

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*
尾嶋 拓***

所属 :

*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

***生命システム研究センター 生命モデリングコア 分子機能シミュレーション研究チーム

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

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 つの研究を行う。

- 1) GENESIS の高度化とベンチマーク
- 2) 大規模構造変化を記述可能な粗視化モデルの開発と、カルシウムイオンポンプの反応機構の解析
- 3) 模式的細胞内環境における GTP 加水分解反応機構の解析

4) WW-domain タンパク質フォールディングのデータ同化解析

5) 新しいレプリカ交換法を用いたタンパク質構造探索

2. 具体的な利用内容、計算方法

1) GENESIS の高度化とベンチマーク

For efficient parallelization of MD, we designed multiple program/multiple data (MPMD) scheme which separates processors responsible for real space and reciprocal space calculations. We also developed multiple time step integration based on MPMD where communication-intensive part is skipped regularly during MD simulations.

2) 大規模構造変化を記述可能な粗視化モデルの開発と、カルシウムイオンポンプの反応機構の解析

特に全原子モデルを用いた計算の高速化を目指し、遅い運動を持つ自由度にかかわる長距離相互作用計算や温度・圧力制御を数ステップ置きに計算する Multiple time step integration を導入し、その精度・速度向上を調べた。全原子モデルは CHARMM36 力場を用いた。すべての計算において GENESIS を用いた。

3) 模式的細胞内環境における GTP 加水分解反応機構の解析

擬似細胞環境内にある Ras/GAP 複合体における GTP 加水分解反応の遷移状態の構造を GAMESS に実装された QM/MM RWFE SCF 法で約 10 回の追加の最適化を行った。続いて、結果として得られた反応物、生成物間の自由エネルギー差を自由エネルギー摂動法で見積もるため、各中間状態(20 点; 反応物、生成物含む)で MD 計算を行った。中間状態は全て同時に生成し、一括で計算することも可能ではあるが、より良いサンプリングを行うため、端点(反応物もしくは生成物)から順番に生成させ、周辺の構造緩和をできるだけ取り入れる形を採用した。中間状態の各点では 50 ns の MD 計算を行った。

4) WW domain タンパク質フォールディングのデータ同化解析

昨年に引き続き、1 分子 FRET データと直接比較するために、蛍光色素分子をラベルした WW domain の全原子モデルを構築し、TIP3P 溶媒モデル下でのフォールディングシミュレーションを行った。GENESIS を用いて、88 レプリカを用いた温度レプリカ交換 MD(REMD)を行って効率的に構造空間をサンプリングした。

5) 新しいレプリカ交換法を用いたタンパク質構造探索

レプリカ交換溶質焼きなまし法(REST 法)とその発展形である我々が開発した gREST 法を人工小タンパク質の Trpcage のフォールディング反応に適用した。REST 法については 10 レプリカ、gREST 法は二面角のみに適用し、5 レプリカを用いてシミュレーションを行った。ソフトウェアには GENESIS を用いた。

結果

1) GENESIS の高度化とベンチマーク

MPMD scheme is developed on HOKUSAI GreatWave and final benchmark is performed on the K computer. Graph shown below is the benchmark result for 28 million atom system.

