

タンパク質の立体構造予測

Prediction of Protein Tertiary Structures

角 美智子 (知的システムデザイン研究室)

Michiko Sumi (Intelligent Systems Design Laboratory)

Abstract: Proteins are the most complicated molecules that exist in nature. Since protein structures are closely related to their biological functions, and the prediction of their tertiary structures is very important problem in theoretical molecular science. It is widely believed that tertiary structure of proteins is determined by their amino-acid sequences. However, it is not only because the inclusion of solvent effects in the energy function is formidable but also because there exist huge number of local minima in the energy function, and SA is used at the prediction of their tertiary structures. This paper reports the protein structures and prediction of their tertiary structures.

1 はじめに

タンパク質分子の立体構造とその生化学的な機能の間には密接な関係がある。よって、タンパク質の立体構造の研究は、その機能発現の仕組みを解明することを目指しており、究極的には新薬品や特定の機能を持った人工タンパク質の開発ばかりでなく、狂牛病やアルツハイマー病などのタンパク質の誤った折り畳みに起因する病理の発現機構の解明へとつながっている¹⁾。本発表ではタンパク質の構造とその解析法について述べる。

2 タンパク質の構造

2.1 一次構造

タンパク質は 20 種類のアミノ酸が鎖状に連結して作られる。タンパク質の造りにあたるこのアミノ酸配列のことを一次構造という。アミノ酸配列は多様で鎖の長さもさまざまであることから、多種類のタンパク質が存在する。しかし、1 種類のタンパク質は固有のアミノ酸配列を持っており、その結果独自の特異的な働きをする。アミノ酸は、1 個の炭素原子にアミノ基 (-NH₂)、カルボキシル基 (-COOH)、水素原子 (-H)、さらに側鎖 (-R) が結合した構造を持つ。側鎖には、様々な分子構造があり、これがアミノ酸の性質を決めている。アミノ酸のカルボキシル基と、隣のアミノ酸のアミノ基との間で縮合反応がおこり、ペプチド結合 (-CONH-) が形成される。これによりできあがった分子をペプチドと呼ぶ。アミノ酸が多数つながった分子をポリペプチドといい、これがタンパク質である²⁾。図 1 がポリペプチド、つまりタンパク質の一次構造である。

2.2 高次構造

アミノ酸配列が、熱力学的作用により折りたたまれると複雑な構造と機能を持つタンパク質となる。ポリペ

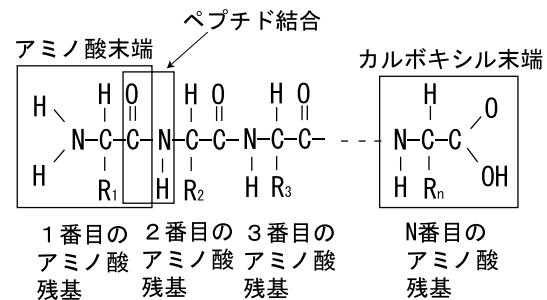


Fig. 1 タンパク質の構成

チド鎖内の結合には、結合を中心軸として自由に回転できるものが多いので、タンパク質分子の取り得る折りたたまれ方は原理的には膨大な数に上ることが推測される。しかし実際には、そのアミノ酸配列 (一次構造) によって決まるただ一つの折りたたみ方しかとっていない。これは、タンパク質が、熱力学的に最も安定した状態、すなわち自由エネルギーを最小化するように折りたたまれるためである。この折りたたまれた状態を立体構造 (三次構造) という。つまり、タンパク質はその一次構造が決まれば、立体構造も決まるといことである。

三次構造の一部分にあたる、タンパク質の折りたたみ方を二次構造という。二次構造の代表的なものにはヘリックスとシートがある。ヘリックスは 1 本のペプチド主鎖が規則正しくらせん状に巻いた状態になっている。そこではすべてのペプチド結合が水素結合を形成しており、エネルギー的に最も安定な構造である。シートは、平行に配置された 2 本のポリペプチド主鎖が水素結合によって固定された構造である。3 本以上のポリペプチド主鎖が平行に並ぶことも多く、この場合できるだけ水素結合を形成し安定になる。ヘリックスやシー

