計算効率を向上させることができた。これにより、大規模系を含む様々な系への応用がより現実的になると考えられる。

5) Protein G フォールディングのデータ同化解析
Adaptive sampling を用いることで効率的に Protein G の構造空間をサンプルすることができた。今後、このデータからフォールディングダイナミクスを表現するモデル(マルコフ状態モデル)を構築し、計測データを取り入れてデータ同化を行っていく予定である。

5. 今後の計画・展望

GENESIS の高速 MD エンジンと、マルチレゾリューション計算を組み合わせることで、これまで計算不可能であった挑戦的な計算を行うことができた。既存の計算とは質的に異なる新しい結果が出てきているが、新たな物理モデル・知見とより直接結つく十分なデータを得るため、また「ポスト京」に向けて、来年度も引き続き挑戦的な課題を実行していく。

平成 29 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. J. Jung and Y. Sugita, "Multiple program/multiple data molecular dynamics method with multiple time step integrator for large biological systems", *J. Comput. Chem.* **38**, 1410-1418 (2017). 査読有
2. C. Kobayashi, J. Jung, Y. Matsunaga, T. Mori, T. Ando, K. Tamura, M. Kamiya, and Y. Sugita, "GENESIS 1.1: A hybrid-parallel molecular dynamics simulator with enhanced sampling algorithms on multiple computational platforms.", *J. Comput. Chem.* **38**, 2193-2206 (2017). 査読有
3. M. Kamiya and Y. Sugita, "Flexible Selection of the Solute Region in Replica Exchange with Solute Tempering: Application to Protein-Folding Simulations", *J. Chem. Phys.*, accepted, in press. 査読有
4. Y. Matsunaga, T. Yamane, T. Terada, K. Moritsugu, H. Fujisaki, S. Murakami, M. Ikeguchi, and A. Kidera, "Energetics and conformational pathways of functional rotation in the multidrug transporter AcrB", *eLife*, accepted, in press. 査読有

【国際会議、学会などでの口頭発表】

1. 小林千草, 松永康佑, Jaewoon Jung, 杉田有治, "Molecular dynamics simulations for conformational changes on E1/E2 transition of Ca²⁺-ATPase", 第 55 回日本生物物理学会年会 シンポジウム 2017 年 9 月 21 日 (招待講演)
2. Jaewoon Jung, Chigusa Kobayashi, and Yuji Sugita, "Development of GENESIS for high performance computing of biomolecular simulations", Symposium of "Next-generation in-silico drug discovery using high-performance computing", 第 55 回日本生物物理学会年会 シンポジウム 2017 年 9 月 21 日 (招待講演)
3. Jaewoon Jung, Chigusa Kobayashi, and Yuji Sugita, "Kinetic energy definition in velocity Verlet integration for accurate pressure evaluation", 第 31 回 分子シミュレーション討論会, Nov. 29-Dec. 1, Kanazawa, Japan (oral)
4. K. Tamura, H. Sugimoto, Y. Shiro, Y. Sugita "Computational modeling of the ATP-bound outward-facing form of a heme importer", 第 55 回日本生物物理学会年会、3Pos069、熊本、2017 年 9 月 (poster)
5. K. Tamura, Y. Sugita "Molecular dynamics simulations of membrane transport proteins", 7th JLESC Workshop, Illinois, 2017 年 7 月 (招待講演)
6. 田村康一, "分子シミュレーションによる生体膜トランスポーターの構造機能解析", 研究会「化学反応のポテンシャル面とダイナミクス」, 沖縄, 2017 年 4 月 (招待講演)
7. 松永康佑 "ストリング法による多剤排出トランスポーターAcrB の機能ダイナミクス解析" 高圧力蛋白質研究センター 近畿大学 2018 年 1 月 22 日 (招待講演)
8. 松永康佑 "機械学習を用いた計測とシミュレーションの統合によるタンパク質動態解析" 京都大学理学研究科「数理を基盤として新分野の自発的創出を促す理学教育プログラム(MACS 教育プログラム)」 京都大学 2017 年 11 月 29 日 (招待講演)
9. 松永康佑 "機械学習を用いた 1 分子計測とシミュレーションの統合とタンパク質動態解析" 新学術領域「柔らかな分子系」若手シンポジウム 東北大学 2017 年 9 月 14 日 (招待講演)
10. Y. Matsunaga "Integrative modeling of protein folding dynamics from experiments and simulations" Telluride Workshop on "The Complexity of Dynamics and Kinetics from Single Molecules to Cells", Telluride, Colorado USA, June 20-24, 2017. (招待講演)