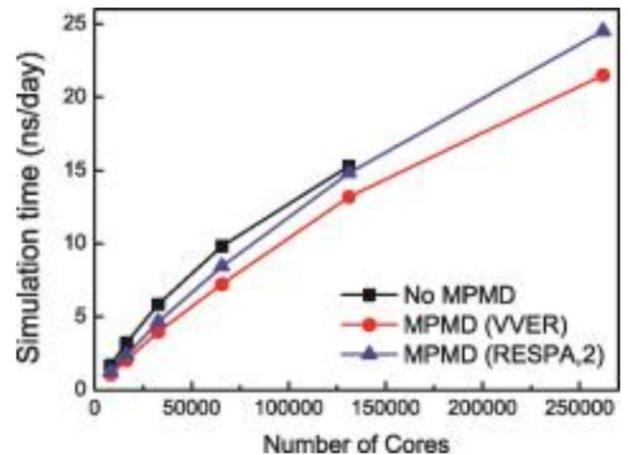


図 1 : MPMD scheme のベンチマーク結果

2) 大規模構造変化を記述可能な粗視化モデルの開発と、カルシウムイオンポンプの反応機構の解析

全原子モデルを用いた計算の高速化に関しては、長距離相互作用や温度・圧力制御の頻度を変化させたシミュレーションによって脂質二重膜、水分子の性質を比較することで、物理的精度を保つパラメータセットを作成した。長距離相互作用や温度・圧力制御の計算は全体の演算時間に対して比較的大きな割合を持っているだけでなく、集団通信を必要とするため、多数のノードを使った際の並列効率を損ねる大きな要因となっていた。Multiple time step integration と物理的精度を保つパラメータの導入・開発により、カルシウムイオンポンプの全原子モデルのシミュレーションにおいて並列効率と演算速度が上昇した。

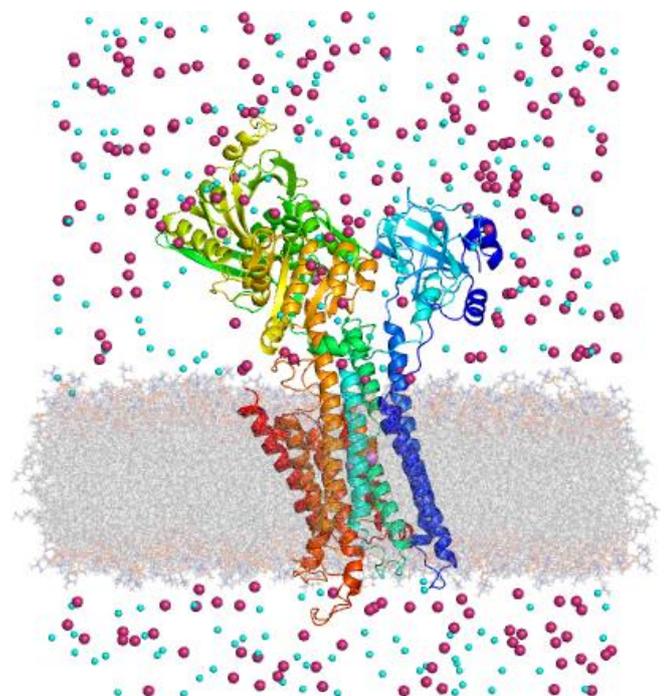


図 2: シミュレーションで得られたカルシウムイオンポンプのスナップショット

3) 模式的細胞内環境における GTP 加水分解反応機構の解析

遷移状態のさらなるリファインメントの結果、高振動モードの解析に耐える程度に最適化された遷移状態の構造を得ることができた。自由エネルギー計算については、現時点ではまだ一部の計算が完了していないために詳細な議論は難しいが、加水分解中の GTP が周囲から感じる静電場は水溶液中の場合とは大きく異なっているため、自由エネルギーもある程度は変わると予想される。その結果から、現実的な細胞環境が化学反応にどの程度の影響を与えうるのかを見積もることができはるはずである。

4) WW domain タンパク質フォールディングのデータ同化解析

昨年に引き続き、88 レプリカを用いた温度レプリカ交換 MD(REMD)を行うことで、天然構造やアンフォールド構造以外の中間構造を広くサンプリングすることができた。昨年同様、REMD シミュレーションデータの一部と、GPGPU を用いた長時間 MD シミュレーションの結果を組み合わせて、マルコフ状態モデルを構築した。そして、構築したマルコフ状態モデルに対して、1 分子 FRET データを用いた機械学習を適用し、遷移確率行列をアップデートした(Baum-Welch アルゴリズムをマルコフ状態モデル向けに改変したものを適用)。昨年は定性的なレベルのモデリングであったが、今年度はパスウェイ解析まで行うことに成功し、フォールディングの中間状態を同定することができた。更に、その構造が別の独立な実験データ(Φ -value analysis)と整合的であることまでわかった。

