

タンパク質配座の探索アルゴリズムの開発

(豊橋技術科学大学 知識情報工学系) ○鎌倉 寿行、 後藤 仁志

【背景】

タンパク質の機能解明においてフォールディング問題は大きな障害の一つである。フォールディング過程において多数の経路が考えられること、そして、例え一つの優位な経路を見つけられるとしてもその通過点として考慮しなければならない配座があまりにも多いことが、この問題を難しくしていると言える。当研究室では、この問題に対して、ポテンシャルエネルギー曲面の極小点である配座異性体、特に低エネルギー領域に存在する配座異性体を徹底的に探索ことから解決を試みている。例えば最近では、配座空間探索法 CONFLEX を用いてアラニン 8 量体のオリゴペプチド約 1 万個の低エネルギー配座異性体を創出し、 α ヘリックス、逆巻きヘリックス、 β ストランドなどの特異的な二次構造を有する配座異性体を特定し、立体構造の類似性指標からそれらの間の変換経路を探索することを可能にした¹⁾。

このように、現時点では、十数残基程度のオリゴペプチドであれば、並列分散やグリッド技術を CONFLEX と組み合わせることによって網羅的な配座探索とそれらの変換経路を予測することは十分可能になっている。しかし、徐々に大きなペプチドやタンパク質を扱うようになり、新たにいくつかの問題が生じている。一つは、残基数が 50 を超えるあたりからタンパク質は複数の二次構造を組み合わせた複雑、かつ密な構造をとるため、小さな分子に適用してきた配座探索法で扱うことは難しくなるということがある。また、技術的な問題ではあるが、Protein Data Bank (PDB) に登録されているタンパク質のデータには分子計算を行う上で多くの不備が散見されるため、何らかの対処をしなければならないということである。

【目的】

そこで本研究では、50 残基を超えるタンパク質の配座を網羅的に創出するために、新しい配座創出アルゴリズムを構築する。また、PDB のタンパク質データ形式 (以下、「PDB フォーマット」) を読み書きでき、分子計算を行う上で適切な情報を取り出すことのできるプログラムモジュールを作成する。

【方法】

PDB に登録されているデータを分子計算に利用する際に問題になる点は、およそ次のとおりである。

- A) 結晶構造における水素の座標情報の欠落があること。また、NMR 構造も含めて、たとえ水素の座標情報があっても、それは解析ソフトが任意に発生した座標である。
- B) 異常な結合長が一部に含まれること。
- C) 5 価の炭素など、不正な価数を示す原子が含まれていること。
- D) 二重結合、S-S 結合、イオン電荷など、様々な化学的付加情報にしばしば不備があること。

本研究では、これらの問題を解決するために、まず PDB フォーマットの全てのレコードに対処できるサブプログラムモジュールを作成した。そして、上述した問題に対処する構造チェックのルーチンを追加することで、PDB フォーマットで提供されるタンパク質構造データから、直接分子計算に利用できる座標情報を容易に取り出すことのできるサブプログラムモジュールを完成させた。このサブプログラムモジュールは Fortran90 言語に準拠して作成されており、様々な分子計算プログラムでも用意に導入可能になっている。

一方、今回新たに考案したタンパク質の配座創出アルゴリズムを以下に示す。

- 1) タンパク質を特徴的な二次構造フラグメントに分割する。
- 2) フラグメント毎に CONFLEX を適用し徹底的な配座探索を行い、フラグメント配座異性体データベース

を構築する。

- 3) 各データベースからあらかじめ決めておいた基準にしたがってフラグメント配座異性体を抽出しタンパク質を再構築することで、多数の試行（サンプル）構造を創出する。
- 4) 再構築した試行構造の構造やエネルギーを調べ、明らかに不備がある、例えば、原子間距離が近すぎる場合や結合が絡まっている場合には、あらかじめ取り除く。
- 5) 全ての試行構造に構造最適化を適用し、配座チェックを経てタンパク質の配座異性体データベースを構築する。

