



文部科学省次世代IT基盤構築のための研究開発  
「革新的シミュレーションソフトウェアの研究開発」

創薦・バイオ新基盤技術開発へ向けた

タンパク質反応全電子シミュレーション・システム(ProteinDFシステム)

Research and Development for Next-generation Information Technology of MEXT,

“Revolutionary Simulation Software”

『All-Electron Based Simulation System for Protein Reactions;

The New Infrastructural Bio-Medicinal Technology—ProteinDF System—』

東京大学生産技術研究所 計算科学技術連携研究センター

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、小池聰、西村民男、佐藤昌宏、小池秀耀、柏木浩

作成年月 2007年12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 と、それを中核とする ProteinDF システムを開発している[1,2]。

ProteinDF システムは ProteinDF の他、4つのソフトウェアから構成されている。擬力ノニカル局在化軌道 QCLO[3]法に基づきタンパク質分子軌道（MO）計算の収束を大幅に改善・自動化するソルバプログラム、ProteinDF によって得られた波動関数のデータベース ProteinQR、タンパク質の分子動力学（MD）計算が行える ProteinMD、そしてグラフィカル・ユーザー・インターフェース（GUI）を持ち、これらソフトウェアを統括する ProteinEditor 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]、統合環境の整備やソルバの性能向上といった様々な工夫により、実用性がさらに増した。ProteinDF システムを用いることによって、この最先端の計算が容易に達成できるだろう。

### ◇ プログラムの概要

**タンパク質 DF 計算エンジン (ProteinDF) :** ガウス型基底関数を用いた、RI (Resolution of Identity) 法に基づくタンパク質のための DF 法プログラムである。公開版では以下の機能を追加している。一つは高度な分子積分ルーチン並列化と行列演算ライブラリ ScaLAPACK の活用による並列性能の向上である[4]。行列を分散配置することにより、メモリ非共有型の並列クラスタでも大規模タンパク質が計算可能となった。また、Anderson 法による新たな収束加速法も実装した。これらにより 300 残基規模タンパク質の全電子計算を容易に達成することができた。

さらに、SVWNなどの局所密度近似汎関数に加え、B3LYPハイブリッド汎関数を実装している(Figure 1)。B3LYPは数多く実験結果と比較が行われ、その高精度の結果から一般化学分子計算の業界標準となっている汎関数である。局所密度近似汎関数よりも高い計算コストが要求されるが、B3LYP汎関数によるタンパク質の全電子計算が可能となった意義は大きい。