5) 新しいレプリカ交換法を用いたタンパク質構造探索

従来の REST 法では、Trpcage 全体を溶質として扱い、その構造変化を促進させた。一方、gREST 法では Trpcage 内部の二面角の構造変化だけを促進させ、必要なレプリカ数の削減を図った。以上の条件で Trpcage のフォールディングシミュレーションを各レプリカあたり 1,500 ns 行った。

その結果、どちらの方法でも安定な天然構造を得ることができた。一例として、gREST 計算の結果を以下の図に示した。初期構造では完全にアンフォールドした構造が、1,000 ns 程度でフォールディングできている。gREST 法と REST 法でフォールディングに要した時間はほぼ同等であり、gREST 法の方が REST 法よりもレプリカ数が少ないため、gREST 法はこのような種類の計算に有用であると考えられる。

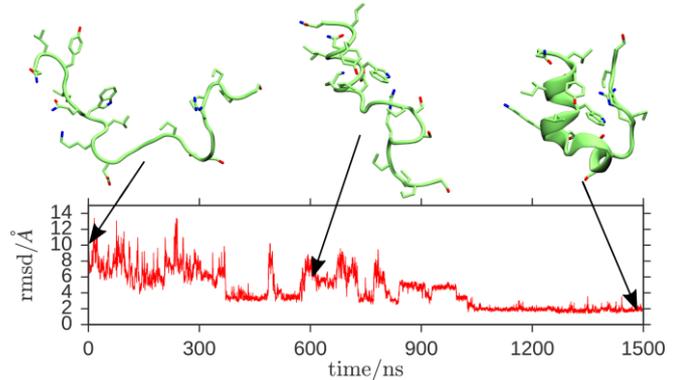


図 3: gREST 法で得られた Trpcage のフォールディングトラジェクトリの一例。縦軸は天然構造との構造差を示す RMSD で、1,000 ns 付近で天然構造(右端)に到達している。

なお、通常の MD 計算では、3,000 ns の計算を行ってもフォールディングは観測できなかった。

3. まとめ

1) GENESIS の高度化とベンチマーク

We have developed new parallelization scheme suitable for massively parallel supercomputers in Hokusai GreatWave or K, suitable for large scale MD.

2) 大規模構造変化を記述可能な粗視化モデルの開発と、カルシウムイオンポンプの反応機構の解析

本年度は、詳細な反応機構の解明のため、全原子モデルの高速化を目指した。遅い運動を持つ自由度にかかわる長距離相互作用計算や温度・圧力制御を数ステップ置きに計算する MTS 法と、精度を保つパラメータの導入により、高速に効率良く求める計算することが可能となった。これらにより、粗視化モデルと全原子モデルの二つの利点を能く取り入れた計算が可能となった。

3) 模式的細胞内環境における GTP 加水分解反

応機構の解析

遷移状態のさらなるリファインメントにより、より詳細な解析が可能となった。自由エネルギー計算の完了にはもう少し時間を要するが、それにより細胞内環境の影響をある程度見積もることができるはずである。また、初期構造を変化させて比較することで、その影響をもう少し定量的に見積もることも計画している。

4) WW domain タンパク質フォールディングのデータ同化解析

構築したマルコフ状態モデルをパス解析することでフォールディング時の構造変化計算を行い、フォールディング機構に関する物理的知見を得ることができた。

5) 新しいレプリカ交換法を用いたタンパク質構造探索

小タンパク質のフォールディングにおいて、開発した **gREST** 法は、従来法である **REST** 法と同程度の結果をより少ない計算コストで得ることができた。今後は、ドッキング等の計算や、多次元化等によるさらなる構造探索の効率化についての知見を得ることが重要である。

4. 今後の計画・展望

今年度は、**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、来年度も引き続き同様の課題を実行していく。

