



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、平野敏行、阿部敏彦、下向智美、谷村景貴、恒川直樹、松田潤一
作成年月日	2012年6月
コード名	ProteinDF
使用言語	C++

#### ◇ ProteinDF System 2012 概要

当グループは、密度汎関数法でタンパク質分子の全電子カノニカル波動関数を計算する ProteinDF システムを開発している[1、2]。計算法の汎用性から、タンパク質に限らず様々な巨大分子の電子状態を解析できるため、ナノスケールものづくり設計に貢献できる。本公開ソフトウェアは、実用的なプリポスト処理機能を追加しつつ、本システムをバイオ・ナノ分子特性シミュレーションシステムとして発展させた2012年度版である。本ソフトウェアはダウンロードして無料で利用することができる[3]。

#### ◇ 注目すべき新機能・開発項目

さまざまな並列計算機に対応した超並列密度汎関数計算エンジンの改良と機能追加：ProteinDFは大規模分子の全電子計算を限られたメモリ上で遂行するために、一度に計算する大規模行列を最低限に絞り、その時点で必要ではない行列はディスク領域に退避させる方法を採用している。1000残基クラスのタンパク質全電子計算の場合、行列のサイズは約80GBにも及ぶ。ディスクを用いる方法は非効率的であるが、計算ノードの搭載メモリ量が限られている現在において、大規模分子の量子化学計算を行う次善の策といえる。京速コンピュータ「京」をはじめとするいくつかの大規模並列計算機では、共有ファイルシステムではなくノード毎にストレージが接続されているが、こうした計算機システムでも動作できるように対応した。ディスク領域に退避させるもう一つのメリットとして、リスタートが可能である点も挙げられる。リスタート機能をはじめ、さまざまな利便性の改善も行った。

大規模分子の密度汎関数計算をより高速に行うために、分子積分・数値積分演算の高速化と高度並列化を施している。分子積分エンジン、特に計算律速となる4中心2電子積分の計算エンジンを見直し、高速化を図ると共に、微分計算も高速に行えるように機能追加した。交換相関項の数値積分では、行列

積を有効に使えるように書き換えた。これにより計算機に最適化された BLAS ライブラリを利用することができ、計算効率を向上させることができる。

いくつもの行列要素を参照・計算・格納しなければならない分子積分・数値積分において、数千規模のノードに巨大大域行列を分散保持させた場合はかなり非効率になる。分子積分・数値積分の超並列計算を効率良く行うため、プロセスの通信待ち時間を極力減らした、新規並列計算方法を導入した。分子積分においては、計算に必要な密度行列を各プロセスに分散保持し、そのローカル行列に対して各プロセスが並行して分子積分を計算できるように実装した。数値積分においては、まず各プロセスにおいてローカルに保有している密度行列を元に全グリッドの電子密度ならびに密度勾配を求め、全通信により集計した後、交換相関項をローカルに算出するように改良した。これにより、ロードバランスが保つことは難しいが、通信待ちを減らして分子軌道計算を行うことに成功し、大規模並列計算機における並列密度汎関数計算ルーチンが整った。

**部分構造最適化機能**：高速なエネルギー勾配計算機能を利用して様々な手法による構造最適化計算を実装した。化学物質の性質や反応を定量的に解析するためには、ポテンシャルエネルギー曲面 (Potential Energy Surface : PES) の情報が不可欠である。系の自由度が極めて多いタンパク質のポテンシャルエネルギーをすべての原子核配置について求めるのは事実上不可能であるが、実際には、活性中心回りの構造において得られれば十分実用的である。任意の構造から出発して最小値に必ず到達できる完全なアルゴリズムは未だ存在しないし、巨大な系ではどの方法が有効であるのか未知である。そこで、本プログラムでは、複数の極小値探索法を実装した。エネルギー勾配を利用する方法として、最急降下法、共役勾配法、準ニュートン法を実装した。これらの方法は量子化学計算で一般的によく使われている方法である。エネルギー勾配を利用しない方法としては、修正 Powell 法と滑降シプレックス法を実装した。これらの方法を組み合わせて、タンパク質全電子計算を行いながら、活性中心回りの部分構造最適化計算が実行できる。今後、遷移状態探索法も実装する予定である。

