



分類	量子化学計算
キーワード	タンパク質、量子化学計算、密度汎関数法
開発者	佐藤文俊、井原直樹、恒川直樹、西野典子、西村康幸、平野敏行、吉廣 保、小池 聡、西村民男、佐藤昌宏、小池秀耀、柏木 浩
作成年月日	2006年6月
コード名	ProteinDF
使用言語	C++

◇ ProteinDF システム

当グループではアミノ酸残基とヘテロ分子で同程度の定量性を持ち、タンパク質をありのまま扱う量子化学計算による解析手段として密度汎関数 (DF) 法プログラム ProteinDF を開発した[1,2]。ProteinDF システムは ProteinDF をベースに分子軌道 (MO) 計算の収束を大幅に改善する擬カノニカル局在化軌道 QCLO[3] を搭載した、密度汎関数法によるタンパク質の全電子計算や分子動力学 (MD) 計算が行える統合ソフトウェアである。初めてタンパク質の全電子計算を行った当時は計算時間も桁違いで、DF 計算の収束も専門家ですら困難を極めたが[2]、様々な工夫により実用性が大幅に向上した。今回公開する ProteinDF システムを用いることによって、この最先端の計算がどなたでもかなり容易に達成できるだろう。

◇ プログラムの概要

ProteinDF システムは5つのサブシステムから構成されている。なお、タンパク質波動関数データベースは一部 CD により別途公開済みである[4]。プロテイン・エディタ (ProteinEditor) : 本システムの実行・制御・解析に渡る全機能を統括する環境を用意した。グラフィカルなエディタ機能を持つワークベンチ、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 (OpenGL)、およびそのグラフィカル・ユーザ・インターフェース (GUI、Windows 対応) で構成されている。自動計算法 (QCLO) : タンパク質の密度汎関数法による全電子波動関数の計算を達成するには、専門家が数々の試行錯誤をしなければならなかったが、前述の通り、当グループでは QCLO を使用して安全かつほぼ自動的にタンパク質の全電子計算が達成できる方法を発案した[3]。本公開版では、さらなる改良を施した QCLO 法を組み込んだ (Fig.1)。構造最適化・*ab initio* MD (ProteinMD) : 新たにスクラッチ

から MD プログラムを開発した。これまでの AMBER によるデータベース力場以外にも、電子状態の計算には ProteinDF を用いて力を解析的に計算し、核の運動には ProteinMD を用いてタンパク質の *ab initio* MD および *ab initio* 構造最適化を実行できる。大規模タンパク質計算エンジン (ProteinDF) : 300 残基 (5000 原子、20000 電子、30000 軌道) 規模の世界最大の全電子計算が達成できる (Fig.2)。1.3GHz Itanium2 の 32CPU (理論ピーク性能 166.4GFlops) を使用した実測では 1ヶ月ほどで計算可能。100 残基クラスの計算ならば試行錯誤無しで同システムを用いて半日。タンパク質波動関数データベース (ProteinQR) : タンパク質の代表的な構造や、同じタンパク質に対して一連の遺伝子工学的な操作がなされた構造をあわせて数十種類選出し、これらの全電子計算を系統的に実行している。得られた結果を収集し、データベースを構築した。

タンパク質の電子状態計算が実用化の域に突入した。生命現象の基礎過程解明といった研究のみならず、医薬品、触媒、遺伝子治療、遺伝子改変、環境有害物質の解析等に有用なツールを充実させて、次年度公開する予定である。

