

タンパク質量子化学計算支援統合環境 ProteinEditor の開発

○吉廣保¹, 西村康幸¹, 佐藤文俊^{1,2}¹東京大学生産技術研究所(〒153-8505 東京都目黒区駒場 4-6-1)²東京大学情報基盤センター(〒113-8658 東京都文京区弥生 2-11-16)

【緒言】

本研究グループはタンパク質のための密度汎関数法プログラム ProteinDF^[1]を基にタンパク質の統合的な量子化学計算シミュレーションシステムの開発を行っている^[2]。本システムは「計算エンジン ProteinDF」を中心に「自動計算法」、「構造最適化・*ab initio*分子動力学(MD)」などのプログラムで構成されており、これらを統括するグラフィカル・ユーザ・インターフェース(GUI)が「ProteinEditor」(図 1)である。現在ではタンパク質の基底状態全電子計算達成をサポートする機能をほぼ全て装備している。本発表では、ProteinEditor の特徴的な機能について報告する。

【目的】

一般に、タンパク質のシミュレーションには初期構造に Protein Data Bank (PDB)^[3]で配布されている 3 次元構造データを用いることが多い。PDB に登録されている立体構造は X 線構造解析や多次元 NMR などによって求められたものであるが、その 8 割以上が X 線構造解析によるデータで、水素原子の座標データを持っていない。さらに、化学的な見地からは異常な構造歪みや原子間衝突などを残したまま構造精密化が行われていることも多く、これらは量子化学計算には不向きな構造といえる。適切なタンパク質の計算構造を構築することは、今やタンパク質量子化学計算にとって最も重要なプロセスの 1 つであると言っても過言ではない。そこで ProteinEditor に、PDB データへの水素付加やアミノ酸置換、モデリング支援機能、MD と連動させて構造歪みを緩和する機能、および構造妥当性のチェックを支援する機能の実装を行う。

【方法】

開発言語は C++ を用いており、GUI は Visual Studio .net 2003 を、分子モデル表示には OpenGL を用いており、Windows 2000/XP 上で実行可能である。

【結果】

複雑なタンパク質立体構造の 3 次元空間における位置関係を視覚的に把握するため、各種の機能表現に必要なタンパク質のための大規模分子グラフィックスを OpenGL で実装し、裸眼立体視デバイスにも対応した。大規模分子グラフィックスには、タンパク質モデリング支援機能も実装しており、分子グラフィックスからインタラクティブ

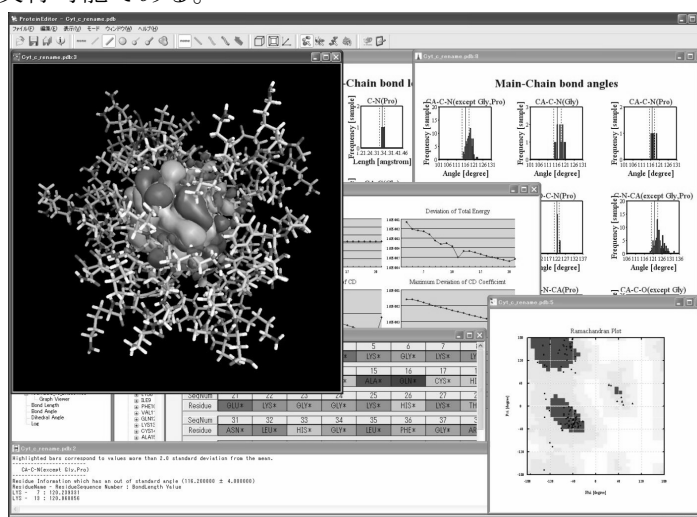


図 1 ProteinEditor のスナップショット

に、原子(単独および特定原子種集団)、基、側鎖、アミノ酸残基(群)、基質、配位子といった様々な単位で構造を編集することが可能である。

タンパク質の構造編集後には、構造精密化を適切に実行することが必要である。加えて、ユーザが独自に力場を指定して局所部分構造最適化を行う手法も簡便かつ有効である。ProteinEditorは、この局所部分構造最適化機能を実装している。また、PROCHECK^[4]に相当する機能として結合長分布、結合角分布、二面角分布、ラマチャンドラプロットの妥当性評価支援機能を実装しており、構造確定後には、この機能により構造妥当性を評価し量子化学計算の初期値として適切な構造になっているかを判断することができる。これらの妥当性評価支援機能は分子グラフィックスと連携でき、構造に問題がある部位を分子グラフィックスに強調表示することも可能である。強調表示された部位に関しては再修正などの構造精密化を直ちに施すことが可能である。

国際蛋白質構造データバンク(wwPDB)^[5]と日本蛋白質構造データバンク(PDBj)^[6]が提供しているPDBデータは2007年8月1日よりIUPAC命名法に従う化合物の記述の標準化等を含む新しいフォーマットに変更された^[6]。ProteinEditorは、この新しいフォーマットにも対応した。

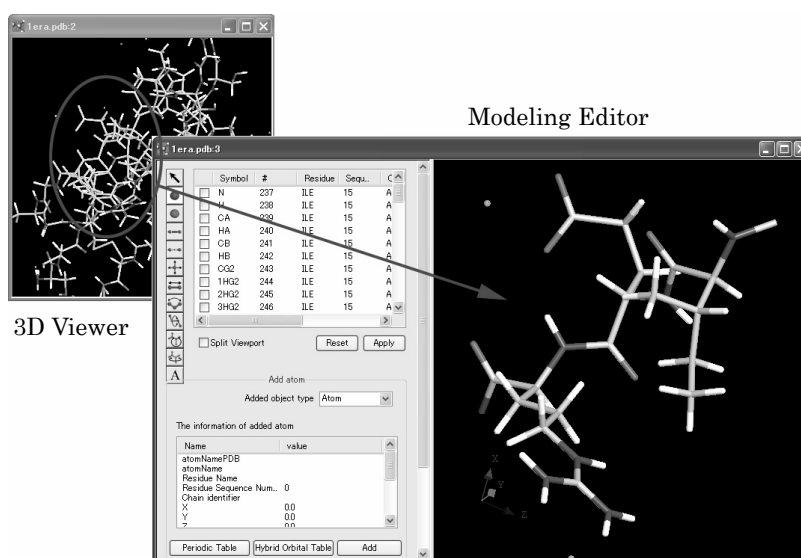


図2 タンパク質モデリング支援機能
選択した構造を抜き出し、Modeling Editorで編集

【謝辞】

本研究は、文部科学省 IT 基盤構築のための研究開発プログラム「革新的シミュレーションソフトウェアの研究開発」において実施された。

【参考文献】

- [1] F.Sato, T.Yoshihiro, M.Era, H.Kashiwagi, *Chem. Phys. Lett.*, **341** (2001) 645.
- [2] 革新的シミュレーションソフトウェアの研究開発, <http://www.rss21.iis.u-tokyo.ac.jp/>.
- [3] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *The Protein Data Bank, Nucleic Acids Research*, **28** (2000) 235. <http://www.rcsb.org/pdb/>.
- [4] Laskowski R A, MacArthur M W, Moss D S, Thornton J M, *J. Appl. Cryst.*, **26** (1993) 283
- [5] wwPDB: <http://www.wwpdb.org/>; PDBj: <http://www.pdbj.org/>
- [6] Helen M Berman, Kim Henrick, Haruki Nakamura, John Markley, Philip E Bourne, John Westbrook, *Nature Biotechnology*, **25**, (2007) 845. <http://www.wwpdb.org/>