

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年5月27日現在

機関番号：14301  
 研究種目：基盤研究(B)  
 研究期間：2010～2012  
 課題番号：22300100  
 研究課題名（和文） 液体論を基盤とした蛋白質立体構造予測法の構築  
 研究課題名（英文） Construction of protein-structure prediction method based on liquid state theory  
 研究代表者  
 木下 正弘 (KINOSHITA MASAHIRO)  
 京都大学・エネルギー理工学研究所・教授  
 研究者番号：90195339

研究成果の概要（和文）：水及び疎水環境中における蛋白質の立体構造安定性を解析し、その違いを明らかにした。また、両者に対する液体論に根ざした自由エネルギー関数を構築し、それらの有用性を確認した。実験で得られた蛋白質天然構造モデルのキャラクタリゼーション法を開発した。生物情報科学分野で用いられる手法を用いて多数の候補構造を発生させ、水環境中における蛋白質用の自由エネルギー関数に最低値を与える構造を選定する方式を試みたところ、幾つかの蛋白質において有望な結果が得られた。

研究成果の概要（英文）：We analyze the structural stabilities of proteins in aqueous and nonpolar environments, thus revealing their differences. Free-energy functions are constructed for these proteins on the basis of liquid-state theories and their usefulness is demonstrated. After generating a number of candidate structures using typical bioinformatics techniques, we select the structure giving the free-energy function for a protein in aqueous environment: Promising results are obtained for several proteins tested.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	5,000,000	1,500,000	6,500,000
2011年度	4,100,000	1,230,000	5,330,000
2012年度	4,200,000	1,260,000	5,460,000
年度			
年度			
総計	13,300,000	3,990,000	17,290,000

研究分野：理論生物物理学，液体の統計力学

科研費の分科・細目：情報学，生体生命情報学

キーワード：蛋白質，フォールディング，立体構造予測，水和，自由エネルギー関数

1. 研究開始当初の背景

蛋白質は固有の立体構造（天然構造）に折り畳み、その立体構造に基づいた機能を発揮している。ポストゲノム計画の中心課題は、それらの蛋白質の立体構造と機能を明らかにし、蛋白質を主役とする生命現象の発現機構を解明することである。この課題が解決す

れば、新しい機能を持つ人工蛋白質の創製や、新薬の設計、画期的な治療法の開発などが可能となる。しかし、アミノ酸配列から天然構造を予測可能にすることは、60年来未解決の超難問である。その主な理由は、天然構造の候補となる構造の数が膨大であり、かつ、決定的に重要な水の役割を的確かつ効率的に

考慮できる液体論に根ざした方法を持ち合わせていなかったことにある。水-蛋白質系に全原子モデルを適用し、原子間の静電相互作用やレナード・ジョーンズ相互作用などを考慮して、計算機シミュレーションによって立体構造空間を探索する試みがある。しかし、膨大な計算労力が大きな問題となる。さらに大きな問題は、観測される結果が、用いる蛋白質用力場パラメータによって大きく変わり得ることである。そのため、現在では、情報科学的方法が主流となっている。立体構造が既知の蛋白質に対するアミノ酸配列のデータベースを検索・参照し、情報科学的手法を駆使して候補となる複数の立体構造を作り出すのであるが、最良と判断される立体構造を特定する最終段階で行き詰る傾向がある。水の効果を満足に考慮していないし、折り畳み機構の解明にもつながり難い。計算労力や力場パラメータの不確定性の問題が深刻化せず、かつ、折り畳み機構の解明につながる Physics-based な立体構造予測法の確立が強く望まれる。

一方、我々は、蛋白質折り畳みの推進力が水の並進配置エントロピー利得（水分子に対する並進の自由度の制限が緩和されることに起因）であり、この因子が蛋白質分子の構造エントロピー損失と脱水和のペナルティー（折り畳みに伴って蛋白質-水間の水素結合が切断されるが、それを分子内水素結合で必ずしも補償できないことに起因するエンタルピー上昇）を凌駕することを既に示している。また、以上の概念に基づき、蛋白質立体構造予測用の自由エネルギー関数を既に開発している。精密な水分子のモデルを用いた分子性流体用積分方程式論を用いているにも拘らず、形態熱力学的アプローチと統合することにより、1つの立体構造当たり0.1秒程度でこの自由エネルギー関数を計算できる。以上の成果に基づいて、Physics-based な立体構造予測法の確立に向けた新たな動きを作り出した。

