

機関番号：31201

研究種目：若手研究（B）

研究期間：2009～2010

課題番号：21700322

研究課題名（和文） 溶液 X 線散乱法による天然解鎖蛋白質の精密構造解析

研究課題名（英文） Accurate structural analysis of natively unfolded proteins by the solution X-ray scattering

研究代表者

関 安孝 (SEKI YASUTAKA)

岩手医科大学・薬学部・講師

研究者番号：30377220

研究成果の概要（和文）：解鎖蛋白質の構造特性を研究するために、ポリペプチド・コンホメーションの新しい生成法を開発した。この方法からの予測と溶液 X 線散乱測定の結果を比較して、高濃度変性剤溶液中の蛋白質の構造特性を解析した。ポリペプチド鎖の平均自乗半径の残基数依存性と Kratky プロファイルの形の解析から、高濃度変性剤溶液は、蛋白質に対して良溶媒であることが確認された。また構造集団が実測データを再現するためには、主鎖二面角が α 領域内にある確率が 20~40% でなければならないことが分かった。

研究成果の概要（英文）：A new method for generating polypeptide conformations has been developed for studying structural characteristics of unfolded proteins. Those of proteins in high-concentration denaturant solution were analyzed by comparing predictions from this method with results from solution X-ray scattering measurement. Analysis of the dependence of mean square radius on the number of residues and the shape of Kratky profiles for polypeptide chains has confirmed that a high-concentration denaturant solution serves as a good solvent. We found also that, for a conformational ensemble to reproduce experimental data, the probability that main-chain dihedral angles are in the α -region must be in the range of 20 to 40%.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009 年度	2,800,000	840,000	3,640,000
2010 年度	600,000	180,000	780,000
年度			
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：分子生物物理学

科研費の分科・細目：情報学・生体生命情報学

キーワード：分子シミュレーション、天然解鎖蛋白質、溶液 X 線散乱法

1. 研究開始当初の背景

近年、生理的条件下で、鎖が解けて特異的な立体構造を形成しない (Natively Unfolded, NU) 蛋白質・領域が、生物機能に重要な役割を果たしていることが明らかになった。更に、ゲノム解析の結果、NU 蛋白質・領域が真核生物に多量に存在することが分かり、これらの構造・機能解析が、今後さらに重要になると予想される。

NU 蛋白質の設計指針や機能発現の分子機構を明らかにするには、アミノ酸配列に基づいた構造特性の解析が不可欠である。そのための手法として、多くの研究実績をもつ変性状態蛋白質の構造解析法が有効である。また、得られた結果の解釈には、NU 蛋白質と変性球状蛋白質の構造情報の比較が必須である。(図 1)

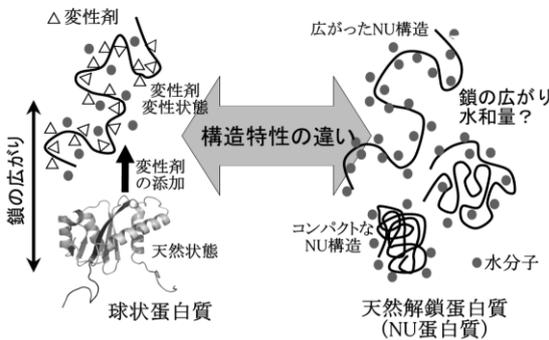


図 1 特異的な立体構造を形成する球状蛋白質と比べて、NU 蛋白質・領域は親水性アミノ酸残基の比率が高い。球状蛋白質の変性構造の解析法を用いて、NU 蛋白質の構造情報を得ることで、両者の構造特性の違いを明らかにする。

高濃度の変性剤により解鎖した球状蛋白質は、小角 X 線散乱法から得られる、平均二乗半径 (R_{sq}) の残基数依存性と散乱プロフィール (散乱強度の角度依存性) の形状、及び多次元 NMR 法による $^3J_{HN\alpha}$ 結合定数から、所謂ランダムコイル鎖であると考えられている。

