

論文審査の要旨及び審査委員

(2, 000字程度)

報告番号	甲 第 29 号		氏 名	安保 勲人		
論文審査 審査委員	氏 名		職 名	氏 名		職 名
	主 査	本間桂一	教授	委 員		
	委 員	太田元規	教授			
	中村建介	教授				
		佐川孝広	准教授			
		福地佐斗志	教授			

タンパク質はアミノ酸がペプチド結合により数珠状につながった分子であり、生体中ではこの紐状の分子が立体的に折り畳まれて機能する、という考え方が常識とされてきた。しかし、今世紀に入り細胞中で立体構造を形成しない領域、天然変性領域が機能を持ち、天然変性領域を持つタンパク質、天然変性タンパク質が転写制御やシグナル伝達といった機能に深く関わっていること、天然変性タンパク質の異常が重篤な疾患を引き起こすことが指摘されるに至った。天然変性領域中には相互作用パートナーと結合する数残基から数十残基の比較的短い領域が存在し、この機能部位を介した相互作用が天然変性タンパク質の機能に重要な役割を果たしている。天然変性領域および機能部位は、コンピュータを用いた予測プログラムによってアミノ酸配列より予測可能だが、天然変性領域予測が実用レベルの測精度に達する一方、機能部位の予測精度は充分ではない。実験的証拠が担保された機能部位データを提供するデータベースが存在するが、構造領域や天然変性領域のデータ数と比較して、機能部位データの数は圧倒的に少ない。このデータ数の不足が機能部位予測を難しくしている原因の1つとしてあげられる。

本論文は、天然変性領域中の機能部位は天然変性領域に存在する領域でありながら、構造領域的性質を持つことに着目し、天然変性領域中の機能部位予測プログラム NeProc の開発について述べたものである。論文は6章からなっており、第1章では天然変性タンパク質の構造や機能を概観し研究の意義を述べている。第2章では、既存の天然変性領域予測プログラムや天然変性領域中の機能部位予測プログラムについて、アルゴリズム及び特徴を概観している。これらを踏まえ、第3章ではNeProcの開発手法を詳しく述べた後、既存プログラムとの性能比較を行っている。性能比較において、NeProcは機能部位データを学習に用いていないにも関わらず、機能部位データを学習している既存の予測プログラムを上回る予測精度を達成している。特に、これらの機能部位の多くが分布する10残基から50残基の範囲の機能部位に対する予測精度が高いことを報告している。第4章では、未だ検証例が少ない天然変性領域中の機能部位データを補うため、タンパク質データベースの情報を利用し未発見の機能部位候補を選別した上で、予測精度の検証を行っている。この結果も第3章同様に、既存プログラムを上回る性能を見せている。第5章では、NeProc及び既存プログラムをヒトの全タンパク質に適用することで、NeProcや他のプログラムの予測傾向を解析すると共に、NeProcの改善すべき点も見出している。第6章では、全体の結果のまとめ及び今後の展望について述べている。

博士学位予備審査において、統計検定について、第4章に利用された生物種について、開発手法の記述について、背景の記述について、論文の体裁について、用語について、など、多様な意見が述べられた。これらの意見に対し、統計検定をNeProcとその他のプログラムの全ての組み合わせについて実施、第4章でヒトタンパク質を対象に未発見の機能部位候補を解析していたが、ここにヒト以外のモデル生物種を追加、第3章の予測モデルの構築の説明をより詳細にし、その補助となる図も追加、第1章において、より詳しい背景の説明及び図の追加、論文の体裁を整えると共に論文全体の構成が理解しやすくなるような図の追加、不明瞭な用語の定義の明確化、など本論文では修正が行われた。本研究は、構造領域を学習することで天然変性領域中の機能部位の予測が可能であることを示している。また、比較的単純なニューラルネットの組み合わせで、高い予測性能を発揮している点もユニークである。この論文は、既知の天然変性領域中の機能部位例が十分ではないという予測モデル作成の困難を克服する可能性を示す興味深い結果を記している。

以上、博士論文の研究内容と英語による公表論文を含む参考論文の公表結果、審査における審査委員との質疑応答の結果等から総合的に評価し、博士学位論文として合格と判断した。