

氏名	小副川 博也		
学位の種類	博士(情報工学)		
学位記番号	生工博甲第81号		
学位授与の日付	平成20年3月25日		
学位授与の条件	学位規則第4条第1項該当		
学位論文題目	蛋白質の二次構造帰属法 ProSSA の応用とアプリケーションの開発		
論文審査委員	主査	教授	横井 博一
			” 西尾 一政
			” 花本 剛士
			” 尾川 博昭
			” 岡元 孝二

学位論文内容の要旨

この研究は、主鎖二面体角の連続性パターンによる二次構造帰属法 ProSSA を使って、蛋白質についての詳細な二次構造の解析を行ない、それらの結果について以下の8章からなる論文構成にわけて論じた。

第1章では、これまでの蛋白質二次構造の帰属法についての問題点を論じ、この論文の位置付けを示した。

第2章では、プロテイン・データバンクファイルを使って、二次構造を帰属したファイルをアーカイブするデータベースを構築した。二次構造帰属法としては、IUPAC の二次構造の定義である「ほかのセグメントと相互作用しない繰り返し構造」にもあてはまる主鎖二面体角の連続性パターンによって二次構造を帰属する ProSSA 法を使った。ProSSA 法は、二次構造解析によく用いられる水素結合様式パターンによる二次構造帰属法 DSSP も内包する。また、ProSSA 法は、コンホメーションの規則性を使って二次構造を帰属するため、 β シート構造に属する β ストランド構造だけではなく、単に伸びた構造も β 構造と帰属することができる。さらに、コンホメーションの規則性を用いることから、規則的な構造と規則的ではない構造の境界を確定できる。このように、ProSSA 法は様々な利点があることを明らかにした。

第3章では、第2章で構築したデータベースから α ヘリックス構造の末端領域における統計的解析の結果を使って、 α ヘリックス構造の末端領域を予測した。従来の二次構造予測は、二次構造の核となる領域の解析データを使って予測していたため、二次構造の末端を予測できなかった。そこで、二次構造の末端領域において、規則的な領域と規則的ではない領域の境界を確定できる ProSSA 法を使って統計的な解析によって得られた α ヘリックスの末端領域の傾向指数を使って、二次構造の形成傾向スコアを算出し二次構造を予測した。その結果、実際の α ヘリックス構造の末端と高い相関を得た。一方で、予測から外れた領域が存在した。すなわち、実際には α ヘリックス構造の末端であるにもかかわらずスコアが低い領域である。このような領域に対しては、二次構造上、不安定な領域であると考えた。つまり、二次構造の末端予測法を、構造既知の蛋白質の配列へ適用することで、二次構造の不安定領域の検索の手法に転換した。そして、不安定領域を検索した結果から、不安定因子となっているアミノ酸残基を探し出し、 α ヘリックスの末端領域の傾向指数と照らし合わせ、スコアが高くなるようなアミノ酸残基の候補を探しだし、アミノ酸置換することで二次構造の安定化を図ることができる。すなわち、二次構造予測法から二次構造設計法への転換を図り、この手法のアルゴリズムを組み込んだプログラムを開発した。このプログラムを使って、実際に部位特異的変異導入による実験で得られた変異蛋白質の立体構造を解析した。その結果、野生型のときのスコアよりも変異型のときのスコアの方が高いときは、熱に対する安定性が向上し、その逆のときは、熱に対する安定性が低下した。つまり、この手法が熱安定化を図る方法として有用であることを示した。

