

課題名 (タイトル) :

全電子計算に基づくタンパク質反応シミュレーションの研究

利用者氏名 : ○木寺 詔紀*, 佐藤 文俊*, 平野 敏行**, 松田 潤一**

所属 : *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

**東京大学生産技術研究所

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

本課題は「次世代生命体統合シミュレーションソフトウェアの研究開発」プロジェクト、分子スケール研究開発チームに参画中の、全電子計算に基づくタンパク質反応シミュレーションプログラム ProteinDF のチューニングのために実施された研究である。「京」次世代スーパーコンピュータは 8 万ノード・64 万コア以上の計算ノードから構成されており、高度に並列化されたプログラムが必要とされる。ProteinDF が「京」次世代スーパーコンピュータ上で効率的に並列計算できるように開発・テストすることを目的とした。

2. 具体的な利用内容、計算方法

1000 残基クラスのタンパク質全電子カノニカル分子軌道計算の場合、行列のサイズは 80GB にも及ぶ。そこで ProteinDF は大規模分子の全電子カノニカル分子軌道計算を限られたメモリ上で遂行するために、巨大行列の分散保持を行い、それに伴う行列要素計算(分子積分)を並列化している。ProteinDF はすでに MPI/OpenMP ハイブリッド並列化が行われており、この分子積分計算もハイブリッド並列計算により行われる。

分散メモリ型並列計算機に分散保持された行列に対する分子積分計算は、共有メモリ型に比べてメモリアクセスに時間がかかるため、高速計算に著しく不利である。そこで総メモリ量は必要になるが行列要素の計算に高速な行列積が利用できるコレスキー分解法を開発し、ProteinDF に実装した。

3. 結果

本研究で開発・チューニングした ProteinDF を実際に「京」上でビルド・動作確認をすることができた。ただし、単体性能が芳しくないために大

規模計算に至っていない。単体性能の向上が課題であった。C++コンパイラ開発元の富士通とも協力し、ProteinDF のチューニング作業を行った。現在、C++コンパイラの最適化が適切に行われるように改良を行なっているところである。

4. まとめ

次世代スーパーコンピュータ「京」上で効率良く ProteinDF を実行できるようにするため、コレスキー分解法を開発を行った。また、単体性能の向上を図り、分子積分ルーチンのチューニングを行った。

5. 今後の計画・展望

FX10 を利用し、引き続き ProteinDF の単体性能向上を図る。計算コストの高い分子積分において、その計算回数を大幅に削減するこれスキー分解法を開発できたため、これを FX10 上で最適化する。次世代スーパーコンピュータ「京」や FX10 に適した ProteinDF が完成し、前人未踏の大規模タンパク質量子化学計算が達成したいと考えている。

6. 利用がなかった場合の理由

本研究の目的は、「京」スーパーコンピュータへのチューニング・ポータリングであり、理化学研究所外からの VPN アクセスに利用した。計算機リソースは次世代スーパーコンピュータ「京」を利用した