

タンパク質分子の標準構造の生成

5T-1

安部晴男 水戸三千秋

西日本工業大学

1. まえがき

我われは、タンパク質分子のX線結晶構造から標準構造への変換問題を取り扱っている。

分子モデリング・システムでは、最初に、コンピュータの内部に、取り扱うタンパク質分子の立体構造を生成する必要があるが、X線結晶構造解析によるデータは、分解能不足のため、分子内の原子結合距離・結合角が誤差を含んでいて、しかも水素原子が欠落している。そこで、標準的な原子結合距離・結合角をもつ原子団からなる剛体と見なせるユニット(標準ユニット・・・水素原子を含む)を要素として連結し、X線結晶構造(参照構造)と類似の構造(標準構造)を生成することを考える。この問題は、一般的には、全ての原子の位置を変数とする非線形多変数最小2乗問題となる。

我われは、この近似解法として、既に、ローカルマッチング法による生成法を提案した^{1) 2) 3)}が、N末端からC末端へ順々に生成するにつれて、対応する参照構造から隔たってくる傾向がある。そこで、個々のアミノ酸残基の主鎖のユニットの生成は、すでに生成した主鎖のユニットの集まりを新たなユニットとみなして、参照構造とグローバルに一致するように補正を行ないつつ生成する改良ローカルマッチング法を用い、側鎖のユニットの生成は、従来通りローカルマッチング法を用いた。今回、この方法を実際のタンパク質分子に適用した結果について報告する。

2. X線結晶構造(参照構造)から標準構造への変換

タンパク質の骨格構造を形成している主鎖のユニットの生成は、すでに生成した主鎖のユニットの集

まりを新たなユニット(集合ユニット)とみなし、それに対応する参照構造の集合ユニットに最適マッチングするように補正を行う(図1)。側鎖のユニットの生成は、従来通り、2つの参照ユニットに最適マッチングさせるように行う。図2に示しているように、

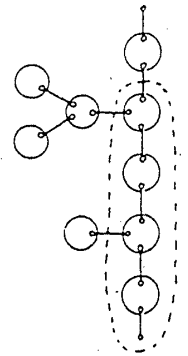


図1 主鎖の集まりま、(i-1)番目の側鎖のユニット U_{i-1} まで生成を終わっているとす。この生成ユニット U_{i-1}

の腕に、参照ユニット U^x_i に対応する標準ユニット U^s_i の脚を重ねて連結することによって U_i を生成する手順を述べ

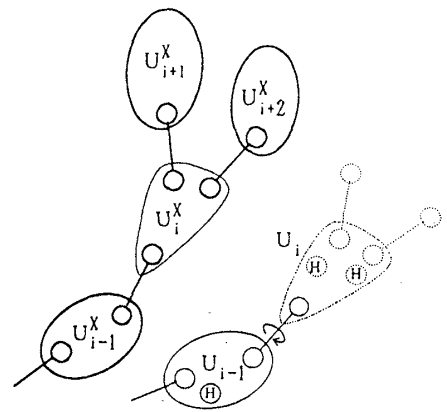


図2 参照ユニットと生成ユニット

(1) 参照ユニット U^x_i の脚がZ軸、足が原点になるように U^x_{i-1} (主鎖の場合は、(i-1)番目までの参照ユニットの集まりを U^x_{i-1} とみなす) と U^x_i の2つのユニットを座標変換する。(2) U_{i-1} (主鎖の場合は、すでに生成した(i-1)番目までの主鎖のユニットの集まりを U_{i-1} とみなす) を座標変換して、次の生成ユニット U_i の脚(= U_{i-1} の腕)がZ軸に、足が原点になるようにする。(3) 前生成ユニット U_{i-1} が、前参照ユニット U^x_{i-1} と最適マッチングするように参照ユニットをZ軸回転する。(4) U^x_i に対応した標準ユニット U^s_i をもってきて、足と脚軸を一致させる。(5) 標準ユニット U^s_i と参照ユニット U^x_i との最適