一方、(ala)₃ などの短いペプチドでは、分光学的測定から主鎖 2 面角がプロリン領域 ($\phi = -75^\circ$, $\psi = 145^\circ$ 付近) に局在していることが示され、変性剤非存在下の解鎖蛋白質ではポリプロリン II ヘリックスが主要な構造だと示唆されている。

これらの情報を統合し、NU 蛋白質の天然状態と球状蛋白質の変性状態の解鎖構造の差とその成因を解明することが強く望まれている。

2. 研究の目的

我々は、解鎖蛋白質の構造特性を解析するため、ランダムペプチド鎖の高速・高精度な分子モデリング法を新規に開発した。このモデリング法は、主鎖 2 面角頻度分布を用いることによって、ランダムペプチド鎖の構造集団の主鎖構造を特徴付けることができる。更に、これとペプチド鎖の原子座標から溶液 X 線散乱強度の散乱角依存性 (散乱プロフィール) を予測する計算法を組み合わせ、鎖状高分子構造解析システムを構築した。

分子モデリング法から生成されたランダムペプチド鎖の構造集団が、実測される溶液 X 線散乱プロフィールを再現すれば、その構造集団の構造特性は、実際の蛋白質の構造特性を反映していると考えられる。

まず、NU 蛋白質と比較するため、天然球状蛋白質の変性状態の 1 つである高濃度変性剤存在下における変性状態の一般的な構造特性、特に溶媒の親和性について解析を実行した。

3. 研究の方法

(1) 主鎖 2 面角頻度分布

ランダムペプチド鎖を生成する際、主鎖 2 面角頻度分布は、"Dali Domain Dictionary 2.0" データベースから抽出したドメイン構造から得た。その基準は、残基数が 50 以上であること、解像度が 0.24nm より良いこと、主鎖 2 面角のコア領域への出現頻度が 94% より高いことである。データベースに登録されている 526 ドメイン中、379 ドメインが上記の基準を満たした。

選ばれた 379 ドメインの全主鎖 2 面角を計算するとともに、"DSSP" による 2 次構造同定を行なった。これらアミノ酸残基名、主鎖 2 面角 (ϕ , ψ)、2 次構造情報 (helix, sheet, turn, coil など) を使い、主鎖 2 面角頻度分布を作成した。

具体的には、2 次構造を規定し、その 2 次構造に含まれる残基の分布を作成した。作成した分布は、all=($\alpha+\beta+t+c$), ($\alpha+t+c$), ($\beta+t+c$), ($t+c$), ($\beta+t$), coil=(c) 分布の 6 種類である。例えば、($t+c$) 分布とは天然構造中の水素結合ターン (DSSP では T が識別子) と 2 次構造外に含まれる残基の 2 面角を集めて分布としたものである。同様に α は、DSSP においてヘリックス (H, G, I) と同定された残基を、 β は、ストランド (E, B) と同定された残基を示す。

図 2 に得られた all 分布を示す。この分布は全アミノ酸残基を使っているが、実際のランダムペプチド鎖を作成する際は、アミノ酸種類ごとに分けている。なお、主鎖 2 面角 ω は 180° に、プロリン残基の ϕ は -75° に固定している。

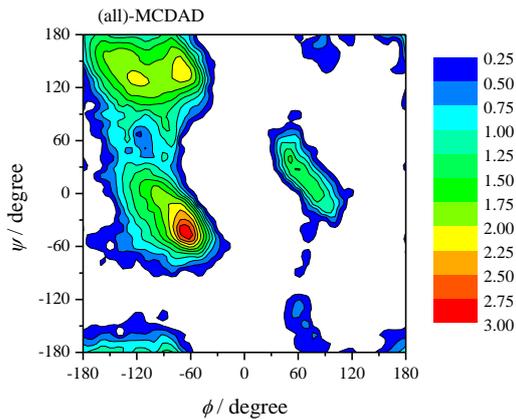


図2 all-分布出現頻度の等高線図

(2) 解析に使用した蛋白質

高濃度変性剤存在下における変性蛋白質の R_{sq} の残基数依存性を解析するため、horse cytochrome *c* (104 アミノ酸残基), bovine α -lactalbumin (123), bovine ribonuclease A (124), hen egg white lysozyme (129), sperm whale myoglobin (153), avian sarcoma virus integrase core (162), dihydrofolate reductase (191), MutY catalytic domain (225), triosephosphate isomerase (250), *EcoRI* endonuclease (276), UDP-galactose 4-epimerase (338), creatine kinase (379)

