

先端研究施設共用イノベーション創出事業【産業戦略利用】『みんなのスパコン』TSUBAME によるペタスケールへの飛翔
利用報告書 平成 20 年度 課題種別 課題 ID :i08ne087

利用課題名 生体高分子用シミュレーションソフトウェア DS CHARMM の大規模系における並列性能評価
英文: Scalability Evaluation of DS CHARMM on huge systems.

利用課題責任者 福島信弘

Fiestname Surname Nobuhiro Hukushima

所属 サイエンス・テクノロジー・システムズ株式会社

Affiliation Science & Technology systems, inc.

URL <http://www-st-systems.co.jp>

邦文抄録

本プロジェクトでは, Accelrys 社の CHARMM を使用し, TUBAME で本製品が性能を発揮できるのか, また十分に性能を発揮するためにはどの程度の大きさの系が適当なのかを評価することを目的としている。そのために DNA 2 重らせん構造の分子動力学シミュレーションを実施した。また分子動力学シミュレーションとしても大型なものに属する HIV 逆転写酵素阻害剤/ヌクレオチド基質複合体 PDB ID:1T05(17500 原子), およびショウジョウバエヌクレオソームコア PDB ID: 2PYO(22200 原子)のモデルを構築し, 計算を試みた。

英文抄録

The purpose of this project is to evaluate how much size of system is necessary for TUBAME whose performance is demonstrated when the CHARMM of Accelrys Inc. is used. So this project uses the DNA double helix, HIV-1 reverse transcriptase cross-linked to template-primer (PDB ID: 1T05) and drosophila nucleosome core (PDB ID 2PYO) for calculating model and tries to simulate with molecular dynamics.

Keywords: CHARMM, DNA double helix, HIV-1 reverse transcriptase cross-linked to template-primer, drosophila nucleosome core, Molecular Dynamics.

背景と目的

現代の化学は物質の構造/物性を原子/分子レベルで探求する学問である。計算機シミュレーションは, 計算機上に系のモデルを作成し, 原子/分子の構造や性質をミクロのレベルで解析することができる有効な手法であり, ミクロとマクロをつなぐ研究ツールとして広く用いられている。しかし, パソコンやワークステーションを用いて計算機シミュレーションを行うのはモデルのサイズや計算時間の問題から限界があり, 計算機シミュレーションはどうしても, まだまだ次世代技術だと言われてしまうレベルから脱出できていない。

だが, 国内最高レベルの処理性能を誇る TUBAME を活用することにより, 実用レベルの計算機シミュレーションの実現が可能であり, 実験の後づけというレベルを超えた, 実験に先行した計算機シミュレーションによる最先端の研究が可能になると期待される。

本プロジェクトでは, Accelrys 社の CHARMM を使用し, TUBAME で本製品が性能を発揮できるのか, また

十分に性能を発揮するためにはどの程度の大きさの系が適当なのかを評価することを目的としている。そのために DNA 2 重らせん構造の分子動力学シミュレーションを実施した。また生体分子の分子動力学シミュレーションで扱えるレベルのものの中でも大型なものに属する HIV 逆転写酵素阻害剤/ヌクレオチド基質複合体 PDB ID:1T05(17500 原子), およびショウジョウバエヌクレオソームコア PDB ID: 2PYO(22200 原子)の TUBAME を用いた実験に先行する計算機シミュレーションによる予測を実施した。

概要

米国 Accelrys 社製 Discovery Studio は, ライフサイエンス分野の次世代型モデリング・シミュレーションソフトウェアであり, 約 20 程のモジュール群から構成されている。その中でも, 古くから定評があり, デファクトスタンダードなシミュレーションモジュールである DS CHARMM を用いて大規模生体高分子のスケラビリティ

(様式第 20)

ティの検証を行う。今までに開発元である米国 Accelrys 社及びアクセルリス株式会社では、大規模コンピュータシステムを所有しておらず、またユーザにおいては、ライセンス料金の問題やシミュレーション内容自体が、開示できないこともあり、高々64 並列程度しか検証できていないのが現状である。貴学の TSUBAME システムを利用させていただくことで、128, 256, 512, 1024 並列といった非常に高並列までの性能向上率検証を行うことができ、CHARMm と関連モジュールを用いて、逆転写酵素+DNA+RNA+阻害剤+溶媒といった大規模複合系について計算可能となることが期待される。同時に Discovery Studio の機能である力場自動生成モジュールについての大規模系への精度検証もできることも期待される。

