

3U-3

# タンパク質の標準立体構造の高速生成法

安部晴男、水戸三千秋、本田宗生、廣田悦司

西日本工業大学

## 1. まえがき

タンパク質の分子は、剛体とみなせる少数の原子からなる構造(ユニット)から構成されているとする。

分子モデリング・システムでは、まず最初に、処理対象のタンパク質分子の3次元構造を内部データとして生成する必要がある。

このためには、タンパク質分子のX線結晶構造解析によるデータを使用するが、分解能不足のため水素原子が欠落していて、分子内の原子結合距離・結合角も誤差を含んでいる。そこで、低分子の研究結果に基づいた標準的な原子結合距離・結合角をもつユニット(標準ユニット)を要素として連結し、X線結晶構造に類似した構造(標準立体構造)を生成することが必要となる。この問題を、原子位置を変数とする最小2乗問題として考えると、多変数最適化問題となる。

我われは、この近似解法として、以前に、ユニット毎のマッチングにより、標準ユニットを次々に連結してタンパク質の標準構造を生成する方法を提案した<sup>1)</sup>。

今回は、これをさらに改良した、より効率の良い生成法について報告する<sup>2)</sup>。

## 2. 標準ユニット

ユニットを、図1の様

に定義する。タンパク質分子は、多種類の標準ユニットからなり、2つのユニットは、腕と脚を共有する形で接続されていて、この腕(=脚)は軸回りの回転の自由度をもっている。

標準ユニットの局所座標系は、足を原点、脚が常にZ軸の正の向きになるようにとる。

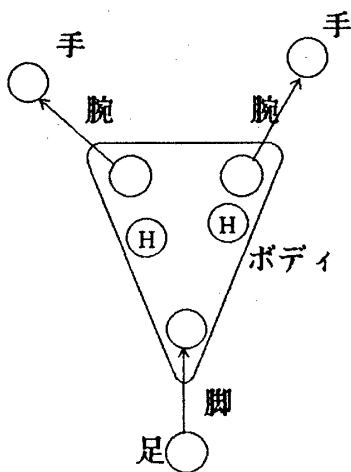


図1 ユニットの定義

## 3. ローカルマッチング法による標準立体構造の生成手順

生成しつつあるタンパク質の標準立体構造を生成ユニットと呼び、いま、(i-1)番目のユニット  $U_{i-1}$  まで生成されているものとする。また、X線結晶構造でのi番目のユニット(参照ユニットと呼ぶ)を  $U^*_i$ 、対応する標準ユニットを  $U^s_i$  で表す。  $U_i$  を生成する手順は次の通りである。

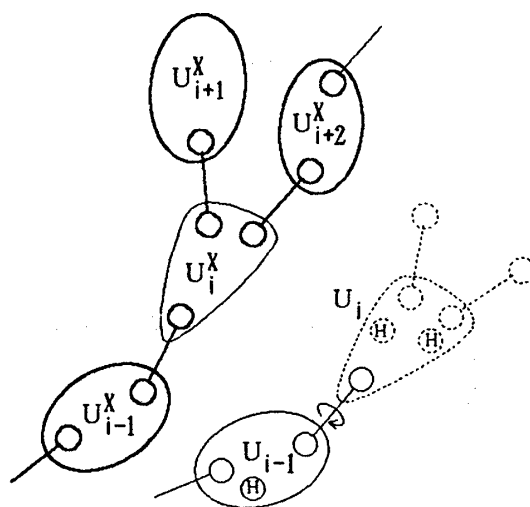


図-2 参照ユニットと生成ユニット

### 【生成ユニット $U_i$ の脚の準備: Z軸方向】

- (1)  $U_{i-1}$  を平行移動して、手の付け根(=次の  $U_i$  の足: 位置ベクトル  $\tau$ ) を原点に一致させる。この平行移動のための座標変換行列を  $M_1(-\tau)$  とする。
- (2)  $U_{i-1}$  の腕(=  $U_i$  の脚) がZ軸の正方向に一致するように  $U_{i-1}$  を直近回転する。このとき、回転角を  $\phi$  とし、回転マトリックスを  $R_1(\phi)$  とする。

### 【参照ユニット $U^*_i$ のZ軸への立ち上げ】

- (3) 参照構造の  $U^*_{i-1}$  と  $U^*_i$  の部分を連結したまま平行移動し、  $U^*_{i-1}$  の手の付け根(=  $U^*_i$  の足) を原点に一致させる。
- (4)  $U^*_i$  の脚(=  $U^*_{i-1}$  の腕) がZ軸の正方向に一致するように、  $U^*_{i-1}$  と  $U^*_i$  とを連結したまま、直近回転して立ち上げる。

An Effective Method for Generation of Protein Conformations with the Standard Geometry

Haruo ABE, Michiaki MITO, Yoshio HONDA, Etuji HIROTA

Nishinippon Institute of Technology

【前参照ユニット  $U^{s_{i-1}}$  と前生成ユニット  $U_{i-1}$  のマッチング補正】

(5)  $U^{s_{i-1}}$  と  $U_{i-1}$  の対応する原子間の距離の2乗和が最小となるように、 $U^{s_{i-1}}$  と  $U^{s_i}$  をZ軸回転する(付録参照)。

