

PDB の水素付加・アミノ酸置換・化学修飾プログラムの開発 ータンパク質の量子化学計算ソフトウェア支援のためにー

(東大生研) ○牟田 元・大槻浩幸・西村康幸・連 宗旺・吉廣 保・佐藤文俊

【序】

本研究グループはこれまでに、密度汎関数法によるタンパク質のための量子化学計算ソフトウェア ProteinDF^[1]を開発してきた。ProteinDF は巨大分子であるタンパク質のカノニカル分子軌道計算を効率よく実行するためのアルゴリズムを搭載しており、シトクロム *c* の全電子計算を世界で初めて達成した実績がある。^[2] 現在、ProteinDF を基に、タンパク質の統合的な量子化学計算シミュレーションシステムを開発している。

タンパク質のシミュレーションには、タンパク質の全原子を含んだ三次元構造が必要である。Protein Data Bank (PDB)^[3]は現在 2 万件を超えるデータを有するタンパク質構造データベースの標準である。しかし、PDB で配布されているデータはその約 85%は水素原子が欠落しており、構造ひずみなどの問題も抱えているため、そのまま理論計算に用いることはできない。水素付加などの PDB データの構造補正を包括的に支援するソフトウェアが必須である。

本研究では、このようなニーズに応えるため、PDB のデータを編集するプログラムを作成した。タンパク質の水素付加に加え、アミノ酸置換、化学修飾などの模擬実験を支援する機能も開発した。開発は C++を用いたオブジェクト指向プログラミングにより行い、他のプログラムに容易に組み込めるようクラスライブラリ形式を採用した。

【方法】

PDB データへの水素付加は、①既存の水素以外の各原子について結合データを補完し、②付加する水素原子の数と立体構造を原子ごとに判定し、③水素原子の座標を決定することで行われる。

本研究では段階①において、PDB で配布されている PDB Hetero Group Dictionary (mmCIF 形式) ファイルを用いた。このファイルには PDB に登録されているデータに含まれる全てのアミノ酸・核酸・ヘテログループの、水素原子も含めた結合情報（結合原子のペア、および結合の種類）が含まれており、原子名を照合することで容易に結合情報を補完することが可能である。

段階②は、①で補完された結合情報を各原子について評価することにより、各原子の混成を判定した。例えば、単結合を 4 本持つ炭素原子は sp^3 炭素であり、二重結合が含まれていれば sp^2 炭素である、という要領で判定を行う。これにより、各原子について足りない水素原子の数および立体構造が明らかとなる。

段階③では、②の判定結果を元に、XYZ 座標で表された PDB データに水素原子を追加する。この際、各々基準となる原子から、水素原子の基準原子に対する Z-matrix（結合長、結合角、二面角）を生成し、これを XYZ 座標に変換して水素原子座標を決定した。

アミノ酸置換および化学修飾はどちらもアミノ酸の側鎖部分の交換であり、プログラム上はほぼ同一の処理となる。側鎖の交換は XYZ 座標のままでは困難なため、対象アミノ酸の XYZ 座標を

一旦 Z-matrix に変換し、Z-matrix に対して置換操作を行った後、XYZ 座標に再変換した。

【結果】

タンパク質-ヘテログループ複合体への水素付加の結果例を Fig.1 に示した。また、100 種類の PDB データに対する水素付加テストの結果、本プログラムによる水素付加処理は、同機能を持つ他ソフトウェアに比べほぼ確実に水素付加処理を行うことが出来るという良好な結果を示した。本ソフトウェアにおいて水素付加に失敗したケースは、段階①において、その PDB に原子名の登録違いがあったことが原因であった。このように PDB Hetero Group Dictionary を用いる本方法により、PDB 内の多様な構造を持つヘテログループへの水素付加に強みを発揮したと考えられる。

また、アミノ酸置換の例を Fig.2 に、化学修飾の例を Fig.3 に示した。現在、これらの機能をさらに充実させることで、タンパク質に関する種々のシミュレーションの解析を容易に行うことができるプログラムを開発している。

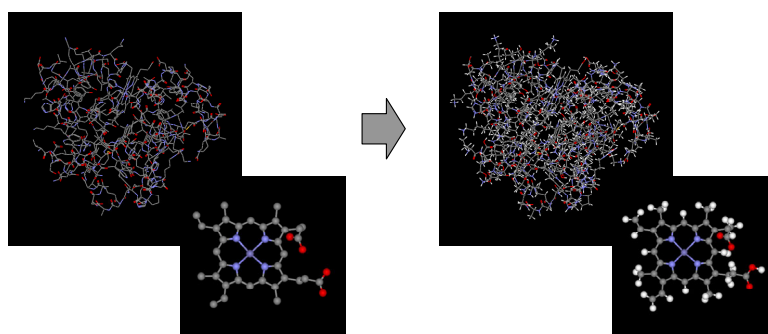


Fig.1 水素付加の例 (myoglobin, 1A6N)

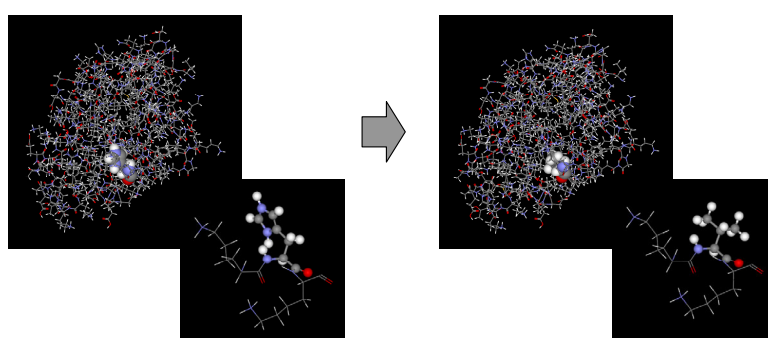


Fig.2 アミノ酸置換の例 (myoglobin, 1A6N, H97V mutation)

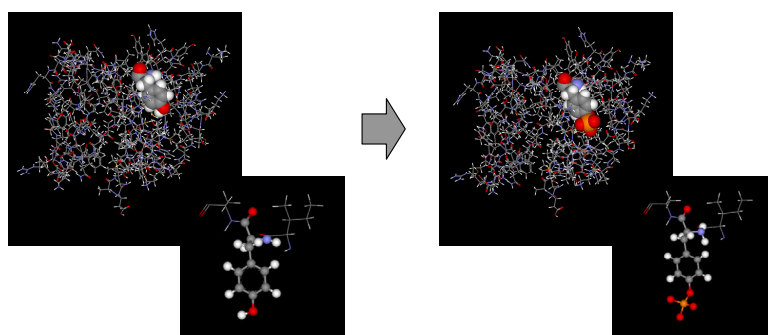


Fig.3 化学修飾の例 (SH2 domain + slam peptide, 1D4T, Tyr281(slam peptide) → Phosphotyrosine)

【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」の支援の下におこなわれました。

【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.
- [2] F. Sato, T. Yoshihiro, M. Era, H. Kashiwagi, Chem. Phys. Lett., 341 (2001) 645.
- [3] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, The Protein Data Bank, Nucleic Acids Research, 28, pp. 235-242 (2000) <http://www.rcsb.org/pdb/>