## 2. 研究の目的

本研究の目的は以下の通りである：水環境中における蛋白質と疎水環境中における蛋白質の立体構造安定性を比較・検討する；同時に、疎水環境中にある膜蛋白質用の自由エネルギー関数を開発する（1. で述べた、既に提案した水環境中にある蛋白質用の自由エネルギー関数を修正する）；上記2通りの自由エネルギー関数の性能を詳細に検討する；実験で得られた蛋白質天然構造モデルのキャラクタリゼーション法を提案し、より良

いモデルを作成するための指針を与える；水環境中にある蛋白質用の自由エネルギー関数と情報科学的手法を組み合わせることにより、立体構造予測に向けた基礎解析を行う。

## 3. 研究の方法

我々の自由エネルギー関数  $F$  は次式で与えられる：

$$F = (-TS + \Lambda) / (k_B T_0), \quad T_0 = 298 \text{ K}. \quad (1)$$

$S$  は水和エントロピー、 $\Lambda$  は伸びた構造を基準にした「水和エネルギー+構造エネルギー」に該当し、構造のコンパクト化に伴う脱水和のペナルティーの総和である。水分子と最大数の水素結合を形成しているが分子内水素結合は持たない完全に伸びた構造に比べると、よりコンパクトな構造では、水分子との水素結合 ( $\text{CO} \cdots \text{W}$ ,  $\text{NH} \cdots \text{W}$  など) の切断と共に、幾つかのドナー ( $N$ ) とアクセプター ( $O$ ) が内部に埋もれている。その場合、分子内水素結合 ( $\text{CO} \cdots \text{HN}$  など) が形成されていれば問題ないが、必ずしも形成できているとは限らず、脱水和のペナルティーを被ることになる。

$T = T_0$  に設定して考える。脱水和に対して、水分子との水素結合の切断に起因するものだけを考える。水環境中および蛋白質内部中における  $\text{CO} \cdots \text{HN}$  なる水素結合形成のエネルギー低下を各々  $0$  および  $-14k_B T_0$  に設定する。ドナー ( $N$ ) とアクセプター ( $O$ ) が水分子との水素結合を切断して埋もれた場合、分子内水素結合を形成すれば損得なしとする。しかし、ドナーまたはアクセプターが埋もれても分子内水素結合の相手がいない場合、 $7k_B T_0$  のペナルティーを課す。以上述べた方式を図1に示す。埋もれているか否かは、露出表面積がゼロか否かで判断する。主鎖と側鎖のすべてのドナーとアクセプターに対して埋もれているか否かを判定し、主鎖-主鎖、主鎖-側鎖、側鎖-側鎖に対する分子内水素結合形成をチェックする。このようにして、脱水和のペナルティーの総和をとったものが  $\Lambda$  である。 $-S$  と  $\Lambda$  (共に正の量) は蛋白質立体構造の関数である。

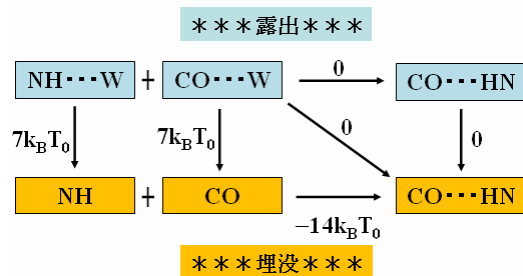


図1:  $\Lambda$ を計算するための熱力学サイクル (水環境中)

水の並進配置エントロピーの効果は、水を大きさを持つ粒子の集まりとしてモデル化しなければ考慮することができない。本研究では、水に対して多極子モデルを採用（水分子間相互作用は、中心間距離および2分子の方向の多変数関数）し、蛋白質の複雑な多原子構造を4つの幾何学的指標（排除容積、露出表面の表面積・平均曲率の積分値・ガウス曲率の積分値）で代表させる「分子性流体用積分方程式論と形態熱力学のアプローチの統合型方法論」を用いて $-S$ を計算する。