の計 12 種類の蛋白質の分子モデリングを行った。これらの蛋白質は、高濃度変性剤変性状態において、溶液 X 線散乱法により R_{sq} が決定されている。また、これらの蛋白質はジスルフィド結合やリガンド分子との共有結合を含まない。なお、cytochrome *c* は、apo 状態での R_{sq} 測定がなされているので、分子モデリングでも同様とした。

(3) 溶液 X 線散乱プロファイルの計算

溶液 X 線散乱の散乱強度の散乱角度依存性 (散乱プロファイル) の計算は、以下の式を用いた。

$$I(K) = \sum_{i=1}^N \sum_{j=1}^N f_i(K) f_j(K) \cdot \frac{\sin Kr_{ij}}{Kr_{ij}}$$

ここで、 N は原子数、 $f_i(K)$ は、原子 i の構造因子、 r_{ij} は原子 i と j の距離である。また、 K は散乱ベクトルの絶対値で、以下の式で表される。

$$K = \frac{4\pi}{\lambda} \sin \frac{\theta}{2}$$

ここで、 λ は X 線の波長、 θ は散乱角を示す。

分子モデリングによって構築されたコンホメーション (各原子の空間座標) から 1 分子の散乱プロファイル ($I(K)$) を計算し、それをアンサンブルで平均をとって、実測との比較をした。

4. 研究成果

(1) 新規分子モデリング法の高速度性

我々が開発した新規分子モデリング法には、任意残基から鎖を成長させるように付加させながらコンホメーションを構築するアルゴリズムを採用した。その具体例を図 3 に示す。

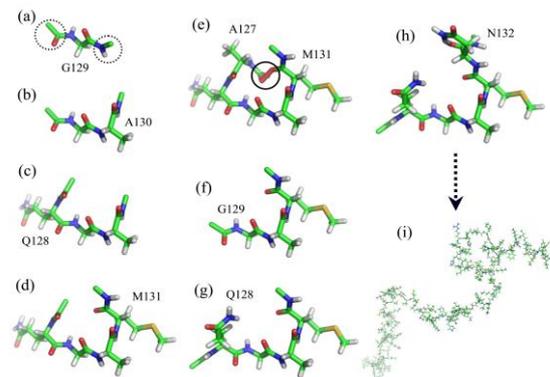


図3 ペプチド鎖成長アルゴリズム

(a) から (i) に向かって、ペプチド鎖を成長させるように残基を付加させていく。(e) に示すように、途中で原子衝突が発生した場合、一旦、残基を削除し、再付加させる。

このアルゴリズムの採用により、残基間に発生する原子間衝突を調べる計算時間と原子間衝突の無い別のコンホメーションへの再構築をする計算時間を最小化することが出来ると期待される。そこで、本法と従来法との計算時間の比較を行った。その結果を、図 4 に示す。

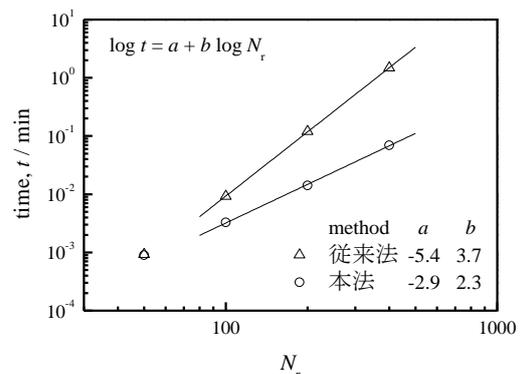


図4 ポリアラニンの残基数と構造生成時間の関係

図4の縦軸は、ポリアラニンのランダムペプチド鎖構造を生成する際の平均時間を示す。横軸は残基数であり、50, 100, 200, 400について計算した。従来法に比べ、本法は100残基以上で計算時間の短縮が確認できる。また、生成時間の対数と残基数の対数には線形関係が認められ、線形解析の結果得られる傾きは図4中に示した。このように、本法で採用したアルゴリズムによって、特に残基数が100残基を超える場合、コンホメーションの生成時間を短縮することが明らかとなった。

