

専攻	材料システム 工学	学籍番号	869701	指導教官氏名	高橋由雅 宮下芳勝
申請者氏名	唐澤肇				伊藤健兒
					亀頭直樹

論文要旨

論文題目	環境記述ベクトル法にもとづくタンパク質の二次構造予測に関する研究
------	----------------------------------

(要旨 1,200字以内)

タンパク質の立体構造の決定は、現在、そのほとんどがコストと時間のかかる X線結晶解析法によって行われている。その一方で、タンパク質のアミノ酸配列からその構造を予測するという問題に対し、さまざまな試みがなされてきているが、いまだに十分な手法は現れていない。本研究では、これまで提案された種々の二次構造予測法の問題点等を考慮しながら、タンパク質アミノ酸配列上での各アミノ酸残基の物理化学的周辺環境の新たな記述法を提案するとともに、これにもとづくパターンクラス分類の観点から、二次構造識別へのパターン認識の応用を試みた。また、二次構造予測への応用の可能性についても検討を行った。

まず、タンパク質中の各アミノ酸残基の配列上での物理化学的周辺環境を記述するために、配列上の注目するアミノ酸残基(Target Amino acid residue;TA)についてN, C両末端側各5残基を考慮し、これら周辺アミノ酸残基をその対応するアミノ酸の持つ5種の物理化学的特性値(分子量、 $pK_{\alpha}\text{-COOH}$ 、 $pK_{\alpha}\text{-NH}_2$ 、等電点、側鎖の疎水性のフラグメント値 π)を用いて特徴づけた。これにより、TAの周辺環境を50個の要素からなる多次元ベクトル(環境記述ベクトル)として記述できる。このような多次元ベクトル表現を導入することにより、これまで

の多くの二次構造予測の研究で用いられてきたアミノ酸残基名による残基配列の統計解析とは異なった残基配列環境の定量的（数值的）な記述が可能となった。次に、この環境記述ベクトルにもとづき、3つの二次構造(Helix, Sheet, Turn)クラスに対する識別モデルの開発を判別分析法によって行った。また、ランダムコイル残基に関しては、特定のクラスモデルを想定することが困難であることから、3つの二次構造型ランダムコイルへのサブクラス化を図り、それぞれのサブクラスに対して識別モデルの開発を行った。この方法によりタンパク質中のアミノ酸残基を4つの状態(Helix, Sheet, Turn, Random coil)に同時識別することが可能となった。さらに、このようにして得られた識別モデルをもとに、モデルの開発に使用した97タンパク質に対して、配列上のアミノ酸残基の二次構造認識実験を行ったところ、97タンパク質のトータルで約51.3%の認識率を得ることができた。

以上、本研究ではタンパク質配列アミノ酸残基の環境記述に対する新たな定量的記述法を提案するとともに、本法のタンパク質二次構造予測問題における有用性を明らかにした。このことは、今後のタンパク質アミノ酸配列と高次構造との相関解析等の問題に対する新たなアプローチを示すものと確信する。