結果および考察

まず性能評価のために DNA の 2 重らせん構造にて分子動力学シミュレーションを行った。

DiscoveryStudioの”Nucleic Acid Modeling” ツール(図 1)を使用し、図 2のようなDNAの 2 重らせん構造を作成した。

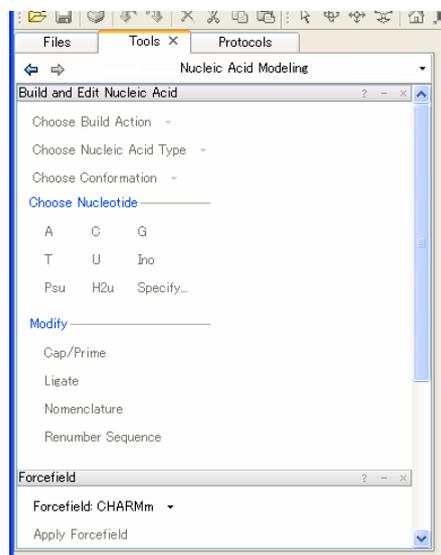


図 1 DiscoveryStudio の”Nucleic Acid Modeling” ツール

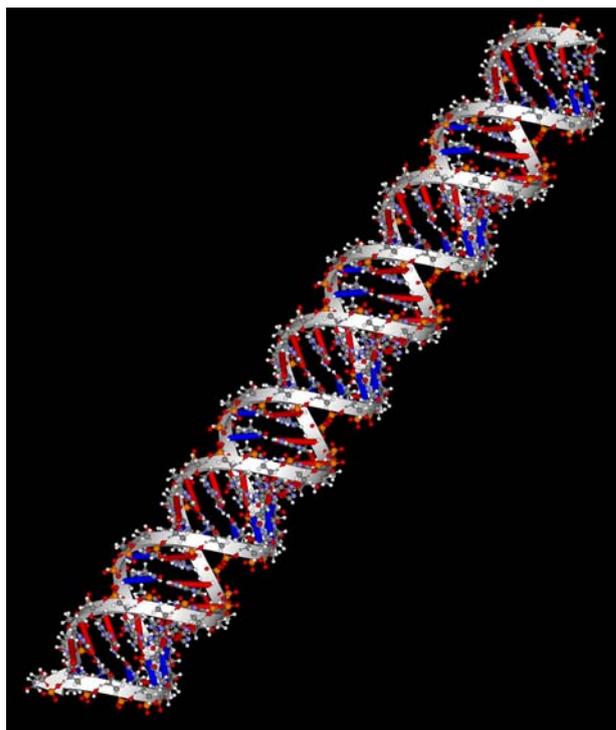


図 2 DNA の 2 重らせん構造

残基数については 100 残基と 300 残基のものを用意した。まずそれぞれに対して Minimization を行った。この Minimization については TUBAME を使用せず、弊社の計算マシンにて行った。Minimization 後に、TUBAME にて、Dynamics を計算した。その際のパラメーターは以下の通りである。

表 1 Production parameter

Production Steps	100000
Production Time Step	0.001 [nsec]
Production Target Temperature	300.0 [K]
Production Type	NVT
Nonbond List Radius	14.0 [Å]
Nonbond Higher Cutoff Distance	12.0 [Å]
Nonbond Lower Cutoff Distance	10.0 [Å]

使用 CPU 数を変えて、性能向上率検証を行った。結果は以下の通りである。なお 1 ノード当たり 1CPU を使用した。

① 100 残基の場合

表 2 CPU 数と計算時間 (100 残基)

CPU 数	経過時間
4	25.28 [min]
8	15.93 [min]
16	27.68 [min]
32	25.64 [min]

② 300 残基の場合

表 3 CPU 数と計算時間 (300 残基)

CPU 数	経過時間
4	88.2 [min]
8	53.16 [min]
16	40.59 [min]
32	68.4 [min]
64	68.4 [min]
128	714 [min]

100 残基の計算結果については 8CPU, 300 残基の計算結果は 16CPU 程度を境にそれ以上 CPU 数を増やしても、逆に性能は低下する結果となった。これは系の大きさが自身の大きさに対して、並列数が多いため、プロセスの分割に返って時間が掛かってしまうためである。

以上の結果から TSUBAME にて 128, 256, 512CPU を使用して性能を発揮するためには 2000 残基を超えるような大きなモデルで計算する必要があることが分かったが、計算するまでには至らなかった。

次に生体分子の分子動力学シミュレーションで扱えるレベルの系の中でも大型なものに属する HIV 逆転写酵素阻害剤/ヌクレオチド基質複合体(図 3)およびショウジョウバエヌクレオソームコア(図 4)のモデリングを DiscoveryStudio にて行った。しかし、採択直後のメンテナンスによるシステム停止を始め、たびたび発生する障害および学内需要の上昇による混雑等によって TUBAME の活用を行うことができなかった。



図 5 HIV 逆転写酵素阻害剤/ヌクレオチド基質複合体

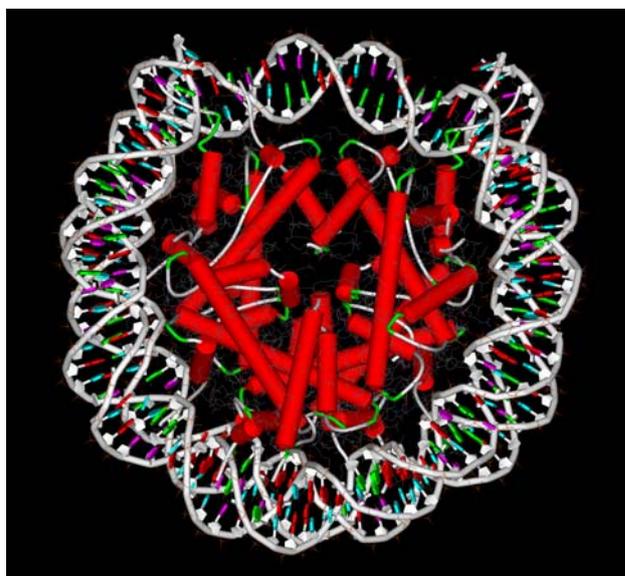


図 6 ショウジョウバエヌクレオソームコア

まとめ、今後の課題

今回はモデルの大きさが小さかったため、TSUBAME の性能を十分に発揮できなかった。TSUBAME の性能を十分に発揮するには 2000 残基、5000 残基といった大きなモデルが必要であると考えられる。また、今回は大きな系として HIV 逆転写酵素阻害剤/ヌクレオチド基質複合体及びショウジョウバエヌクレオソームコアの計算を試みたが、残念ながら結果を出すことができなかった。本法は今後の TUBAME の活用方法例としては有効な方法であると考えられる。