

タンパク質全電子計算プログラム ProteinDF を基盤にした 次世代量子化学計算システムの開発計画

佐藤 文俊[†], 上野 哲哉[†], 田原 才静[†], 恒川 直樹[†], 牟田 元[†], 吉廣 保[†], 柏木 浩^{†*},
稲葉 亨[‡], 村松 伸哉[‡], 西川 宜孝[‡], 小池 秀耀[‡]
東大生産技術研[†], アドバンスソフト[‡], 九工大^{*}

1. はじめに

タンパク質は体温という温和な条件下で反応が進行する、わずか 0.1eV 程度のエネルギー変化で機能する精密分子機械である。そのため、タンパク質の機能を理論的に明らかにするためには、量子論に基づく解析方法を編み出す必要がある。当グループはタンパク質の全電子計算が可能な密度汎関数法プログラム ProteinDF を開発し、タンパク質の量子化学計算に力を注いできた。本発表ではこれまでの結果、ならびにこれを基盤とした発展型システムの開発計画について報告する。

2. ProteinDF とヘムタンパク質シトクロム *c* の全電子計算

近年、タンパク質の本格的な量子化学計算を目指した研究が登場し、いくつかの純粋なペプチド鎖の計算が行われるようになった。この中で、発表者らは特にヘムタンパク質や光合成反応中心のように金属錯体がタンパク質の活性中心として働き、さまざまな反応性を示す金属タンパク質に興味があり、電子レベルでそのメカニズムを明らかにすることを目標としている。このような系における定量的な量子化学計算では電子相関効果が重要である。そこで発表者らは *ab initio* Hartree-Fock 法と同程度の計算量で電子相関を取り込むことができる Kohn-Sham-Roothaan 方程式に基づく密度汎関数法に着目し、プログラム ProteinDF を開発した¹⁾。発表者らは、これをワークステーションクラス上に用いて、ヘムタンパク質の1つであるシトクロム *c* の全電子計算を 2 ヶ月かけて達成した²⁾。これは密度汎関数法による世界初の金属タンパク質全電子計算である。

計算分子は馬心筋シトクロム *c* で、1つの *c* 型ヘム (Fe プロトポルフィリン) を持つ 104 残基の

典型的なヘムタンパク質である。生体内では呼吸鎖中でヘム Fe の酸化還元を利用してタンパク質間の電子授受を担っている。原子数、電子数、使用した軌道および補助基底関数の総数はそれぞれ 1,738、6,586、9,600 および 17,578 であった。図 1 のグラフィックスは最高占有軌道 HOMO を様々な等値面の値で描いたもので、等値面の値は左上から順に、 ± 0.00005 , ± 0.0005 , ± 0.005 , ± 0.05 である。A はスケールを倍にして描いている。HOMO は d_{xy} を主成分とする軌道であるが、Fe 原子の $3d$ 軌道に比べて複雑で、比較にならないほど広がっており、タンパク質分子表面をはみ出したところでもかなりの値を持っている。この特徴がシトクロム *c* の電子授受機構に重要な役割を果たすと考えられる。このように、タンパク質の軌道は小分子のものとは異なる観点から注意深く観察しなければならない。

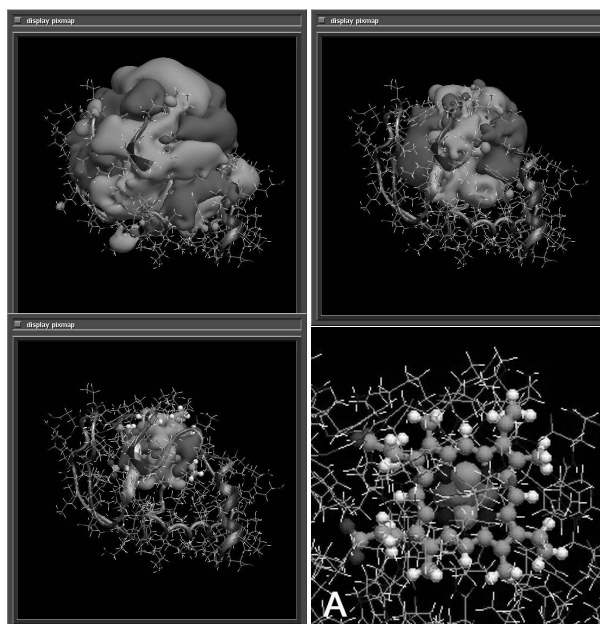


図1 シトクロム *c* の HOMO

Development Project of the Next-Generation Quantum Chemistry Calculation System Based on ProteinDF for All-Electron Calculations on Proteins

[†]Fumitoshi Sato, Tetsuya Ueno, Saisei Tahara, Naoki Tsunekawa, Hajime Muta, Tamotsu Yoshihiro, Hiroshi Kashiwagi; Institute of Industrial Science, University of Tokyo