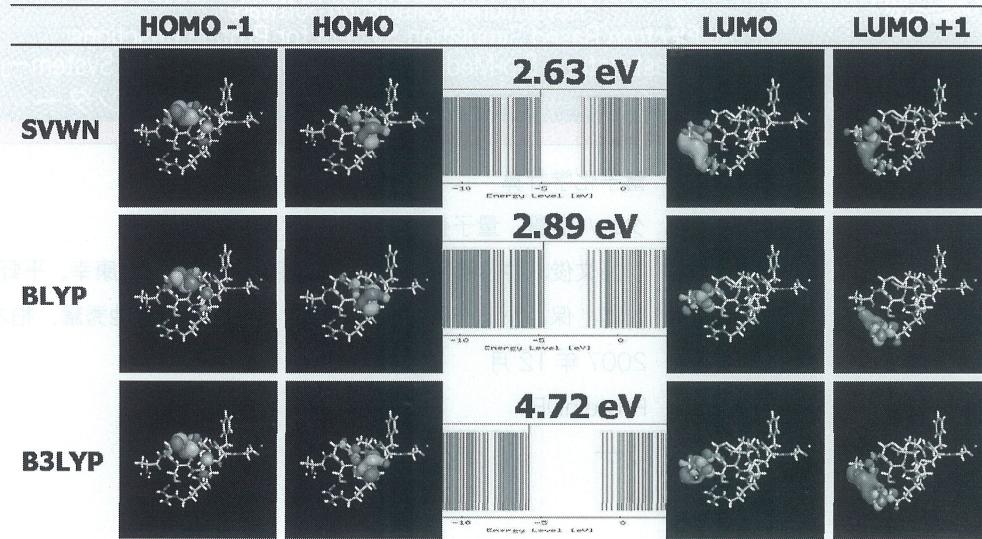


Figure 1. 各汎関数における分子軌道の比較（モデル 1FUL での例）

**擬力ノニカル局在化軌道 (QCLO) 法**：タンパク質は複雑な立体構造を形成する巨大分子であるため、その全電子計算を達成するには、専門家による試行錯誤が必要であった。これまでに、MO 計算の収束を大幅に改善する QCLO 法を開発し、多くのタンパク質全電子計算を達成することができた。タンパク質はアミノ酸から構成される一本鎖であるというこれまでの考え方方に加え、タンパク質の複雑な立体構造に注目した解法を考案した[5]。この手法により、複雑な構造のタンパク質においても初期波動関数は大幅に改善され、さらに安全かつ効率よくタンパク質全電子計算を行うことが可能となった。また、取り込んだ相互作用を計算科学的に評価する手法を開発し、計算前に理想的なフレーム分割方法を調べることが可能となった。

**古典分子動力学計算 (ProteinMD)**：分子動力学法および分子構造最適化法のシミュレーションを担う。タンパク質などの生体分子シミュレーションのパッケージとして普及している AMBER の分子力場データを利用することができます。構造変化データやエネルギーなどの時系列データの出力ファイルはそのパッケージのフォーマットに準拠している。ProteinMD は ProteinDF システムにおけるタンパク質電子状態計算の前処理の一躍を担う。計算パフォーマンスおよびソフトウェアの簡便性を考慮して、計算対象となる系の境界を自由境界条件と水分子球状境界条件 (Cap) に限定している。また、前処理として必要とされる、ユーザーが指定する一部の原子のみ可動となる部分シミュレーションが可能である。タンパク質を取り巻く溶媒分子の構造決定には古典分子動力学法が有効である。TIP3P などの明示的なモデルの取り扱いはもちろんのこと、暗示的水溶媒効果の取り扱いも可能である。分子動力学法における運動方程式の時間積分はベルレ法を含むシンプレティック差分法を基本としている。アンサ

ンブルの指定は NVE 一定と NTV 一定を選択することができる。温度制御のアルゴリズムとして拘束法と拡張法を装備している。

タンパク質の自由エネルギー計算において、長時間の分子動力学法が必要とされる。先のシンプレクティック差分法により、長時間のシミュレーションに対するエネルギーなどの保存量の安定性が得られている。また、アンブレラサンプリング法などに必要とされる、正規の力場に対する補助ポテンシャルを導入することができる。今回のソフトウェア公開に向けて、計算パフォーマンスを考慮してデータ構造からソースコードを書き改めた。従来の機能に加え、相互作用高速計算専用ハードウェア(MDGRAPE-3)との効率のよいデータの受け渡しが可能となった。

**タンパク質波動関数データベース (ProteinQR)**：代表的なタンパク質や、あるタンパク質に対するアミノ酸置換が行われた構造を合わせて、100種類程度のタンパク質の全電子計算を系統的に実施した。得られた結果を収集し、継続的にデータベースを構築中である (Table 1)。これらの一部は既に CD に収め、公開済みである[6]。

Table 1. タンパク質波動関数データベース登録リスト (抜粋)

タンパク質	残基数	原子数	軌道数
ユビキチン (1C3T)	76	1231	6685
クランビン (1EJG)	46	642	3627
インターロイキン (1BBN)	133	2157	11909
サイトカイン受容体 (1C8P)	102	1660	9329
コクリン LCCL ドメイン (1JBI)	100	1475	8196
インスリン単量体 (1A10)	51	786	4461
インスリン 6 量体 (1A10)	306	4716	26766
クランビン類似体 (1CRN)	46	642	3627
RGD ペプチド (1FUV)	136	543	2606