疎水環境中にある蛋白質用の自由エネルギー関数も式(1)で与えられるが、 $-S$ と $\Delta$ の中味が異なる。 $-S$ は、脂質二重層を形成するリン脂質中の炭化水素基( $\text{CH}_2$ ,  $\text{CH}$ ,  $\text{CH}_3$ )の並進移動に起因する。膜内環境を剛体球系溶媒としてモデルし、回転対称積分方程式論と形態熱力学のアプローチの統合型方法論を用いて計算する。 $\Delta$ は分子内水素結合形成に伴うエネルギー低下を表す負の量となり、図2に示す方式で計算する。

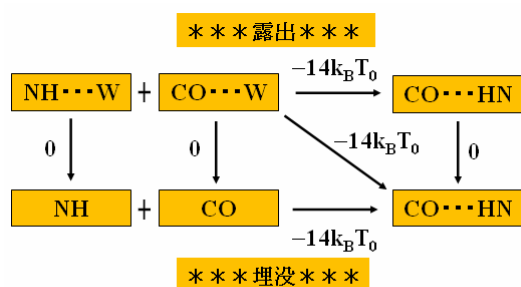


図2:  $\Delta$ を計算するための熱力学サイクル (疎水環境中)

#### 4. 研究成果

(1) 水環境中における蛋白質と疎水環境中における蛋白質の立体構造安定性を比較・検討した。疎水環境中における蛋白質（膜蛋白質）に適用可能な自由エネルギー関数は本研究で新たに開発したものである。

水環境下において側鎖の充填が2次構造選定に決定的な影響を与えることを水のエンタロピーに着目して示した。2次構造は主鎖の排除体積を減少させ、水のエンタロピー利得をもたらす。しかしながら、側鎖同士の密な充填も極めて大きな水のエンタロピー利得をもたらす。蛋白質の立体構造安定性に決定的な影響を及ぼす。そのため、折り畳み過程において側鎖の密な充填がむしろ優先され、それを達成できるような $\alpha$ -ヘリックスや $\beta$ -シートの含有率が選定される。その結果、天然構造中の2次構造の含有率は必ずし

も高くなく、さらに、 $\alpha$ -ヘリックスのみ、 $\beta$ -シートのみ、あるいはその両者を有するパラエティーに富んだ天然構造が安定化される（図3）ことが分かった。

水環境下では、2次構造は「水のエンタロピー利得」と「分子内水素結合形成による、蛋白質-水分子間水素結合の切断の補償」の両方を確保できる格好の構造ユニットであり、できるだけ多く形成することが望まれる。しかし、側鎖の充填に伴う水のエンタロピー利得が支配的に効くため、上で述べた結果はエンタルピー成分を組み込んでも基本的には変わらない。一方、疎水環境中では溶媒のエンタロピーの効果が弱くなる。そのため、側鎖の密な充填よりも分子内水素結合形成が重要となる。 $\alpha$ -ヘリックス構造は $\beta$ -シート構造よりも、より効率良く分子内水素結合を形成でき、 $\alpha$ -ヘリックスが好んで選定されることが分かった。その結果、膜蛋白質は $\alpha$ -ヘリックスを形成することが多いが、 $\alpha$ -ヘリックス同士の会合は脂質2重層膜の疎水鎖を構成する炭化水素基( $\text{CH}_2$ ,  $\text{CH}_3$ ,  $\text{CH}$ )の並進移動に起因するエンタロピー利得により駆動されることが示された（図4）。