【標準ユニット  $U^{s_i}$  と参照ユニット  $U^{s_i}$  とのマッチング】

(6)  $U^{s_i}$  に対応した標準ユニット  $U^{s_i}$  (脚はZ軸)を、最小2乗の意味で最適マッチングするようにZ軸回転する。この時の回転行列  $T_i(\theta)$  は、 $U^{s_i}$  と  $U^{s_i}$  の2つのユニットの構成原子位置座標より決定される(付録参照)。

【生成ユニット  $U_i$  の生成】

(7) 標準ユニット  $U^{s_i}$  の構成原子の局所座標系での位置行列  $P^{s_i} = [p^{s_i}_1 \ p^{s_i}_2 \ \dots \ p^{s_i}_n]$  を、次の式で変換して、生成原子位置行列  $P_i$ 、すなわち  $U_i$  を生成する。

$$P_i = M_i(\tau) \cdot R_i(-\phi) \cdot T_i(\theta) \cdot P^{s_i}$$

この手順を最初のユニットからスタートして、すべてのユニットに対して繰り返すことによって、参照構造と類似の標準立体構造を生成する。

4. 実行例

本方法に基づいて、マイコン PC-9801RA 上で Borland C++ によりプログラムを作成し、X線結晶構造データの少数の残基に対して標準立体構造の生成実験を行った。

例として、リゾチーム分子の一番目のリジン残基についての結果を図3、図4に示す。

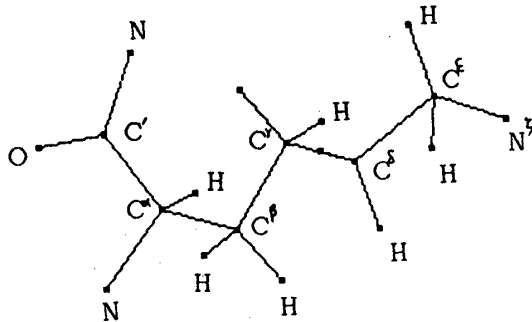


図3 リジン残基の生成構造

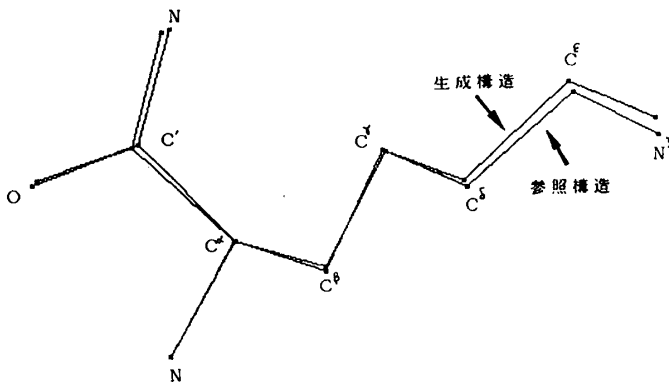


図4 リジン残基の参照・生成構造の比較

5. あとがき

タンパク質分子の標準的な立体構造を生成する方法として、より効率的で、高速なアルゴリズムを提案した。

タンパク質分子の数残基への適用例ではあるが、マイコン上で実用的な速度で実行できることが確かめられた。

今後、実際のタンパク質分子に適用して、有効性を確認することが必要である。

6. 参考文献

- 1) 安部晴男、水戸三千秋、久保康司：“ローカルマッチング法によるタンパク質の標準構造の生成”、電子情報通信学会春季全国大会、D-205(1991)。
- 2) 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)。

【付録】

互いに対応する  $n$  個の原子からなる2つのユニットにおいて、一方をZ軸回転することによって、対応する原子間の距離の2乗和を最小化するような、正規直交回転行列  $T(\theta)$  を求める。

いま、2つのユニットの原子位置行列を  $P$ 、 $Q$  とすると、これらは、次式で表せる。

$$P = [p_1 \ p_2 \ \dots \ p_n]$$

$$Q = [q_1 \ q_2 \ \dots \ q_n]$$

ここで、

$$p_j = \begin{bmatrix} p_{jx} \\ p_{jy} \\ p_{jz} \\ 1 \end{bmatrix}, \quad q_j = \begin{bmatrix} q_{jx} \\ q_{jy} \\ q_{jz} \\ 1 \end{bmatrix}$$

Z軸回りの正規直交回転行列  $T$  は、

$$T(\theta) = \begin{bmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$\Delta = PT - Q$  において、 $\Delta^t \Delta$  を最小化する。

$$\frac{\partial \text{tr}(\Delta^t \Delta)}{\partial \theta} = 0, \quad \frac{\partial^2 \text{tr}(\Delta^t \Delta)}{\partial \theta^2} > 0$$

これは容易に解けて、 $T(\theta)$  の成分が次のように求まる。

$$\cos \theta = \frac{\alpha}{(\alpha^2 + \beta^2)^{1/2}}, \quad \sin \theta = \frac{\beta}{(\alpha^2 + \beta^2)^{1/2}}$$

ここで、

$$\alpha = \sum_{j=1}^n (p_{jx} q_{jx} + p_{jy} q_{jy}),$$

$$\beta = \sum_{j=1}^n (p_{jy} q_{jx} - p_{jx} q_{jy})$$