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. I. Yu, T. Mori, T. Ando, R. Harada, J. Jung, Y. Sugita, and M. Feig, “Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm.” *eLife* **5**, e19274 (22 pages) (2016) 査読有
2. J. Jung, C. Kobayashi, T. Imamura, and Y. Sugita, “Parallel implementation of 3D FFT with volumetric decomposition schemes for efficient molecular dynamics simulations” *Computer Physics Communications* **200**, 57-65 (2016) 査読有
3. J. Jung, A. Nourse, C. Kobayashi, and Y. Sugita, “Graphics Processing Unit Acceleration and Parallelization of GENESIS for Large-Scale Molecular Dynamics Simulations” *Journal of Chemical Theory and Computation* **12**, 4947-4958 (2016) 査読有
4. Y. Matsunaga, Yasuaki Komuro, Chigusa Kobayashi, Jaewoon Jung, Takaharu Mori, and Y. Sugita, “Dimensionality of Collective Variables for Describing Conformational Changes of a Multi-Domain Protein” *The Journal of Physical Chemistry Letters* **7**, 1446–1451 (2016) 査読有
5. 松永康佑 “実験と計算の時間スケールが重なったことで見えてきたこと”
蛋白質科学会アーカイブ http://www.pssj.jp/archives/essay/Es_05/Es_05.html 査読無

【国際会議、学会などでの口頭発表】

6. Yuji Sugita and Jaewoon Jung, “Macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm”, The third Korean-Polish conference on “Protein Folding: Theoretical and Experimental Approaches”, Korea (oral)
7. 小林千草、松永康佑、Jaewoon Jung, 杉田有治、”P-type ATPase の大規模構造変化を伴う反応メカニズム解析のためのマルチレゾリューションシミュレーション手法” 第 89 回日本生化学会大会 仙台 2016 年 9 月 26 日 (oral)
8. Motoshi Kamiya and Yuji Sugita, “In silico folding simulation of Trp-cage using the REST method and its variant”, 第 5 4 回生物物理学会年会 つくば 2016 年 11 月 25-27 日 (poster)
9. 松永康佑 “1 分子 FRET データと分子動力学シミュレーションによるタンパク質ダイナミクス解析” 第 16 回日本蛋白質科学会年会 ワークショップ「生命分子計測のスパースモデリング：限られた計測データから生命分子の構造情報を引き出す」 福岡 2016 年 6 月 7 日-9 日 (oral)
10. 松永康佑 “1 分子 FRET データと分子動力学シミュレーションによるタンパク質ダイナミクス解析” 第 16 回日本蛋白質科学会年会 若手奨励賞シンポジウム 福岡 2016 年 6 月 7 日-9 日 (oral)
11. 松永康佑 “生体分子におけるデータ同化” 理研データ同化ワークショップ 神戸 2016 年 10 月 14 日 (oral)
12. 松永康佑 “タンパク質構造変化における反応自由度探索法” 第 10 回分子シミュレーションスクール -基礎から応用まで- 岡崎 2016 年 10 月 19 日 (oral)
13. Yasuhiro Matsunaga “Markov state modeling of protein folding dynamics by combining single-molecule experiments and simulations” Simulations Encounter with Data Science, Institute of Statistical Mathematics, Tokyo, Japan, March 9-11, 2017. (oral)
14. Yasuhiro Matsunaga “Drug extrusion mechanism of the multidrug transporter AcrB studied by molecular dynamics simulation” 日本化学会第 97 春季年会・アジアシンポジウム 慶応日吉キャンパス 2017 年 3 月 18 日 (oral)
15. 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” 第 10 回京阪奈計算

平成 28 年度 利用報告書

生物物理学セミナー 京都大学 2017 年 3 月 22 日 (oral)

【その他 (プレスリリース、学術会議以外の一般向けの講演など)】

1. 講習会 2016 年度第 16 回 AICS 公開ソフト講習会「GENESIS」 計算科学振興財団(FOCUS) 神戸 2017 年 1 月 13 日