

3D05

密度汎関数法によるタンパク質全電子計算システム ProteinDF の展開

(東大情報基盤¹, 東大生研², 日立機械³, アドバンスソフト⁴)

○ 佐藤文俊^{1,2}, 井原直樹², 恒川直樹², 西野典子², 西村康幸², 平野敏行², 吉廣保², 稲葉亨³, 小池聡⁴, 佐藤昌宏⁴, 下堂靖代⁴, 西村民男⁴, 小池秀耀⁴, 柏木浩^{2,4}

【序】当グループでは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[1]、平成 17 年 7 月より文科省「戦略的革新シミュレーションソフトウェアの開発」プロジェクトでこれを基にした統合システムをさらに発展させている[2]。高コストで複雑なタンパク質の全電子計算も、100 残基クラスのタンパク質であれば実用的な時間でそれほど試行錯誤することなく達成できるまでに成長した[3]。本発表では新しいプロジェクトの全体像ならびに研究内容と現状、今後の計画を報告する。

【新プロジェクトの全体像】新プロジェクトの目的は、これまで開発してきた ProteinDF をベースとした統合システムをさらに発展させて、タンパク質反応の全電子シミュレーションシステムを開発することである。これにより機能や生命現象の基礎過程解明といった基礎研究のみならず、古典論による手法が主であった創薬、生体触媒、生体分子素子、環境物質などへの応用研究に信頼性の高い基盤技術を提供することも目標としている。

以下の 5 つの研究開発項目を統合し、システムを構築する予定である。A. 密度汎関数法による超大規模タンパク質全電子計算エンジン、B. タンパク質の構造最適化、分子動力学、自由エネルギー計算機能、C. 密度汎関数法-配置間相互作用法による電子励起・電子移動計算機能、D. タンパク質波動関数データベース、E. シミュレーションを支援し、物性を評価・解析できる高品位 GUI。さらに、実証計算を通してシミュレーションの方法を堅固な技術へと発展させたい。

【本年度の研究内容】以下の 3 つについて述べる。

1. 密度汎関数法による超大規模タンパク質全電子計算エンジンの改良

ProteinDF エンジン本システムの本シミュレーションの基盤である。すでにスプリットバレンス基底関数系で、300 残基クラスのタンパク質の世界最大規模の全電子計算を計算する能力を有している。本プロジェクトでは 1,000 残基規模のタンパク質の密度汎関数法による全電子計算が実行できるよう ProteinDF エンジンをさらに高性能化したい。これによりタンパク質立体構造データベース PDB に登録されている総数 31,971 サブ構造うち 31,833 サブ構造の全電子計算が可能になる (2005 年 8 月 2 日現在)。文字通りほとんどすべてのタンパク質を計算対象にすることを保障する最重要課題である。これには各計算律速ルーチンの高速化とともに、1,000CPU 規模の計算機に対応できるように並列化をレベルアップする必要がある。従って、次世代の超大規模並列スーパーコンピュータの性能を大いに引き出すことも期待できる。

2. タンパク質の計算構造決定

タンパク質の全電子状態計算において計算構造の決定は重要である。一般にタンパク質の立体構造は、Protein Data Bank などのデータやそれから出発した通常の古典分子動力学法や構造最適化法のシミュレーションデータが用いられる。しかしながら、これらの方法が必ずしも化学的に

適切な構造を保障していないことが判明した。特に規模が大きくなるほどこの問題を避けることが難しい。そこでいくつかの方法を試み、得られた構造の比較および調査を行った[4]。一方、当グループは ProteinDF の統括環境として ProteinEditor を独自に開発・発展させている[5]。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 GUI (Windows 対応) で構成され、安全なシミュレーション達成をサポートする機能の開発を進めている。これをベースに上記の知見をフィードバックし、タンパク質の計算構造決定機能を VR・立体視・MD GRAPE を組み込んだ PC システムで開発している。

3. タンパク質波動関数データベースの構築

ProteinDF によるタンパク質全電子計算からタンパク質の精密な波動関数や電子密度が得られ、これらから様々な物理量が計算できる。現在のところ個人のレベルでは大量に計算できる環境にはないため、これまでプロジェクトで計算してきた結果は公開に値するデータである。その第一弾として 2 次元の画像データによるタンパク質波動関数計算事例集を出版する[6]。当グループではさらに、誰もが自由にタンパク質の解析結果を利用できるように公開したいと考えている。このような特徴を持ったデータに特化した計算結果検索サービス、取得サービス、登録サービス、データ再利用サービスを行う、タンパク質波動関数 DB を構築する方法を検討している[7]。

【今後の計画】今年度得られた研究結果を基に、近い将来 3 つの機能を開発・強化したい。

① タンパク質の本格的な構造最適化機能

特に古典法では、金属を含んだタンパク質の構造最適化や配位子・基質との結合エネルギーの見積もりが不得手である。例えば、自由度が極めて高いヘムポケットを持つ P450 (CYP) の解析には本機能が必須である。

② 自由エネルギー計算機能

量子化学計算をベースにしたエントロピー解析法を発展させる。例えば、長期的な糖尿病の治療に使用されるインスリンは自己組織化の親和力が強く、製剤中で 6 量体が安定となる。そのため薬効のある単量体に解離する過程が律速となっている。本機能はインスリン製剤の改良に役立つ可能性がある。

③ 電子励起・電子移動計算機能

DFT-CI 法、TDDFT 法によるタンパク質電子遷移・励起状態計算の実現が見込めるようになってきた。これによりスペクトルを実験値と直接比較ができるため、さらに有用性の高いシステムへと発展するものと期待できる。

【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] 戦略的革新シミュレーションソフトウェアの開発, <http://www.rss21.iis.u-tokyo.ac.jp/>.
- [3] T. Inaba, S. Tahara, N. Nisiwaka, H. Kashiwagi, F. Sato, *J Comput Chem.*, **26** (2005), 987.
- [4] 恒川直樹, 稲葉亨, 平野敏行, 吉廣保, 柏木浩, 佐藤文俊, 分子構造総合討論会, 3P106 (2005).
- [5] 西村康幸, 吉廣保, 西野典子, 佐藤文俊, 第 28 回情報化学討論会, (2005) in press.
- [6] “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) in press.
- [7] 平野敏行, 佐藤文俊, 分子構造総合討論会, 2P047 (2005).