2SC2

膜タンパク質の構造認識、分類

○諏訪 牧子（ヘリックス研 第二研究部門）

多くの生物で完備したタンパク質セット（プロテオーム）が得られるようになったが、これらを立体構造の観点から整理することは極めて重要である。タンパク質の立体構造形成や機能発現の一般的な法則性に対して本質的な情報が得られるからである。膜タンパク質は、プロテオーム中で20～30%も存在し、創薬等の観点からも重要であるにも関わらず、立体構造が殆ど解明されていないため、そもそも立体構造分類の対象にはならず、ファミリーやフォールド数が未だ不明である。今までに、膜貫通ヘリックスの予測本数で分類が試みられた例もあるが、同一本数中で異なる構造、機能ファミリーを分類するまでには至っていない。本研究では、アミノ酸配列情報のみから膜タンパクの立体構造の差異を認識して分類する試みを紹介する。私達はこれまで、ヘリックス間の極性相互作用が立体構造と相関性が有ることを示してきた。とすれば、立体構造は判らなくても極性表面で特徴付けたヘリックスのセットを比較すれば、元の立体構造の類似性が認識できる。この考えを基に、ヘリックス周囲の極性表面の形を理論級数に近似し、表面の形を反映する3係数を得る。タンパク質の構造をN本のヘリックスの係数が一次構造上の順番に従って並んだ6N個の疑似座標($N \times 2$ 末端側 $\times 3$ 係数)として記述し、構造間の類似度を疑似座標間の平均二乗距離として定義した。これを、立体構造既知の膜タンパク質同士、あるいは、実験から関係が保証されている7本膜貫通ヘリックス型受容体に適用したところ、高感度、高選択性を保って同一構造を認識できた。これは、配列しか判らない膜タンパク質に対して、信頼度が高い立体構造認識、分類ができる可能性を示唆している。現在、この方法でゲノム配列が決まったモデル生物に対し、7本膜貫通ヘリックス型受容体の分類を試みているので、その結果も紹介する。

M. Suwa : Structural recognition and classification of membrane proteins

2SC4

ゲノム配列からのタンパク質折りたたみ

○高田 彰二（神戸大学理学部）

ゲノム配列からのタンパク質折りたたみ神戸大学理学部 高田彰二例えは *M. Genitalium* ゲノムのコードしているタンパク質の25～35パーセントは、構造データベース中のタンパク質との配列ホモロジーによりその立体構造を特定できる。膜タンパク質を除く残り40～55パーセントのタンパク質構造を配列情報のみから予測することは、実用的見地から言ってポストゲノム時代の最重要課題の一つであるとともに、生物物理学における理念的な意味でも大きな挑戦である。我々は、(広い意味で)この問題に対し、いくつかの戦略を並行して進めている。1) フォールディング機構を理解し、それを立体構造探索に生かすこと。天然構造を見出すことは、確かに難問題であるが、細胞内でタンパク質は実際にこれを行っている。フォールディング機構を理解し、真似ることが、立体構造探索に役立つ。2) 物理化学的手法による立体構造予測。タンパク質水溶液系にはたらく物理化学的相互作用を、できる限り忠実に取り込んだ粗視化モデルを構築し、それによって構造予測を行う。3) 進化論的情報を加味した立体構造予測。構造データベース中のタンパク質とグローバルな配列相関性が見出せない場合でも、局所的なフラグメントに相関性をもつものが(殆ど)いつも見出せる。2)のモデルに、このデータベース中のフラグメント情報を生かして、構造予測の精度を上げる試みを紹介する。4) 逆フォールディングとしての配列設計原理の研究。構造予測のモデルと同一モデルで、ある鋳型構造にフォールディングするアミノ酸配列をデザインする。

S. Takada : Protein folding simulation for genome sequence

2SC3

タンパク質モジュールに着目したゲノム機能予測

○由良 敬、郷 通子（名大・院理・生命理学）

現在30以上の生物種で全ゲノムが決定されている。しかし、一番重要な情報のひとつであるコードされているタンパク質の機能はまだわからない場合が多い。現在までに、コードされているタンパク質の約半数の機能しか推定されていない。ORFの機能を推定する方法の開発が急がれている。既存の方法は、大きく3種類に分類することができる。1)すでに機能がわかっているタンパク質のアミノ酸配列全体とORFにコードされているタンパク質とのアミノ酸配列の類似度検索、2)機能既知の類似アミノ酸部分配列から構築されたモチーフによる機能部位検索、3)ORFにコードされているタンパク質の立体構造を予測または決定し、立体構造データベースから類似立体構造をもつタンパク質の検索である。上記方法の中には、タンパク質の部分構造類似性から機能を推定する方法が存在しない。そこで我々はタンパク質の部分構造であるモジュールを用いてゲノムの機能予測をおこなっている。立体構造既知のタンパク質において、金属などの補因子およびDNA/RNAとの相互作用部位の分布を見ると、機能を担う部位は数少ないモジュールで構成されていることがわかった。これは機能をモジュールにアサインすることができることを意味する。これらの機能を持つモジュールを立体構造で分類すると、各機能ごとに類似の立体構造をもついくつかのグループに分類することができる。しかし同じグループに分類されるモジュールのアミノ酸配列には明確な類似性は見られない。我々は、立体構造および機能の類似性を3Dキーノートと命名したアミノ酸配列のパターンに変換することで、これらのモジュールと類似のモジュールをゲノム配列から見だし、機能を推定する方法を新たに開発している。

K. Yura and M. Go : Genome function prediction based on protein modules

2SC5

ゲノム配列に対する蛋白質の構造、機能予測

○水口 賢司（Cambridge 大 Biochemistry (UK)）

Fully sequenced genomes of various organisms offer a rich source information about molecular and cellular processes. A significant fraction of genome sequences, however, show either no similarity, or similarity only to unknown, uncharacterised proteins, giving little information about the function of these proteins. To tackle this problem, we try to exploit the knowledge of protein three-dimensional (3D) structures and identify distant evolutionary relationships, which are often difficult to recognize by conventional sequence comparison methods. Our new fold recognition method FUGUE (<http://www-cryst.bioc.cam.ac.uk/fugue/>) can associate genome sequences with known structures by utilising the structural environment-specific amino acid substitution tables and structure-dependent gap penalties derived from the alignments in the HOMSTRAD database (<http://www-cryst.bioc.cam.ac.uk/homstrad/>). Comprehensive benchmarking results demonstrate that FUGUE outperformed many of the most widely used fold recognition programs. FUGUE and other techniques are now being used to associate sequences from various genomes with distantly related homologues with known 3D structure and make inferences about function. Several examples from the genome of the fruit fly *Drosophila melanogaster* will be discussed.

K. Mizuguchi : Protein Fold Recognition and Functional Assignment for Genome Sequences