今回の実験では、PDBに登録されている副甲状腺レセプターのN末端ペプチドフラグメント（PDB ID:1BL1）を用いた。1BL1は3つの α ヘリックスを持っており、これらを2つのループが繋いでいる。そこで今回は、5つのフラグメント（3つの α ヘリックス、2つのループ）に分けて、それぞれの配座異性体データベースをCONFLEXにより作成した（Fig1-a）。各配座異性体データベースからエネルギーの低い5つの配座異性体を抽出し、全ての組み合わせでフラグメントを繋ぎ合わせることで、3,125個の試行構造を再構築した（Fig1-b）。試行構造の中で非結合原子間が $0.4\text{\AA}$ 以下の構造はあらかじめ排除し（構造群1）、残った125個の試行構造について構造最適化を行った（構造群2）。

【結果と考察】

ループ領域の α 炭素周りの平均ねじれ角（ ϕ, ψ ）をFig.2に示す。ここで、Fig2(a)は構造群1、Fig2(b)は構造群2についてプロットした。また、それぞれに元のPDB構造に対応する値を赤い点でプロットした。この図を見ると、今回のアルゴリズムによる探索が、広域な配座空間を網羅していることが解る。しかし、元のPDB構造近傍の配座空間は網羅できていない。これは、フラグメントの配座異性体を創出する際に、溶媒効果を考慮していないため、低エネルギー領域にある配座異性体と元のPDB構造に隔たりがあるためである。したがって、より広範囲にフラグメント配座異性体を取り出すか、フラグメント配座異性体を創出する際に溶媒効果を考慮することで、元のPDB構造に近い配座異性体が低エネルギー領域に存在するような配座異性体データベースを構築する必要があるだろう。

【謝辞】

本研究は文部科学省「産官学連携イノベーション創出事業費補助金」の援助を受けて行われました。また、本研究の一部は、本研究は科学技術振興事業団「計算科学技術活用型特定研究開発事業」の援助を受けて行われました。

【参考文献】

- 1) H. Goto, T. Takahachi, K. Ohta, U. Nagashima, Nanotech 2003, Vol. 1, pp32-35.

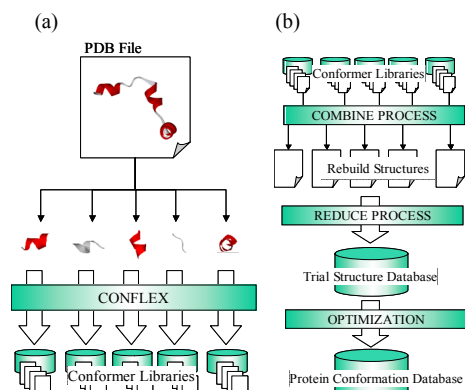


Fig1. 処理の流れ：(a) タンパク質をフラグメントに分割し、各フラグメントの配座異性体データベースを作成する (b) データベースから抽出した配座異性体を柔軟に繋ぎ合わせ、タンパク質構造を再構築する

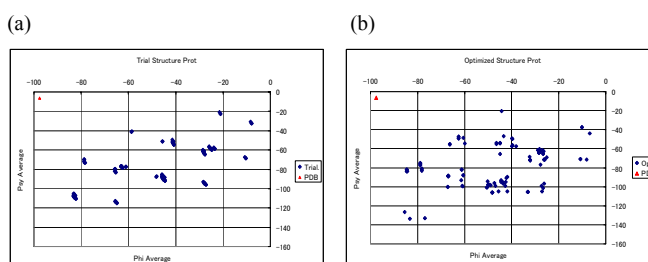


Fig2. 配座空間を示したプロット図：(a) 再構築したタンパク質の平均(ϕ, ψ)角のプロット、(b) (a)の構造最適化後の平均(ϕ, ψ)角のプロット