

論文審査の結果の要旨

申請者氏名 石田 貴士

現在ではタンパク質の機能解析やタンパク質をターゲットとした創薬にその立体構造情報が欠かせないものとなっている。タンパク質の立体構造予測はアミノ酸配列から立体構造を計算によって推定するものであり、実験的な立体構造決定を補助するだけでなく、実験的手法の限界に捉われないことから、従来では決定が難しかったタンパク質の構造を明らかにすることも期待されており、新規の人工的なタンパク質のデザインにおいても力を発揮すると考えられている。通常タンパク質の立体構造予測はそのタンパク質が安定な構造をとることを前提とする。しかし、構造ゲノミクスの進展と共に、安定な構造をとらない不安定な disorder 領域が数多くのタンパク質に存在することがわかってきた。また、これらの disorder 領域は単に一定の構造をとらない領域というだけでなく、分子認識などの機能を持つものが存在することが明らかとなってきた。そのため、この disorder 領域を同定することは、立体構造の決定が不可能な領域を同定するだけでなく、そのタンパク質の機能の推定にも重要な意味をもつと考えられる。

本論文は構造未知のタンパク質に対し、まずそのアミノ酸配列からタンパク質の disorder 領域を予測することでそのタンパク質が安定な立体構造を形成しうるかを判定し、さらに安定な立体構造をとると考えられる領域に対してその立体構造の予測を行うという一連の予測手法の開発を試みている。

第 3 部では、機械学習を用いたアミノ酸配列からのタンパク質 disorder 領域の予測手法について述べている。第 3 章において、クリスタルパッキングの影響を考慮することで機械学習に用いる訓練セットの質を向上させ、また従来の disorder 予測手法で用いられてきた局所的な配列情報に加えて、大域的な配列アライメントの情報を入力情報に加えることで予測精度の向上を図った新たな予測手法を提案した。第 4 章ではこれらの予測手法の性能評価を行った。その結果、局所配列情報からの予測精度が改善したことを示し、また既知構造との配列アライメントからの予測を加えることでより感度の高い予測を実現し、予測立体構造の情報を利用することで予測の特異度を向上させたことを示した。

第 4 部では、アミノ酸配列からのタンパク質立体構造の予測手法について述べている。タンパク質立体構造予測手法の中でも、新規のフォールドのタンパク質についても予測が可能な最小エネルギー構造を探索する方法について、その一手法であるフラグメントアセンブリ法に基づいた手法の開発を行った。第 8 章ではタンパク質の各残基の埋もれ度を配列から予測された埋もれ度と比較することで構造の評価を行う、予測埋もれ度を用いたポテンシャルを提案し、また、そのためにその残基の周辺の残基やマルチプルアライメント

による進化的な情報などから、機械学習の手法の一つであるサポートベクター回帰アルゴリズムを利用して残基の埋もれ度を予測する手法を提案した。第 9 章では、構造サンプリングの効率を改善するために、構造サンプリングに用いるフラグメントの長さを各部位毎にその候補フラグメントとの配列類似によって最適化するアルゴリズムを提案した。第 10 章ではこれらの手法について、デコイセットを用いて性能評価を行った。その結果、予測埋もれ度を利用した新たなポテンシャルが従来のポテンシャルに比べ天然構造の識別能を向上させ、構造フラグメントの長さを最適化するアルゴリズムが構造サンプリングを効率化することを示した。また、それらの結果として予測システム全体としても、従来に比べより高い予測性能を示し、より天然構造に近い高精度の予測構造が得られることを示した。

本論文はアミノ酸配列情報から disorder 領域と立体構造を予測する手法について改良を行い、その両者について予測精度を向上させることに成功している。これらの成果は従来では実験的手法によって決定する必要のあったタンパク質の構造についての情報を、そのアミノ酸配列情報のみからより低コストで得ることへの可能性を示した。よって審査委員一同は本論文が博士（農学）の学位論文として価値あるものと認めた。