トの間をつなぐ部分をループと呼び、これはポリペプチド鎖の長い折れ曲がりであり、長さや構造がまちまちである。ループの中にはターンと呼ばれる3~4アミノ酸残基から成るU字型の二次構造が見られることがある。

タンパク質の中には、ポリペプチド鎖が複数寄り集まった構造をしているものがあり、この寄り集まっているポリペプチド鎖の一つ一つをサブユニットという。サブユニットの構造のことを四次構造と呼ぶ。これら二次構造と三次構造と四次構造をまとめて高次構造という²⁾

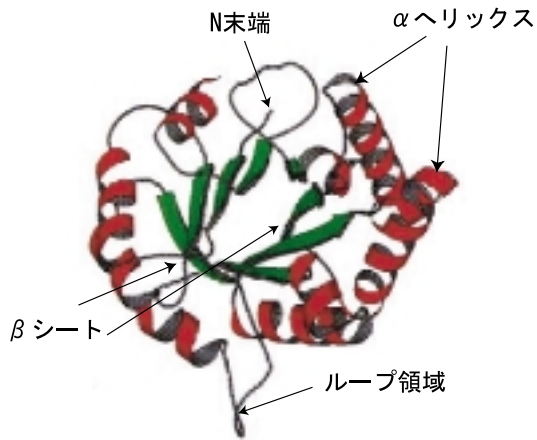


Fig. 2 タンパク質の立体構造

3 立体構造予測

タンパク質の自然の立体構造(自由エネルギーが最小の状態)をそのアミノ酸配列及びタンパク質を取り巻く環境因子の情報のみから観測することを、第一原理からの立体構造予測と呼ぶ¹⁾。

タンパク質のモデルはあまりに複雑なため、最適化の問題では数多くの同程度に最適な構造を与える可能性がある。つまり多くの極小点をもつことを意味している。このことから線形計画法の降下法は有効ではない。一方、モンテカルロ法は目的関数を小さくする方向ばかりではなく、ある程度大きくなる方向へも進み極小点から脱出できるが、計算量は飛躍的に増大することとなる³⁾。

そこでタンパク質の広い構造空間を効率よく探索できることが要件となり、シミュレーテッドアニーリング(SA)が導入された。そのことにより、実際にかなりタンパク質の構造決定が効率化された。現在もSAを用いた構造予測が広く行われている。以下に、タンパク質の構造予測を行うときのSAの各過程、パラメータおよびエネルギー関数の定義を示す。

エネルギー関数 タンパク質分子の構造エネルギーと溶媒和の自由エネルギーの和

生成 一回のアニーリングにおいて、全ての回転角を一回ずつ変更

受理判定 メトロポリス基準

クーリング 指数型アニーリング

初期値 乱数で二面角をランダムに準備したもの

最高温度 1000K

最低温度 250,100,50K

SAは、十分高温(最高温度)でシミュレーションを始め、徐々に温度を下げていき、最終的に最小エネルギー状態に達する方法である。タンパク質は、最適な立体構造上の自由度ごとに異なる構造をつくるが、これらの回転角(二面角)¹⁾の数を設計変数の数とする。1回のアニーリングにおいて、全ての回転角を1回ずつ変更し、エネルギーが最小となる構造、つまり最適な立体構造に向かう経路を探索する。

4 今後の予定

これまでの研究において、基本的なタンパク質の構造、及びSAアルゴリズム、FORTRANプログラミングの理解を図った。今後は分子科学研究所で使用されているプログラムを用いて、SAによる構造解析の更なる効率化、高性能化を図ると共に、ビジュアル化することも目標とする。

参考文献

- 1) 岡本 祐幸 『モンテカルロシミュレーションで探るタンパク質の折り畳み機構』
(物性研究, 1998)
- 2) 池内俊彦 『生命を学ぶ タンパク質の科学』
(オーム社, 1999)
- 3) 小寺 詔紀 『コンピュータ解析 - 最適化をめぐる』
(共立出版, 1995)
- 4) 小長谷明彦 『遺伝子とコンピュータ』
(共立出版, 2000)

¹⁾空間内で2つの平面が互いに交わるときの角度を、二平面交角または二面角という