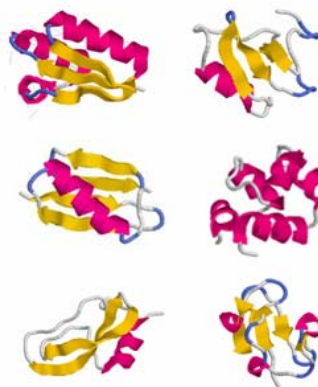


図3: 水環境中で安定化される蛋白質の例



図4: 疎水環境中で安定化される蛋白質の例

(2)水環境中にある蛋白質用の自由エネルギー関数が非常に高い性能を有することを「デコイ構造と呼ばれる非天然構造と天然構造のセットの中から天然構造を射当てるテスト」により検証した。テストは、8つのデコイセットの133種類に及ぶ蛋白質に対して行った。その結果、現在までに報告されているどのエネルギー関数をも凌駕する成績(ほぼ100%的中率)を収めた。この結果は、木下らの自由エネルギー関数が天然構造の特徴を的確に捉えており、蛋白質の立体構造予測に適用できることを示している。

以下に述べる興味深い結果が得られた。天然構造よりも挿入に伴う水のエン트로ピー損失を小さく出来る(充填効率が低い)構造は、全体として深刻な脱水和のペナルティーを被る。逆に、天然構造よりも全体として脱水和のペナルティーを小さくできる構造は、挿入に伴う水のエン트로ピー損失が深刻となる。天然構造は、「水のエン트로ピー損失+脱水和の全ペナルティー」に対して最適化されている。

(3)水環境中にある蛋白質用の自由エネルギー関数とそのエンタルピー及びエン트로ピー成分を用いて、タンパク質の天然構造モデルのキャラクターゼーションを行う方法を提案した。タンパク質の立体構造は、X線結晶解析やNMR実験を通して決定される。得られた天然構造モデルは、実験条件や構造計算の方法の影響を大きく受けるため、それらが適切なものかを評価する必要がある。木下らの自由エネルギー関数は天然構造の特徴をよく捉えているため、各々の天然構造モデルの特徴や弱点を明らかにすることができる。提案された方法を用いることにより、より良い天然構造モデルを作成するための有用な指針が得られる。

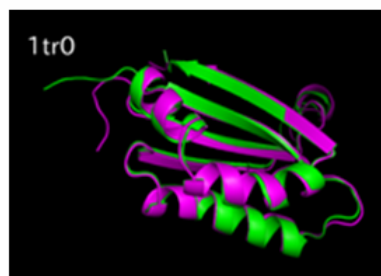
(4)蛋白質立体構造予測に向けた基礎解析を行った結果を紹介する。

1ptf (PDB Code) に対して、粗視化基準振動解析(十分近い距離にある $C_{\alpha}$ をバネで連結して連成振動させ、天然構造の主鎖構造からあまり大きくは異なる数多くの主鎖構造を発生させ、それぞれに側鎖を付けて真空中でエネルギーのミニマイゼーションを行う)を用いて、天然構造からあまり大きくは異なる数多くの立体構造を発生させ、各構造に対する我々の自由エネルギー関数を計算した。天然構造を基準にした $C_{\alpha}$ についてのRMSDが小さいほど(天然構造に近いほど)、自由エネルギー関数値が低くなる強い傾向が認められ、この観点から、我々の自由エネルギー関数は立体構造予測に対して非常に

有力であることが分かる。

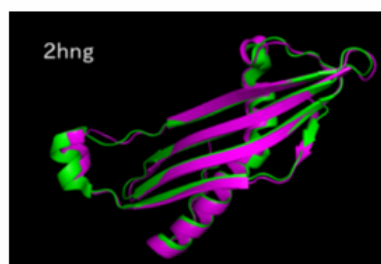
比較モデリング法あるいはフラグメントアセンブリ法を用いて数多くの候補構造を発生させ、我々の自由エネルギー関数を用いてそれに最低値を与える構造を選定した。

比較モデリング法を用いた場合の結果は以下の通りである。1tr0に対して我々が選定した構造(図5)は、天然構造(X線結晶構造)を基準にした $C_{\alpha}$ についてのRMSDが1.30 Åとなった。緑が天然構造、紫が選定構造で、両者はかなり良く一致している。一方、CASPで優勝したチームが提出した構造のRMSDは2.08 Åであり、我々が選定した構造の方が天然構造に近い。2hngに対しては、図6のような結果が得られている。ただし、この蛋白質の場合、CASPが行われた頃に比べるとデータベースが更新されており、我々が選定した構造の方が天然構造に近くて当然であるかもしれない。逆に言えば、データベースが更新されるにつれて、比較モデリング法の有用性が増すということである。



緑:天然構造, 紫:選定構造.

