

京コンピュータを用いたウイルスの全原子シミュレーション

2. 高並列汎用分子動力学シミュレーションソフト MODYLAS の開発

(名大院・工¹, 名大院・工 計算セ², 立命館大・薬³, 金沢大・理工⁴, 分子研⁵, 理研⁶, 富士通⁷)

安藤 嘉倫¹, ○吉井 範行^{1, 2}, 藤本 和士³, 水谷 圭佑¹, 小嶋 秀和¹, 山田 篤志¹,
岡崎 進¹, 川口 一朋⁴, 長尾 秀実⁴, 岩橋 建輔⁵, 水谷 文保⁵, 南 一生⁶,
市川 真一⁷, 小松 秀実⁷, 石附 茂⁷, 武田 康宏⁷, 福島 正雄⁷

Large-scale All-atom Molecular Dynamics Calculation of Viruses using K-computer

2. Development of a Highly Parallelized General-Purpose Molecular Dynamics Simulation Program, MODYLAS

(Graduate School of Engineering, Nagoya Univ.¹, Center for Computational Science, Nagoya Univ.², Department of Pharmaceutical Sciences, Ritsumeikan Univ.³, Graduate School of Natural Science and Technology, Kanazawa Univ.⁴, Institute for Molecular Science⁵, Advanced Institute for Computational Science, Riken⁶, Computational Science and Engineering Solution Division, Fujitsu Limited⁷) Yoshimichi Andoh,¹ ○Noriyuki Yoshii,^{1,2} Kazushi Fujimoto,³ Keisuke Mizutani,¹ Hidekazu Kojima,¹ Atsushi Yamada,¹ Susumu Okazaki,¹ Kazutomo Kawaguchi,⁴ Hidemi Nagao,⁴ Kensuke Iwahashi,⁵ Fumiyasu Mizutani,⁵ Kazuo Minami,⁶ Shin-ichi Ichikawa,⁷ Hidemi Komatsu,⁷ Shigeru Ishizuki,⁷ Yasuhiro Takeda,⁷ and Masao Fukushima⁷

【はじめに】

均一な液相や固相をはじめ、低分子や高分子水溶液、脂質膜やタンパク質からなる複合体など、さまざまな分子集合体についてその物理化学的性質を明らかにするために、分子動力学 (MD) 計算が広く用いられてきた。コンピュータの性能向上に伴って大規模系の MD 計算が行われるようになり、近年では長距離静電相互作用を含む系についても一千万原子を超えるものが実施されつつある。しかしながら、MD ソフトウェアに目を向けると、クラスターマシン程度では高い性能を発揮するものが見受けられるものの、超高並列コンピュータの性能を十分に引き出して大規模系を効率的に実行可能なものは現状でもほとんど開発されていない。そこで我々は「京」スパコンのような超高並列コンピュータを用いて大規模系の MD 計算を高効率に実行することを目指して、高並列汎用分子動力学シミュレーションソフト MODYLAS[1]の開発を進めてきた。

【開発内容】

MODYLAS を使った研究対象として、ウイルス、タンパク質複合体、ベシクルならびに高分子材料や有機・無機材料などといった水溶液系での自己組織化構造体の物理化学を想定している。そのためには近距離 LJ 相互作用計算だけではなく、長距離静電相互作用および各種汎用ポテン

シヤル関数に応じた分子内相互作用計算を行うことができなければならない。また研究での実用上、各種統計アンサンブル、拘束の動力学についても実装が欠かせない。

一方、対象マシンに想定した「京」の特徴として、その膨大なノード数およびノード間を繋ぐトラスネットワークがある。これらハードウェア要請から「京」上にて高並列化性能を実現するためには、(1) 対 MPI 並列スケーラビリティに優れた計算アルゴリズム、(2) トラスネットワークに適した隣接通信方式、を実装する必要がある。(1) について、MD 計算では特に静電相互作用が問題となる。一般的な MD ソフトウェアは高速フーリエ変換(FFT) にもとづく Particle Mesh Ewald(PME)法によって静電相互作用を扱っている。しかしながら、たとえ最新の FFT ルーチンを使用したとしても、一般的には 512 MPI 並列程度で FFT 並列性能が飽和するという問題がある[2]。MODYLAS ではこの点を回避するために従来の PME 法に加え、FFT を行わない高速多重極展開法を選択できる仕様とした。(2) については、まず基本セルをトラスメッシュ位置との対応をとる形でプロセスごとのセル・ブロックに分割した。その上で、必要に応じて周辺プロセスとのみ座標、速度、力および多極子情報を隣接一対一通信する仕様とした。通信/演算両配列の共用を前提としたメタデータ構造を実装し、超高並列時に律速となりうる配列間コピーを隠滅した。並列方式は MPI および OpenMP のハイブリッド型であり、演算量の多い部分は SIMD 化している。力場は生体高分子系で標準的な CHARMM (with CMAP) および AMBER、OPLS に対応している。

【実行性能】

図1はタンパク質分子を含む1000万原子系をインプットとして「京」上で性能を測定した結果である。実際的な系に対して 65,536 MPI 並列 (524,288 コア) 時においても、対 64 ノード比で高いスケーラビリティを達成していることがわかる。さらに 1000 万原子系について、1 MD ステップをわずか 5ms 程度で実行している。これは 3 日で 100 ns の MD 計算が可能であることを意味しており、長時間計算においても MODYLAS が非常に有効であることを示している。

今後さらに発展を遂げるであろう超高並列コンピュータに MODYLAS を適用することにより、HIV やインフルエンザといったウイルス、あるいは細胞小器官といった極めて大規模かつ重要な系についても、全原子モデルとして取り扱えるようになるものと期待される。

【参考文献】

- [1] MODYLAS の詳細については Y. Andoh, *et al.*, *J. Chem. Theory Comput.* **9**, 3201-3209 (2013)を参照のこと。なお、MODYLAS は近日中に公開する予定である。
- [2] C. C. Gruber, J. Pleiss, *J. Comput. Chem.* **32**, 600-606 (2011).

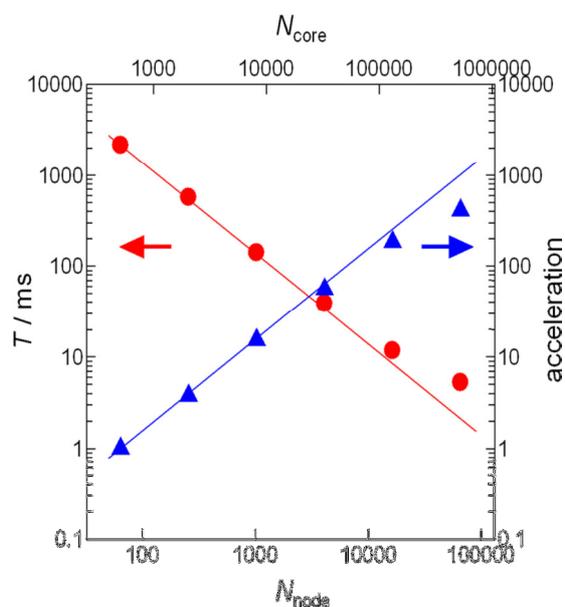


図1. 1000万原子からなるタンパク質水溶液系の1MDステップあたりに要する計算時間(赤)と、並列ノード数 N_{node} に対する加速率(青、対64ノード比)。