

論文審査の結果の要旨

氏名 ハリー アムリ ムサ

本論文では、タンパク質の天然変性領域における化学的組成の進化的な保存と、その性質に基づく機能分類の可能性について述べられている。

タンパク質の天然変性領域 (**intrinsically disordered regions in proteins**) の厳密な定義は容易ではないが、概念的にはタンパク質中で安定的な形では一定の三次構造をとらない領域のことであり、生体内で様々な生物学的機能を果たしていることが知られている。しかし、その領域のアミノ酸配列が進化的に良く保存されていないことが多いため、どのようにしてその機能を維持できているのかが謎であった。そこで論文提出者らは、実験によって決定された、もしくはアミノ酸配列から予測された天然変性領域 (比較的信頼性が高いと言われている) をヒトプロテオームから抽出し、それぞれに対して、7種類の真核生物のオーソログ遺伝子産物の対応領域との配列保存性を調べてみた。その結果、天然変性領域の中には、アミノ酸配列レベルで保存がある程度認められるもの (**HR** グループ) もあるが、そうでない場合でも相当数のものは、その領域のアミノ酸化学的組成 (正荷電 (塩基性)、負荷電 (酸性)、極性、疎水性、特殊 (グリシンとプロリン)) のレベルでは、有意に進化的保存が観察された。以後、**HR** グループには含まれない天然変性領域のうち、この組成レベルでの保存が見られるグループを **HTC** グループ、見られないものを **LTC** グループと呼ぶことにすると、これら3つのグループの領域に好まれる (または避けられる) アミノ酸の種類の中に明瞭な違いがみられた。たとえば、配列保存の **HR** グループではチロシンが好まれるとか、組成保存の **HTC** グループではアルギニンが避けられ、グルタミン酸やシステインが好まれるとか、どれも保存されない **LTC** グループではアラニンが好まれ、アスパラギンやイソロイシン、フェニルアラニンが避けられる等である。また、これらの領域のアミノ酸配列中の存在位置にも若干の傾向の違いが見られた。例えば、**LTC** 領域は他のグループと比較すると、**N** 末部に存在する確率が高く、末端部以外の中央部に存在する確率は低いなどの傾向が観察された。**Gene Ontology** による特定機能の偏りを調べたところ、**HR** グループはホメオボックスドメインや **Zn** フィンガードドメインなどを含む転写制御関係のタンパク質を多く含んでいた。これに対して、**HTC** グループには **RNA** ヘリカーゼを含む **DNA** 複製や修復関係タンパク質が多く含まれており、**ATPase** 活性やヌクレアーゼ活性を持つものに富んでいた。最後に **LTC** グループはイオン結合活性で特徴付けられるものが多く、ペプチダーゼを多く含んでいることがわかった。すなわち、天然変性領域をその進化的保存性レベルで分類すると、配列中の位置の傾向、アミノ酸組成、機能面などで違いが観察された。

論文提出者はさらに、各天然変性領域を上述の5つの性質 (正荷電など) のうちでどの性質が相対的に突出しているかによって、5つのクラスターに分類した。それぞれの

量比は、正荷電 17.5%、負荷電 16.5%、極性 26.2%、疎水性 24.1%、特殊 16.4%であった。それぞれのクラスターと、その領域がアミノ酸配列中に存在する位置（N 末部、中央部、C 末部）とを組み合わせると、興味深い機能的特徴が見られた。たとえば、C 末部にある正荷電領域は mRNA プロセッシング関連タンパク質に多いが、中央部にある正荷電領域はクロマチン組織化関連タンパク質に多く見られた。また、負荷電領域が N 末部にあればそのタンパク質はアミノ酸輸送と関連するものが多く、中央部にあればキナーゼ活性制御と、C 末部にあればタンパク質の折れ畳みや小胞輸送との関連性がそれぞれ強かった。これらの結果は、天然変性領域がその構造柔軟性や機能を進化的に保存してこられたのは、おおざっぱなアミノ酸組成（化学的組成）を保存してきたことが大きい可能性を示唆している。また、この化学的組成は、一次構造上の位置情報や保存の度合いの情報等と組み合わせると、天然変性領域の機能推定や、一般性のある分類法として使えるのではないかと提案している。

なお、本論文は、若林俊一、中井謙太、アシュウィニ・パティルとの共同研究であるが、論文提出者が主体となって分析及び検証を行ったもので、論文提出者の寄与が十分であると判断する。

したがって、博士（生命科学）の学位を授与できると認める。

以上 1,840 字