第4章では、蛋白質の円二色性(CD)スペクトルと二次構造含量の相関について調べた。これま

で蛋白質の CD スペクトル二次構造含量の相関については、多くの研究がなされてきた。しかしながら、CD スペクトルと二次構造含量に相関が見られない場合がある。この論文では、その原因を二次構造帰属法にあると考えた。まず、26 種類の構造既知、かつ CD スペクトルデータが存在する蛋白質を選び、ProSSA 法により二次構造情報を取得した。そして、平均的な二次構造含量をもつ蛋白質 1fec を対照とし、他の 25 種類の蛋白質との関係を調べた。その結果、20 種類の蛋白質において、CD スペクトルの 190nm 付近の正の極値と二次構造含量に相関が見られた。この相関比率は、従来の二次構造帰属法では 60%であるのに対し、ProSSA 法を用いた場合では 80%に向上した。以上の結果から、相関があった蛋白質の CD スペクトルと二次構造含量を用いて二次構造含量を 100%とみなした参照 CD スペクトルを算出し、さらに、蛋白質の CD スペクトルより二次構造含量を推定するアプリケーションを開発した。そして、開発したアプリケーションを使って、実際に測定された蛋白質 CD スペクトルから二次構造含量を推定した結果と、X 線結晶データより ProSSA 法を使って算出した二次構造含量の結果を比較したところ、ほかの二次構造推定プログラムよりも高い相関を得た。さらに、実際に測定された CD スペクトル、測定条件、開発したアプリケーションより推定した二次構造含量の結果、また、過去の論文に掲載された CD データ、測定条件などを集め、CD スペクトルのデータベースを構築した。

第 5 章では、プロテアーゼによって消化を受ける配列について、消化配列と二次構造の関係を調べた。蛋白質には、プロテアーゼによって不活性の前駆体から特定の配列が消化を受け活性化するもの、または、その逆のものがある。このとき、消化を受ける配列と二次構造にどのような関係があるかについて述べられることはない。そして、消化配列は、蛋白質の表面にある、あるいはループ領域に存在することがよく知られている。そこで、ProSSA 法を使って二次構造と消化配列の関係を調べた。まず、構造既知であり、プロテアーゼによって消化を受けることがわかっている蛋白質を 7 種類選出した。これら 7 種類の蛋白質に対して ProSSA 法を適用し、二次構造を帰属した。その結果、従来の二次構造帰属法では、消化配列の近くに二次構造は見られなかったものの、ProSSA 法で二次構造を帰属しときは、消化配列の近くに二次構造が存在した。そして、消化配列は、 β ストランド構造の N 末端の境界領域に多く存在した。また、消化配列と同じ配列が基質蛋白質の中に存在する場所がある。しかしながら、その未消化配列は、 β ストランド構造の N 末端の境界領域ではないところに存在した。したがって、プロテアーゼは、基質蛋白質の表面に存在する認識配列が β ストランド構造の N 末端境界領域の存在する場合に、消化するといえる。

第 6 章では、ProSSA 法における β ストランド構造の解析を行なった。ProSSA 法は、シート構造に属する β ストランド構造と、単に伸びた構造である β ストランド構造の区別がつかない。そこで、ProSSA 法の β ストランド構造を、水素結合様式パターンで二次構造を帰属する DSSP 法を利用してシート形成と一本鎖の β ストランド構造に分けて、統計的な解析を行なった。まず、1081 蛋白質に対して、ProSSA 法を適用し、 β ストランド構造を帰属した。さらに、DSSP 法から得られた β シート構造と照らし合わせて、シート形成と一本鎖の二つの β ストランド構造に分類した。そして、それぞれの β ストランド構造の長さに対する頻度を調べた。その結果、一本鎖 β ストランド構造の長さに対する頻度は、単調に減少していた。しかしながら、シート形成 β ストランド構造の長さに対する頻度は、偶数残基ごとにピークを得た。このことから、シート形成 β ストランド構造が比較的両親媒性の構造を形成している場合が多いと考えた。また、シート形成と一本鎖の β ストランド構造の主鎖二面体角 (ϕ , ψ) において、もっともよく現れる角度を使って、モデルペプチドであるアラニル・アラニンを使って、分子軌道計算を行なった。その結果、一本鎖 β ストランド構造の方がシート形成 β ストランド構造よりもエネルギー的に安定していた。つまり、 β シート構造は、ストランド間の水素結合が安定化に寄与しており、シート形成の β ストランド構造では、エネルギー的に不安定であると考えた。