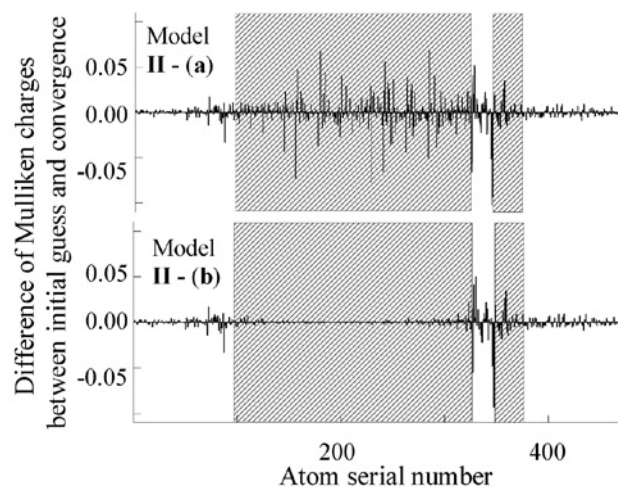


Fig.1 SCF の初期値と収束値のマリケン電荷の差 (上)従来法、(下)本方法

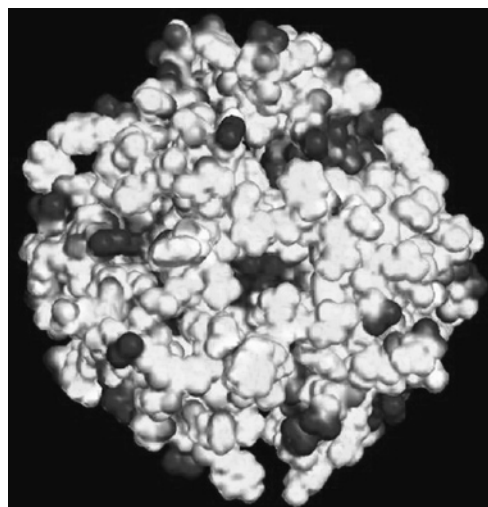


Fig.2 306 残基インスリン 6 量体の静電ポテンシャル

◇タンパク質立体構造解析ツール

ProteinEditor、ProteinMD の各サブシステムは単独で使用可能である。これらを使用して、タンパク質立体構造の解析を高品位に行うことが出来る。

ProteinEditor は ProteinDF を実行するための統合環境としての機能以外にも、PDB[5]が配信する標準タンパク質立体構造データを表示し、解析する機能が備わっている。例えば、Ramachandran プロット機能、ボンド長・ボンド角・二面角の分布表示機能、異常原子間距離表示機能、アミノ酸置換 (ミューテーション) 機能、簡易 MD 組み込みによる局所構造緩和機能、MD アニメーション表示機能など。処々の機能は様々な 3次元グラフィックスと連動しているため、異常な構造がチェックしやすい。