図5:我々が選定した構造では、RMSD=1.30 Å. CASPで優勝したチームが提出した構造では、RMSD=2.08 Å.

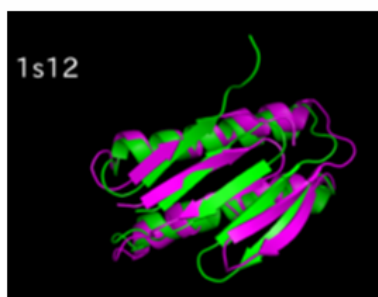


緑:天然構造, 紫:選定構造.

図6:我々が選定した構造では、RMSD=1.23 Å. CASPで優勝したチームが提出した構造では、RMSD=5.38 Å.



フラグメントアセンブリ法を用いた場合 (図7, 図8) には, 我々が選定した構造よりも CASP で優勝したチーム (Baker のグループ) が提出した構造の方が天然構造に近い。しかし, 我々の計算では十分多くの候補構造がサンプリングできていないからではないかと考えられる。



緑:天然構造, 紫:選定構造.

図7:我々が選定した構造では, **RMSD=2.44 Å**. CASPで優勝したチームが提出した構造では, **RMSD=2.08 Å**.



緑:天然構造, 紫:選定構造.

図8:我々が選定した構造では, **RMSD=3.70 Å**. CASPで優勝したチームが提出した構造では, **RMSD=1.58 Å**.

今回, それぞれの蛋白質に対して膨大な数の候補構造を発生させたが, 天然構造 (X線結晶構造) よりも自由エネルギー関数値が低くなる構造は決して見出せなかった。このことは, 自由エネルギー関数自体は極めて有望であり, 候補構造の発生法に問題がある (天然構造に近い構造をサンプリングできていない) ことを示唆している。比較モデリング法を用いることができる場合には候補構造の発生法に大きな問題は見られないが, フラグメントアセンブリ法を用いた場合にそのような問題が生じた。

(5) 今後の展望について要約する。水環境中および疎水環境中における蛋白質立体構造型成のメカニズムを明らかにし, その記述には水や炭化水素基 ( $\text{CH}_2$ ,  $\text{CH}_3$ ,  $\text{CH}$ ) に対して分子モデルを用いた統計力学理論が不可欠であることを示した。前者に対する自由エネルギー関数の有用性を示すと共に, 後者に対する新たな自由エネルギー関数を開発した。また, 実験で得られた情報から天然構造モデルを作成する過程に上記の自由エネルギー関数を応用することも可能である。膜蛋白質は, その立体構造を実験的に決定するのが難しいため, 理論的援用が必要であるが, 本研究で開発した自由エネルギー関数が有用となる。また, 出来る限り少ない NMR 実験情報 (願わくばケミカルシフトのみ) から優れた天然構造モデルを作成できる方法の開発が期待される。天然構造の立体構造予測は実験情報がゼロとなった場合に対応するため, まだ道は長い, 少なくとも天然構造の特徴を的確に捉えた自由エネルギー関数 (計算時間が1つの立体構造あたり0.1秒程度に過ぎないことも利点) が得られたことは大きな進展である。今後は, 生物情報科学分野で培われてきた種々の手法に我々の自由エネルギー関数をいかにして組み込むかが最大のポイントになるであろう。

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 32 件)

