

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 : ○木寺 詔紀, 寺田 透, 森次 圭

理研での所属研究室名 :

社会知創成事業 次世代計算科学研究開発プログラム

次世代生命体統合シミュレーション研究推進グループ 分子スケール研究開発チーム

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

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。

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 ($\mu 2lib$) の開発を行っている。昨年度までに、相互作用計算ループ部分についての OpenMP 化、また、コピー間の MPI 集団通信を避けるための隣接通信版アルゴリズムとその実装を行った。本年度は、更なるスケーラビリティの向上のための後述のアルゴリズム開発とそのテスト計算を RICC の上限である 8192 コアまでコア数を増やしながら実行した。

ベンチマーク計測の計算対象としては、adenylate kinase (3,343 原子)を用いた。相互作用計算には 2 つの手法を適用した。一つは GB/SA と呼ばれる溶媒和モデルであり、溶媒自由度を陽に含まない。No-cutoff のた

め、各プロセスが担当する相互作用数は不変である。もう一つは smooth particle mesh Ewald 法(SPME)で、水分子(TIP3P モデル)・イオンを陽に含んだ計 62,475 原子系で計算を行う。Cutoff 半径を行き来する水分子の存在により各プロセスが担当する相互作用数が変化しロードバランスにばらつきが生じる。

3. 結果

(1) 1 プロセスマルチコピー版の開発

通常のマルチコピーシミュレーションでは 1 プロセスが 1 コピーを担当する。コピー数を固定した strong scaling を計測する場合、使えるコア数を 2 倍にすると 1 プロセスあたりのコア数も 2 倍になるが、これは単体 (1 コピー) での strong scaling を計測しているのと同様である。分子動力学シミュレーションでは系のサイズによりスケールするコア数が限定される、つまり、RICC の上限である 8192 コアまで単体シミュレーションでスケールするためには巨大タンパク質系に適用しなければならない。

このような系のサイズ依存性を解決するため、1 プロセスマルチコピーのアルゴリズムを適用した。この手法では、1 プロセスが 1 コピーを担当して使うコア数を変動するかわりに、コア数を固定しコピー数を変える。使えるコア数を 2 倍にすると、1 プロセスあたりのコア数は変わらないがコピー数が半分になる。これは単体での strong scaling 計測とは異なり、マルチコピーシミュレーションに向けた strong scaling 計測であると言える。

adenylate kinase に適用した結果、GB と Ewald 両方について並列化度 99.999% 以上のスケーラビリティを達成した (図 1)。よって、このアルゴリズムはマルチコピーシミュレーションの高速化に役立つと考えられる。

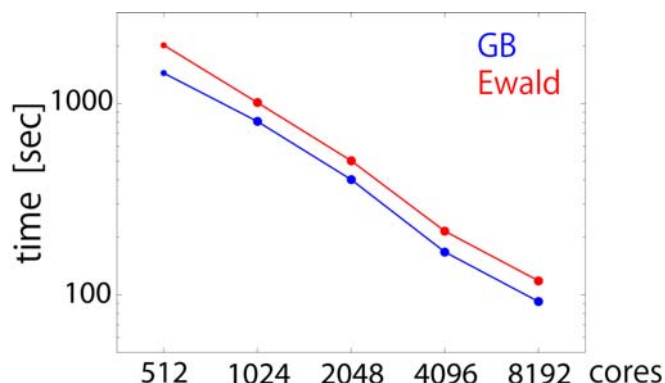


図 1: 1 プロセスマルチコピー版のスケーラビリティ

テスト。512 コピー、64 コア/コピーを固定して strong scaling を計測。

4. 今後の計画・展望 (これまで利用した状況や、継続して利用する際に行う具体的な内容)

来年度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最終年度にあたり、今年度に引き続き「京」でのベンチマークが中心になると考えられる。特にマルチコピーシミュレーションでの $\mu 2lib$ のスケーラビリティ向上をテストするために週末利用時にベンチマーク計測を行い、新規アルゴリズム開発やコード改良を進めていく。

平成 23 年度 RICC 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. Tohru Terada and Akinori Kidera, "Comparative molecular dynamics simulations study of crystal environment effect on protein structure" submitted to Journal of Physical Chemistry B.
2. 森次圭、「粗視化モデルを用いたマルチスケールシミュレーションシステムの構築」、生物物理学会誌、52, 026 (2012).
3. Kei Moritsugu, Tohru Terada and Akinori Kidera, "Disorder-to-order transition of an intrinsically disordered protein revealed by multiscale enhanced sampling", submitted to Journal of the American Chemical Society.

【国際会議などの予稿集、proceeding】

該当なし

【国際会議、学会などでの口頭発表】

1. 寺田 透
「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プログラム開発のためのクラスライブラリ」
日本生物物理学会第49回年会、兵庫、2011年9月
2. Kei Moritsugu, Tohru Terada and Akinori Kidera
“Multiscale enhanced sampling simulation for protein-protein interaction”
日本生物物理学会第49回年会、兵庫、2011年9月
3. 森次 圭、木寺 詔紀
「マルチスケール・マルチコピーシミュレーション (MSES 法) による天然変性タンパク質 sortase の構造サンプリング」
日本物理学会 2011 年秋季大会、富山、2011 年 9 月
4. Kei Moritsugu, Tohru Terada and Akinori Kidera
“Molecular mechanism of allosteric regulation of an intrinsically disordered protein, sortase, revealed by comprehensive all-atom conformational sampling simulations”
Conference on Computational Physics (CCP) 2011, アメリカ、2011 年 10-11 月

【その他】

該当なし