

好熱菌を用いた三残基アミノ酸フラグメントデータベースの構築と解析

黒田 研究室

学籍番号：06251069

町田 恵久海

【背景・目的】

タンパク質の熱安定性は、タンパク質を工業的・産業的に応用する際の大きな課題である。一般的にタンパク質は、それを構成するアミノ酸残基による分子内・分子間での相互作用によって、ポリペプチド鎖が折りたたまれて立体構造を形成し、機能を発揮する。多くのタンパク質は高温環境で熱変性を起こして機能を失ってしまうのに対し、好熱菌由来のタンパク質は高温環境でも変性しない。そのため、好熱菌由来のタンパク質は高温環境でも機能を失わず工業的価値が高い。好熱菌は他の種と比較して、DNAにおけるCG含有量やタンパク質におけるアミノ酸組成に特徴があることが知られている。しかし好熱菌由来のタンパク質を構成するアミノ酸の局所的な構造的長は未だに得られていない。よって本研究では、アミノ酸の局所構造の解析を目的として好熱菌と非好熱菌（常温菌と真核生物の2種）を由来とするタンパク質データセットを作成した。このデータセットから、構造単位を持ちかつサンプル数が最大に得られるアミノ酸三残基フラグメントのデータを抽出し、解析を行った。

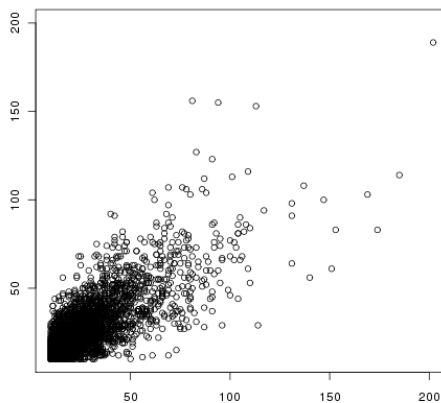
【方法】

PDB（タンパク質構造データベース）から好熱菌・常温菌（好熱菌を除いた真正細菌）・真核生物を由来とするタンパク質のデータを収集し、データ群ごとに配列ファミリーに分割し、それぞれの代表配列を選択した。さらに、好熱菌－常温菌、好熱菌－真核生物の組み合わせで代表配列同士を比較し、互いに類似するアミノ酸配列が存在するもののみを選出し、解析に用いるデータセットとした。得られたデータセットに対し、三残基アミノ酸フラグメントのアミノ酸組成・二面角（PHI・PSI）・溶媒露出表面積の平均値を、好熱菌－非好熱菌（常温菌・真核生物）とで相関係数を算出し、それぞれの相関係数を同様に一残基で算出した場合と比較した。また、得られたフラグメントから統計的手法を用いて、好熱菌由来のタンパク質の予測を行った。

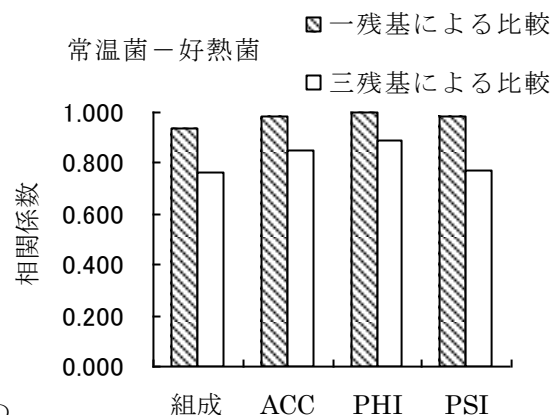
【結果・考察】

データセットに含まれるタンパク質の配列は、好熱菌－常温菌では548ずつ、好熱菌－真核生物では300ずつ得られた。三残基アミノ酸フラグメントの組成・二面角・中央残基の溶媒露出表面積の平均値を取り相関を取った（図1）。またこれらの相関係数は、一残基のアミノ酸で同様に行って得られる相関係数より、どのパラメータにおいても相関係数が減少していた（図2）。これより一残基アミノ酸よりも三残基アミノ酸フラグメントのほうが、好熱菌の特徴をより強く示すことがわかった。

結論として、好熱菌の三残基アミノ酸フラグメントは、組成・溶媒露出表面積・二面角において非好熱菌とは違う分布を持つことがわかった。今回構築した好熱菌由来のタンパク質のデータベースは、熱安定性の高い構造をもつ三残基アミノ酸フラグメントに特化したデータベースであるといえる。



（図1）好熱菌－常温菌による三残基フラグメントの組成。横軸が好熱菌のフラグメント・縦軸が常温菌のフラグメントを表す。



（図2）常温菌－好熱菌間の相関係数の比較