第 7 章では、二次構造解析の基本単位をアミノ酸残基ではなく、ジペプチドとして統計的な解析を行なった。これは、二次構造を考えると、アミド間の水素結合が寄与していることを考えるとアミドを含めたジペプチドを基本単位で考える方が合理的であると考えたためである。また、アミノ酸残基では配列情報を持たないが、ジペプチドは最小の配列情報をもつ。そこで、まず、1081 蛋白質に対して、ジペプチドの数と、ProSSA によって帰属した二次構造の中のジペプチドの数から、400 種

類のジペプチドの傾向指数を算出した。その結果、AC と CA のパラメータが異なるなど、順番を考慮したパラメータを得た。また、 α ヘリックス構造において、傾向指数が 1.5 以上のジペプチドを選び、主鎖二面体角 (ψ , ϕ) の出現頻度を調べた。その結果、ある角度に収束するジペプチドと収束しないジペプチドの 2 つのタイプに分類できた。さらに、モデルペプチドとして、アラニル・アラニンを使って二次構造の中における自由エネルギーを計算した。その結果、 α ヘリックス構造は、 β ストランド構造よりもエネルギー的に安定であった。

第 8 章では、この論文を総括した。

蛋白質の立体構造解析において改めて二次構造解析の重要性を示し、二次構造の帰属法の違いによって解析結果が異なることを、明らかにした。つぎに、二次構造帰属法として ProSSA 法を用いさまざまな二次構造解析を行なった。その結果、 α ヘリックス末端の予測と評価では、予測できなかった末端領域を予測できることを示し、さらに、予測できなかった領域を不安定領域ととらえ安定化するための手法に転換し、その手法が蛋白質の熱安定化の向上を図るための方法として使えることを明らかにした。また、ProSSA 法を用いることで蛋白質 CD スペクトルと二次構造含量との間に、高い相関を得て、さらに相関が得られたデータを使って参照 CD スペクトルを算出し、二次構造含量を推定するアプリケーションを開発した。そして、開発したアプリケーションを使って、実際に測定された CD スペクトルから二次構造含量を推定した。その推定結果は、X 線結晶データから ProSSA を用いて算出した二次構造含量とほぼ一致した。さらに、ProSSA 法を用いることで、プロテアーゼが、配列だけではなく、二次構造も認識していると考えた。ペプチド分子の二次構造研究としては、 β シート構造は、長さに対する頻度の結果より、両親媒性を取りやすいことを示し、さらに、一本鎖 β ストランド構造は、シート形成 β ストランド構造よりもエネルギー的に安定であることや、 β シート構造が水素結合の寄与によってエネルギー的に安定であることを明らかにした。また、二次構造解析の基本単位をアミノ酸残基単位からジペプチド単位にすることで、配列を考慮したパラメータを得ることができた。つまり、この研究では、ProSSA 法が二次構造の研究に対して有用な方法であることを実証した。また、それに関連したアプリケーションを開発した。さらに、二次構造の基本単位としてジペプチド単位を使って、配列を考慮したパラメータを算出した。これらの研究は今後の展開として予測やデザインへの応用が期待できる。

学位論文審査の結果の要旨

本論文に関して、調査委員ならびに公聴会出席者から新しい蛋白質二次構造帰属法である ProSSA 法とそのソフトウェア、蛋白質の二次構造と CD スペクトルの相関性、プロテアーゼの消化配列と二次構造の相関性、さらには、今後の蛋白質の構造解析研究の基本パラメータとなりうるような提案についての質問があったが、いずれも著者の適確な説明によって、質問者の理解が得られた。

以上により、論文調査及び最終試験の結果に基づき、審査委員会において慎重に審査した結果、本論文が博士（情報工学）の学位に十分値するものであると判断した。