**グラフィカル・ユーザー・インターフェース (ProteinEditor)**：量子化学計算に関わる研究者のみならず、生物物理学や生化学の研究者を想定して開発され、ProteinDF による複雑なタンパク質の全電子計算達成をサポートする統合環境としての機能をほぼ全て装備した。Windows に対応し、分子グラフィックスは OpenGL を用いて描画している。システムを統括する機能に加えてグラフィカルな編集機能、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、およびその GUI で構成されている。具体的には、Ramachandran プロット機能、結合長・結合角・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 (ミューターション) 機能、最急降下法 (Steepest Descent Method) による簡易部分構造最適化機能、MD アニメーション表示機能などが組み込まれている (Figure 2)。

さらに、局所的な分子構造編集のためのタンパク質モデリング支援機能も追加した。タンパク質を表示した後、異常原子間距離を発見した場所などを選択し、新たな編集ウィンドウを開くことで局所的な強制修正を加えることが可能である。また生産技術研究所の選定研究の支援を受け、タンパク質モデル

ング支援機能に Virtual Reality (VR) 技術を取り入れた[7]。この VR 技術により、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正確に把握し、インタラクティブに構造（原子、アミノ酸残基、基質、配位子）の編集を支援できる。従来、構造歪み除去は MD による構造最適化を行ってきたが、側鎖や官能基が自由に回転できない局所解に落ち込み、緩和できない構造歪みが残る問題があった。

ProteinEditor 特別版を使用することによって視覚・触覚に訴えた直観的な編集作業ができるため、このような構造から強制的に局所壁を乗り越えさせ、ユーザーが直接適切な構造を指定することができる。

ProteinEditor は、タンパク質量子化学計算のみならず、一般的なタンパク質分子構造編集環境としても使用できる。

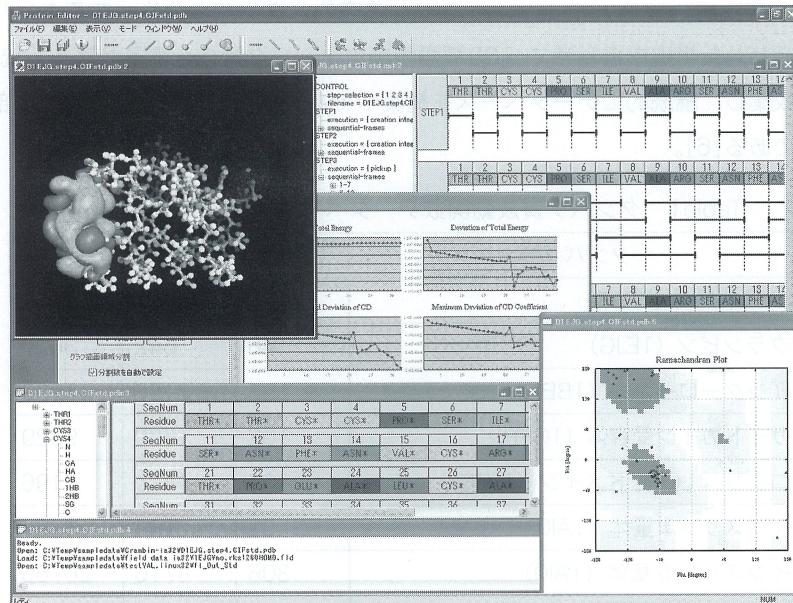


Figure 2. ProteinEditor のスナップショット

### ◇参考文献

- [1] 柏木 浩、佐藤文俊、吉廣 保、稻葉 亨、西川宜孝、小池 聰；“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] T. Inaba, F. Sato; *J. Comput. Chem.*, 28 (2007) 984.
- [5] N. Nishino-Uemura, T. Hirano, F. Sato, *J. Chem. Phys.*, in press.
- [6] 西村民男、稻葉 亨、小池 聰、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.
- [7] 西村康幸、吉廣 保、佐藤文俊；生産研究、59 (2007) 93.