(2) ランダムペプチド鎖の構造アンサンブルの多様性

ランダムペプチド鎖モデリング法が生成するコンホメーションの構造多様性を調べるために、アポミオグロビンのランダムペプチド鎖を生成し、その R_{sq} と溶媒露出表面積(ASA)を計算した。主鎖2面角頻度分布のことなる6つのモデルについて、それぞれ100万個のコンホメーションを生成した。それぞれの頻度分布を図5に示す。なお、以下では、allモデルのように生成時に使用した主鎖2面角分布をモデル名とする。

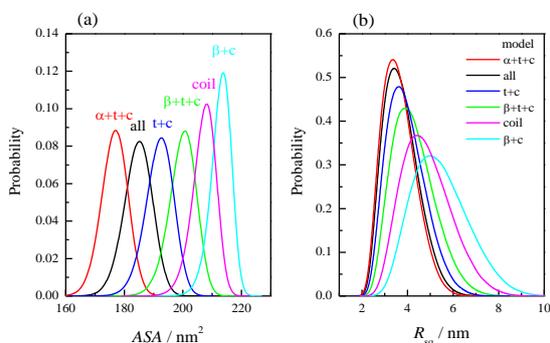


図5 アポミオグロビンのASAと R_{sq} の頻度分布

図5-(a)はASAの、(b)は R_{sq} の頻度分布を示す。また、モデルの違いは線の色で識別している。(a)に示すようにASAの頻度分布は、各モデルで異なるピークを持つとともに、分布の裾では隣り合ったモデル間で重なっていることから、ASAに関して広い構造空間を隈なくサンプルしていることがわかる。

一方、 R_{sq} の頻度分布は、各モデルで異なる分布を持つが、ピーク高さや幅が大きく異なる。また、2.5~6.0 nmの領域は各モデルで共通に取りうる領域である。更に、小さい R_{sq} を示す($\alpha++c$)と(all)の2つのモデルは、ほぼ同一の分布を示す。ここには示していないが、 R_{sq} の平均値を主鎖2面角が α -領域にある割合でプロットするとこの割合が上昇するに伴って平均 R_{sq} 値は減少するが、30%以上の高い α -領域を持つモデルでは、値が収束することが示される。ランダムペプチド鎖の構造特

性として、 R_{sq} の統計平均値には下限値があることを示唆している。

(3) R_{sq} の残基数依存性の実測との比較

12種類の蛋白質についてランダムペプチド鎖の分子モデリングを行った。それぞれの蛋白質に6つの主鎖2面角モデルを、1つのモデルにつき100万個の構造を生成した。溶媒の親和性を解析するため、100万個の R_{sq} 値に対し溶媒和エネルギーから得られるボルツマン因子による重み付け平均を行った。溶媒和エネルギーは、ASAに比例すると仮定し、その比例定数(σ , $\text{kJ mol}^{-1}\text{\AA}^{-2}$)の値を変化させ、実測との比較を行った。 R_{sq} の残基数依存性を図6に示す。

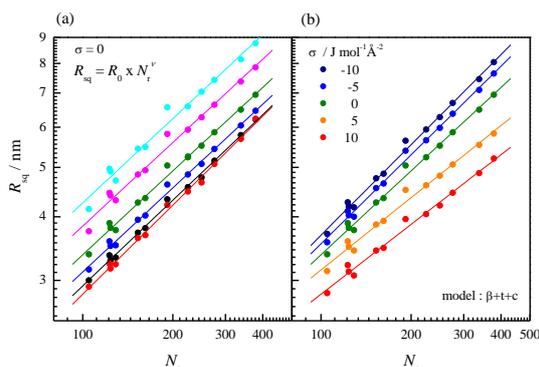


図6 ランダムペプチド鎖の平均二乗半径の残基数依存性

図6-(a)のデータの色は、主鎖2面角頻度分布の違いによる違いを示している。色とモデルの対応は図5と同一である。また、図6-(b)は、($\beta++c$)モデルについて、溶媒の親和性を考慮した場合である。データの色の意味は図中に示した。