- ① M. Kinoshita, “A New Theoretical Approach to Biological Self-Assembly”, *Biophys. Rev.*, in press. (査読有)  
DOI: 10.1007/s12551-013-0100-8
- ② 木下正弘, 永山國昭, 「タンパク質水和理論の新機軸Ⅲ. 理論的考察」, *生物物理* **52**, 300-303 (2012). (査読有)  
[https://www.jstage.jst.go.jp/article/biophys/52/6/52\\_300/\\_pdf](https://www.jstage.jst.go.jp/article/biophys/52/6/52_300/_pdf)
- ③ 木下正弘, 永山國昭, 「タンパク質水和理論の新機軸Ⅱ. 新理論の応用展開」, *生物物理* **52**, 250-253 (2012). (査読有)  
[https://www.jstage.jst.go.jp/article/biophys/52/5/52\\_250/\\_pdf](https://www.jstage.jst.go.jp/article/biophys/52/5/52_250/_pdf)
- ④ 木下正弘, 永山國昭, 「タンパク質水和理論の新機軸Ⅰ. 朝倉一大沢理論を越えて」, *生物物理* **52**, 203-205 (2012). (査読有)  
[https://www.jstage.jst.go.jp/article/biophys/52/4/52\\_203/\\_pdf](https://www.jstage.jst.go.jp/article/biophys/52/4/52_203/_pdf)
- ⑤ S. Yasuda, H. Oshima, and M. Kinoshita, “Structural Stability of Proteins in

Aqueous and Nonpolar Environments”, J. Chem. Phys. **137**, 135103(1-14) (2012). (査読有)

DOI: 10.1063/1.4755755

- ⑥D. Shiqiao, Y. Harano, M. Kinoshita, and M. Sakurai, “Scoring Function Based on Solvation Thermodynamics for Protein Structure Prediction”, BIOPHYSICS **8**, 127-138 (2012). (査読有)

DOI: 10.2142/biophysics.8.127

- ⑦H. Mishima, S. Yasuda, T. Yoshidome, H. Oshima, Y. Harano, M. Ikeguchi, and M. Kinoshita, “Characterization of Experimentally Determined Native-Structure Models of a Protein Using Energetic and Entropic Components of Free-Energy Function”, J. Phys. Chem. B **116**, 7776-7786 (2012). (査読有)

DOI: 10.1021/jp301541z

- ⑧S. Yasuda, T. Yoshidome, Y. Harano, R. Roth, H. Oshima, K. Oda, Y. Sugita, M. Ikeguchi, and M. Kinoshita, “Free-Energy Function for Discriminating the Native Fold of a Protein from Misfolded Decoys”, Proteins **79**, 2161-2171 (2011). (査読有)

DOI: 10.1002/prot.23036

- ⑨S. Yasuda, T. Yoshidome, H. Oshima, R. Kodama, Y. Harano, and M. Kinoshita, “Effects of Side-Chain Packing on the Formation of Secondary Structures in Protein Folding”, J. Chem. Phys. **132**, 065105(1-10) (2010). (査読有)

DOI: 10.1063/1.3319509

[学会発表] (計 69 件)

- ①木下正弘, 「生命現象発現における水分子並進移動の役割」, 新世代研究所第 6 回合同研究会, 明治大学紫紺館, 2011 年 11 月 29 日.
- ②木下正弘, 「分子認識などの生命現象における水の並進エントロピーの重要性」, 日本学術振興会回折構造生物第 169 委員会第 36 回研究会, ゆうぽうと, 東京都品川区, 2011 年 10 月 4 日.
- ③木下正弘, 「「疎水性」の本当の物理起源: その温度依存性を通しての考察」, 第 56 回低温生物工学会大会セミナー, 招待講演, いわて県民情報交流センター, 2011 年 7 月 7-8 日.

[図書] (計 2 件)

- ①木下正弘, 「表面の水和の統計力学理論」,

日本化学会・コロイドおよび界面化学部会 35 周年記念出版, 丸善, 「現代界面コロイド科学の事典—サンプルから宇宙まで—」, 10 章 6 節, pp. 244-245 (2010).

[その他]

ホームページ

<http://www.iae.kyoto-u.ac.jp/centerbunya/kinoshita> (木下研究室—京都大学エネルギー理工学研究所)

## 6. 研究組織

### (1) 研究代表者

木下 正弘 (KINOSHITA MASAHIRO)  
京都大学・エネルギー理工学研究所・教授  
研究者番号: 90195339

### (2) 研究分担者

( )

研究者番号:

### (3) 連携研究者

高田 彰二 (TAKADA SHOJI)  
京都大学・大学院理学研究科・教授  
研究者番号: 60304086