[‡]Toru Inaba, Shinya Muramatsu, Nobutaka Nishikawa, Hideaki Koike; Advancesoft

*Kyushu Institute of Technology

3. 次世代量子化学計算システム

今でこそコスト高なタンパク質の量子化学計算も、今後の計算機の発展を考慮に入れば、まもなく誰でも簡単に実行できる時代が到来することは明らかである。これは、これまで古典論による解析が主であったタンパク質の研究に一大革命をもたらすであろう。

このような時代に先駆けて、ソフトウェアを整備することが急務の課題となっている。そこで、発表者らは、図3で示された次世代量子化学計算システムを5ヶ年計画で開発する。タンパク質の全電子計算用に開発されたソフトウェアProteinDFをベースに、反応解析に必要な新機能を組み込み、使いやすいインターフェースを提供することによって、精密で実用的な、タンパク質のシミュレーションシステムを研究開発する。いくつかの重要なタンパク質の波動関数データベースもあわせて公開し、本システムの普及を目指す。

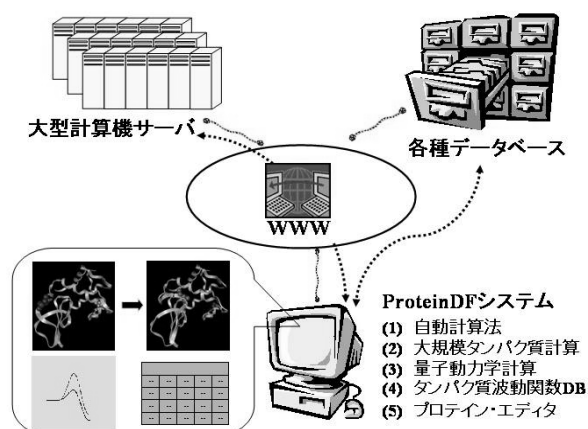


図3 次世代量子化学計算システム概念図

(1) 自動計算法

方程式は自己無撞着 (SCF) 法で解かれるが、タンパク質の電子状態は複雑で、大変よい初期値を用いないと全電子計算が達成できない。本グループは一種の局在化軌道 (QCLO) から初期値を作成する方法を開発し、誤差を従来の数十分の一にすることに成功した³⁾。これを自動的に実行する機能を本システムに組み込んで、ほぼ全ての全電子計算を自動計算する仕組みを提供する。

(2) 大規模タンパク質計算

これまでの研究結果から、倍精度演算で100,000軌道の全電子計算が可能であると見積もられている。これは1,000残基規模のタンパク質に相当し、大部分の重要なタンパク質が計算対象に含まれることになる。本システムを超大

規模計算・超大型計算機サーバにも対応させる。

(3) 量子力学計算

全電子計算を解析した結果、同じアミノ酸残基でもタンパク質中の三次元配置によって、電荷が0.1以上も異なることが判明した。この差は10遠方で0.1eVもの差を生む。タンパク質反応解析のために波動関数から得られた情報を基に分子力学計算を行う方法を研究開発する。

(4) タンパク質波動関数データベース

100残基規模のいくつかの重要なタンパク質の全電子計算を系統的に実施し、タンパク質の高精度の電子状態データベースを構築する。これらの情報はタンパク質の実験結果の理解のみならず、タンパク質の新たな物理化学的知見の獲得、新規タンパク質の設計、機能および反応性の評価に大いに役立つ。

(5) プロテイン・エディタ

本グループはすでに富士総研と共同でタンパク質全電子半自動計算のためのインターフェース (シナリオ・エディタ) を開発し、ユーザ会で公開した実績がある⁴⁾。本研究開発ではさらに、タンパク質の反応解析に役立つツールを作り込み、全機能を容易かつ直感的に操作することができる統括環境へとバージョンアップさせ、ユーザに提供する。

本システムは、ポストゲノム時代のバイオおよびナノテクノロジー研究に役立つ実用的なツールとなるものと期待できる。

なお、本学会ではこれら5つの研究について詳細を発表するので、併せてご聴講願いたい⁵⁾。

謝 辞

本研究は文部科学省リサーチ・レボリューション計画 (RR2002) ITプログラム「戦略的基盤ソフトウェアの開発」の支援を受けている。

参 考 文 献

- 1) F. Sato, Y. Shigemitsu, I. Okazaki, S. Yahiro, M. Fukue, S. Kozuru, H. Kashiwagi: *Int. J. Quant. Chem.*, **63** (1997) 245-256.
- 2) F. Sato, T. Yoshihiro, M. Era, H. Kashiwagi: *Chem. Phys. Lett.*, **341** (2001) 645-651.
- 3) H. Kashiwagi, H. Iwai, K. Tokieda, M. Era, T. Sumita, T. Yoshihiro, F. Sato: *J. Mol. Phys.* (2003) in press.
- 4) <http://www.fuji-ric.co.jp/ccse/ProteinDF/>
- 5) 4C-4, 4C-5, 4F-5, 5F-1, 5F-3