**NMR 計算機能**：核磁気共鳴 (NMR) 分光法は分子構造や電子配置などの詳細情報が得られるため、低分子のみならず、生体分子やナノ分子の重要な観測手段の一つである。そのため、理論計算による結果解析や予測の必要性が高まっている。タンパク質などのナノスケールの分子軌道計算により得られた大規模波動関数から化学シフト計算を実現することを目的として、本バージョンでは Hartree-Fock (HF) 法に基づき、スクラッチからプログラムの開発を行った。磁場の摂動計算を行う上で、ゲージ依存性の問題を避けるために、Gauge Including Atomic Orbitals (GIAO) 法を採用した。GIAO 法では原子ごとにベクトルポテンシャルを定めるゲージの原点を取ることにより、基底関数のゲージ依存性を取り除くため、高精度の計算が可能である。HF 法の摂動計算に GIAO 法を適用する場合に必要となる、種々の分子積分の実装を行った上で、磁気遮蔽定数の反磁性項および常磁性項の計算プログラムを開発した。また常磁性項の計算で必要となる、CPHF 方程式の安定解法を新たに考案・実装した。得られた反磁性項と常磁性項から、最終的に核磁気遮蔽テンソルが求められる。上記の種々の分子積分は高速化中であり、現在のところ、30 残基程度のペプチド鎖 NMR 計算が可能となった。

統括 GUI：全電子計算をサポートする高品位 GUI、ProteinEditor のさらなる機能強化と、新機能の追加を行った。

(a) 大規模タンパク質のための高速化・省メモリ化の実現

光合成反応中心タンパク質などの大規模な分子系を手軽に扱える高品質な GUI は必要不可欠である。ProteinEditor 内部クラスのリファクタリングや、C#・OpenGL による実装により、省メモリ化、処理時間のさらなる高速化、より大規模な分子の計算結果の高解像度化を実現した。

(b) 柔軟な拡張開発のためのプラグイン機能の開発

タンパク質の全電子計算のためのプリポストとして、様々な表現方法が可能であるが、より多様な物理量の表現、解析や分子設計のために、ProteinEditor に自由に拡張開発を行うためのプラグイン機能を追加した。この機能により、ユーザーは ProteinEditor 内部に実装されたインタフェースを通じて ProteinEditor 内部のパラメータに自由にアクセスし、カスタマイズすることができる。プラグインとして扱える外部プログラムの実装言語は、C# と IronPython に対応している。

(c) 表示や動画作成に関する機能強化

産業界が望むレベルの高品位かつ実践的なプリポストに近づけるためには、ProteinEditor の各種機能の強化が不可欠である。本バージョンにおける機能追加として、コマンド処理機構や、動画作成支援機能、Arbitrary slicer の実装、原子の Mulliken 電荷表現、等値面表示の半透明化などを行った。

(ア) コマンド処理機構の追加

ユーザーコントロールによって発行されるイベントを総括して、コマンドとして処理する独立した機構を作成した。これにより、ユーザーコントロールによるイベントと同様の処理が、テキストによる命令でも実行可能となる。

(イ) 動画作成支援機能

上記のコマンド処理機構を利用して作成した。カメラ位置の設定、表示の on/off、画像ファイル出力などのコマンドを処理する。外部で作成したコマンド列ファイルを実行することにより、バッチ処理が可能である。Fig. 1 は、この処理フローを表している。時系列データを扱うために、ProteinEditor 内部データ表現のフォーマット拡張も行った。