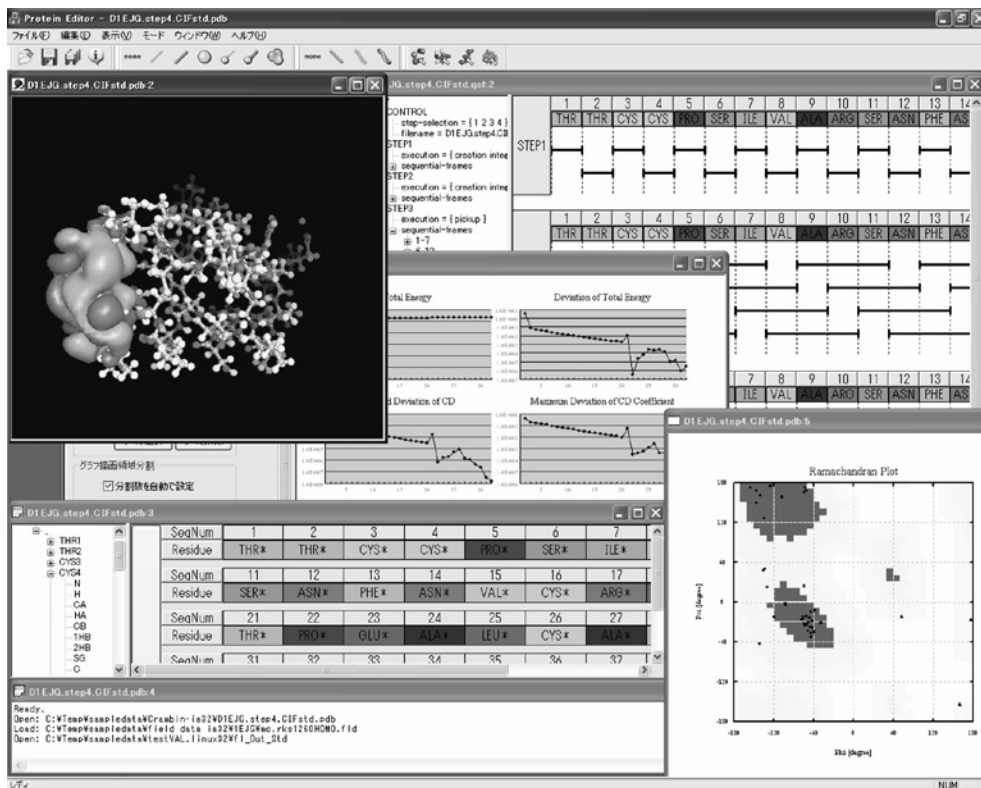


Fig.3 ProteinEditor のスナップショット

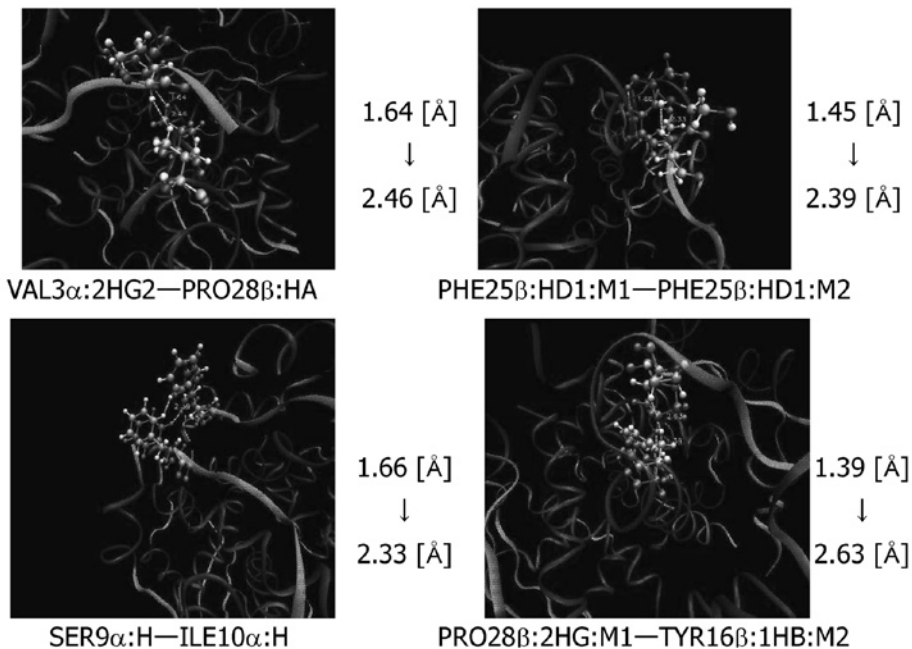


Fig.4 GB 近似による構造最適化で異常近接水素対距離が緩和 (インスリン 6 量体)

ProteinMD はタンパク質構造のリファインメントに使用できる。一般に、PDB のタンパク質立体構造は実験解析法の原理上歪みを持っているもの、得られたデータからのリファインメントが不十分なものが多数存在し、これをもとにシミュレーションを開始するのは危険である。ProteinMD はそのための前処理ツールとして有用である。例えば、PDB 立体構造にはボンドを作らない水素-水素間の距離がファンデアワールス半径の和よりも極めて小さいものが多数存在する。これを緩和する方法は溶媒分子をタンパク質の周辺に頭配置して通常の MD を行うものが一般的だが、この方法ではそもそもの PDB 構造（すなわち実験構造）から乖離してしまう場合が多々見られる。そこで、ProteinMD では Generalized Born 近似による溶媒の取り扱い方法を実装した。これにより、タンパク質構造の差異の RMSD が 1 Å 程度で、歪みを緩和させることができるようになった (Fig.4)。

タンパク質量子化学計算はまだ敷居が高いと思われる方、タンパク質の構造をハンドルするツールが必要な方にこれらのサブシステムをお勧めする。

◇参考文献

- [1] 柏木 浩、佐藤文俊、吉廣 保、稲葉 亨、西川宜孝、小池 聡；“タンパク質量子化学計算－ProteinDF の夢と実現－”、柏木 浩、佐藤文俊監修、アドバンスソフト (2004)、ISBN4-9902143-0-7.
- [2] F. Sato, T. Yoshihiro, M. Era, H. Kashiwagi; Chem. Phys. Lett., 341 (2001) 645.
- [3] H. Kashiwagi, H. Iwai, K. Tokieda, M. Era, T. Sumita, T. Yoshihiro, F. Sato; Mol. Phys., 101 (2003) 81.
- [4] 西村民男、稲葉 亨、小池 聡、平野敏行、田原才静、吉廣 保、西村康幸、西野典子、佐藤文俊、柏木 浩；“ProteinDF によるタンパク質量子化学計算事例集”、アドバンスソフト (2005)、ISBN4-9902143-4-X.
- [5] Protein Data Bank; <http://www.rcsb.org/>