図6-(a)から、主鎖2面角頻度分布の違いは指数係数(ν)には現れず、 R_0 のみに現れることがわかる。一方、(b)に示すように、溶媒効果は ν に反映される。正の σ 値である、所謂“good solvent”では、 ν は大きい値を示し、逆に負の σ 値である“poor solvent”では、 ν は小さい値を示す。

これらのランダムペプチド鎖の特性を使って、実験と比較することにより、主鎖2面角頻度分布と溶媒の親和性に関する情報が得られる可能性がある。そこで、 R_0 と ν の2つのパラメータの実測との比較を行った。その結果を図7に示す。図7-(a)の縦軸は ν であり、横軸はモデル構造の主鎖2面角の α -領域にある割合である。また、プロットの色は溶媒和エネルギーにおける係数、 σ を示している。同様に、図7-(b)の縦軸は R_0 を示している。また、高濃度変性剤変性状態の実験データから得られた、 ν と R_0 の値とその推定誤差を、図中に実線と破線で示した。

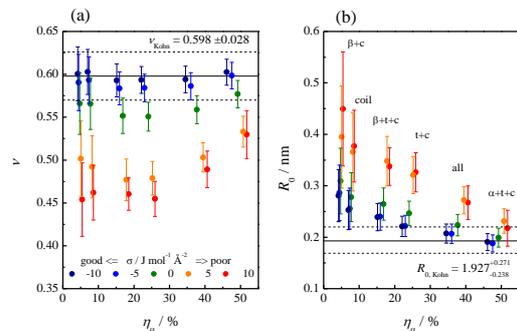


図7 スケーリングパラメータ (ν , R_0) の実測と計算の比較

図7-(a)から明らかなように、実測の ν を再現するためには、負の σ 値を使った重み付け平均が必須である。つまり、実際の高濃度変性剤変性状態の溶液条件は、蛋白質鎖に対して“good solvent”な状態であることを示している。しかし、 ν の値を再現する事に関しては、主鎖2面角頻度分布はどれでも良い。

一方、図7-(b)から、実測の R_0 を満たすためには、主鎖2面角頻度分布は、 α -領域に出現する割合が高い必要がある。この他、アポミオグロビンの散乱プロファイルの解析などから、 α -領域への出現割合は、20~40%と結論される。

このように、我々が開発したランダムペプチド鎖の新しいモデリング法は、溶液X線散乱法と組み合わせることにより、解けた鎖状態の蛋白質の構造特性解析に利用することが出来ることが示された。この方法は、今後天然解鎖蛋白質の構造解析へ有効に利用されると思われる。

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2件)

① Soda, K.; Shimbo, Y.; Seki, Y. & Taiji, M. Structural Characteristics of Hydration Sites in Lysozyme Biophysical Chemistry, 2011, In Press 査読有り

② Nakamura, S.; Seki, Y.; Katoh, E. & Kidokoro, S.-i. Thermodynamic and Structural Properties of the Acid Molten Globule State of Horse Cytochrome c Biochemistry, 2011, **50**, 3116-3126 査読有り

[学会発表] (計3件)

① 関安孝, 溶液X線散乱とNMRの残余双極子結合による解けたタンパク質の構造解析, 第48回日本生物物理学会年会, 2010年9月22日, 仙台

② 関安孝, 解けたタンパク質のNMR残余双極子結合定数の推定法, 第10回日本蛋白質科学会年会, 2010年6月18日, 札幌

③ 関安孝, 分子モデリング法と溶液X線散乱法による天然解鎖蛋白質の構造解析, 2009年10月30日, 徳島

6. 研究組織

(1)研究代表者

関安孝 (SEKI YASUTAKA)
岩手医科大学・薬学部・講師
研究者番号: 30377220