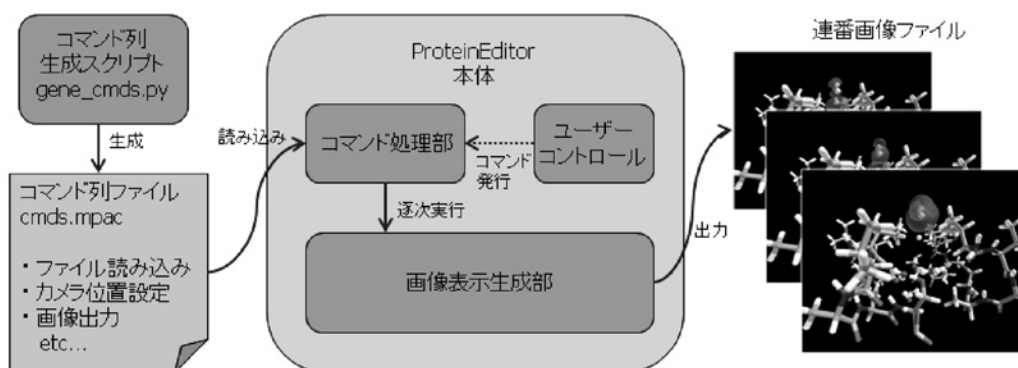


Fig. 1 : The process of making sequential image files

## (ウ) 表示機能の強化

Arbitrary slicer は、計算結果の物理量を面上の等高線マップとして表示するとき、その表示面をマウス操作でインタラクティブに決定できる機能である。この機能により、MbCO ヘム周辺の電子密度を表示したものが Fig. 2 である。Fig. 3 は Mulliken 電荷を分子構造にマッピングした例、および静電ポテンシャルを等電子密度面にマッピングした面を半透明表示した例である。分子構造と物理量の関係性を同時に解析しやすくなるなどの利点がある。

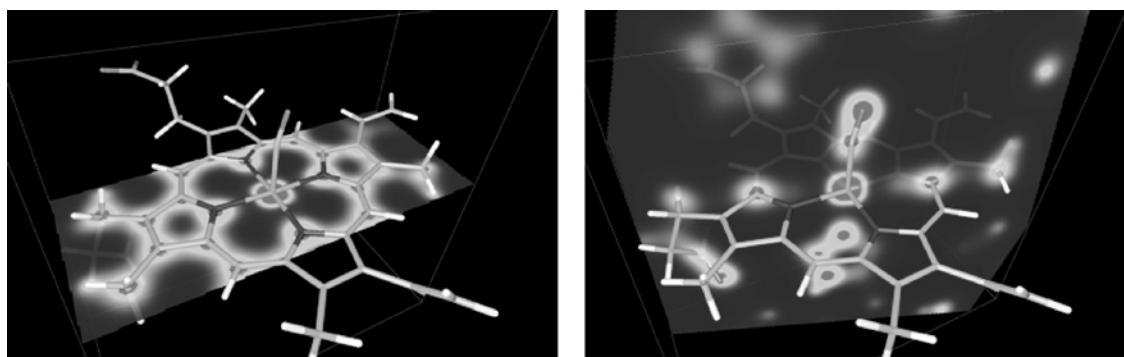


Fig. 2 : The contour plane of electron density around MbCO heme

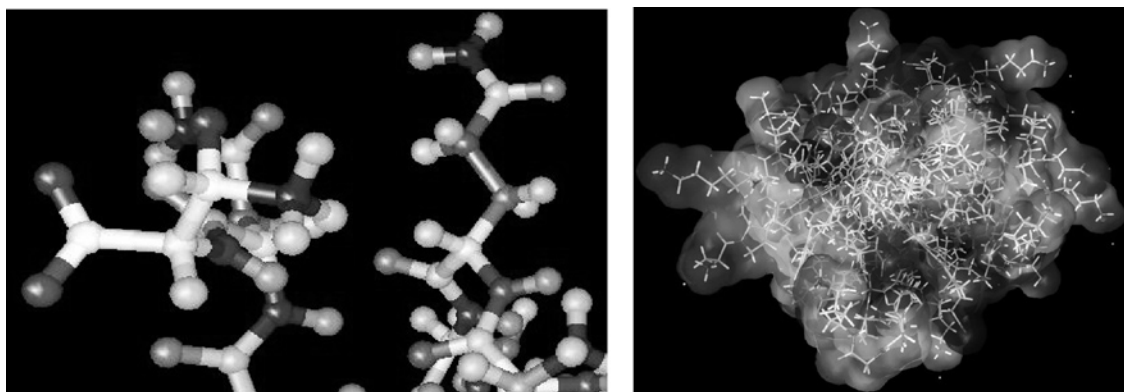


Fig. 3 : (left) Molecular structure mapped by Mulliken charges  
(right) Electrostatic potential of cytochrome  $c_3$  drawn by semi-transparent surface

## ◇参考文献

- [1] 佐藤文俊、恒川直樹、吉廣 保、平野敏行、井原直樹、柏木 浩：“タンパク質密度汎関数法”、柏木浩監修、森北出版（2008）
- [2] 上村(西野)典子、佐藤文俊、恒川直樹、西村康幸、平野敏行、吉廣 保、甘利真司、加藤昭史、小林将人、田中成典、中野達也、福澤 薫、望月祐志：“プログラムで実践する生体分子量子化学計算-ProteinDF/ABINIT-MP の基礎と応用”、佐藤文俊、中野達也、望月祐志編、森北出版（2008）
- [3] <http://www.ciss.iis.u-tokyo.ac.jp/>

（執筆責任者：佐藤文俊）