マッチングを行う。(6) この結果得られた U_i^s が生成ユニット U_i となる。

以上の手順をN末端からC末端へ順々に繰り返す。

3. 参照構造と標準構造の重ね合わせ

参照構造と新たに生成された標準構造とを比較・検討するために、これらの2つの構造を重ね合わせる。

いま対応する n 個の原子からなる標準構造と参照構造を表す二つの行列を考える：

$$\text{標準構造: } P = [p_1 \ p_2 \ \dots \ p_i \ \dots \ p_n]$$

$$\text{参照構造: } Q = [q_1 \ q_2 \ \dots \ q_i \ \dots \ q_n]$$

ここで p_i , q_i は、対応する原子の位置ベクトルである。空間回転行列 R によって二つの構造を重ね合わせようとするとき、対応する原子間の距離の2乗和を最小にするような正規直交回転行列 R を次のように求める。 P を正規直交回転行列 R で回転したとき、対応する Q の原子座標の誤差の2乗和 δ^2 は、 $\delta^2 = \|RP - Q\|^2 = \text{tr} \{(RP - Q)(RP - Q)^t\}$ 従って、 δ^2 を最小にする意味での最適マッチングを与える R は、次の条件のもとでの δ^2 の最小化問題となる： $RR^t = I$ (t : 行列の転置、 I : 単位行列) これは、ラグランジュの未定乗数法により解くことができる。この結果は、 QP^tPQ^t (対称行列) を固有値分解し、固有値を成分とする対角行列を D 、固有ベクトルを列にもつ正規直交行列を W とすると、 R は次式で表される： $R = PQ^tWD^{-1/2}W^t$

4. 実行例

図3は、ここで提案したアルゴリズム(補正ユニット数を30個とした)をグルカゴン(アミノ酸残基数

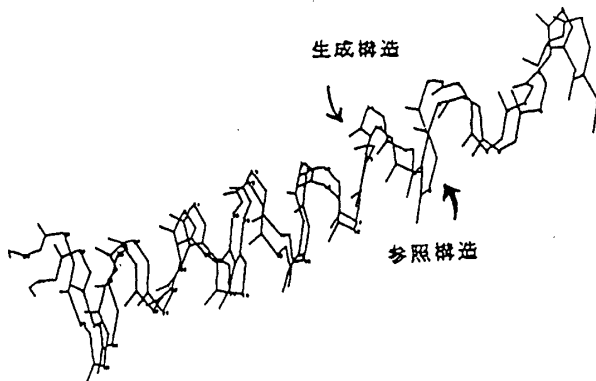


図3 グルカゴンの参照・標準構造の主鎖の比較

29個、ユニット数144個)に適用した後、参照構造と標準構造を重ね合わせた結果を主鎖の骨格構造で示している。全原子のr.m.s.は1.57Å、主鎖のr.m.s.は1.19Åである。図4は、アミノ酸残基内に属している主鎖の原子に対するr.m.s.の残基毎の平均値を表している。

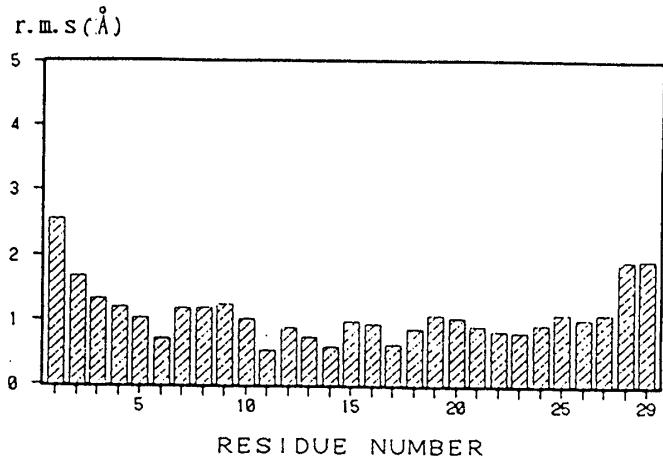


図4 グルカゴンの生成・参照構造の主鎖のr.m.s

5. あとがき

ここで提案したアルゴリズムをマイコン上で、実際のX線結晶構造データを用いて標準構造の生成を行い、参照構造と得られた標準構造との比較・検討を行った。小数残基のタンパク質分子に対しては、マイコン上で実用的な速度でX線結晶構造から標準構造へ変換できることが確かめられたが、他の多数残基のタンパク質分子に対しては誤差の蓄積が表れる傾向にある。このことに関しては講演で述べる。

6. 参考文献

- 1) 安部晴男、水戸三千秋、久保康司：“タンパク質分子の標準構造の生成法”、情報処理学会九州支部研究会報告、第5巻、pp9-15(1991).
- 2) 安部晴男、水戸三千秋、日比野真也：“改良ローカルマッチング法によるタンパク質分子の標準構造の生成”、第46回情報処理学会全国大会、7D-7(1993).
- 3) 安部晴男、水戸三千秋：“タンパク質分子のX線結晶構造から標準構造への変換”、平成5年度電気関係学会九支連大、No. 232(1993).