B105

単体・複合体の立体構造変化の解析による分子認識機構解析

○加藤 雅樹¹、輪 博²、木下 賢吾¹、木寺 詔紀¹ (¹横浜市立大・院総合理・生体超分子、²早大・社会科学)

蛋白質の立体構造と機能の相関を解明することは構造ゲノム科学の最重要課題の一つであるが、そのためには分子認識機構の解明が不可欠である。古典的な考え方として“鍵と鍵穴モデル”が一般的であるが、現実の蛋白質は柔らかく、外部からの刺激に応じて構造変化を引き起こすことの重要性も指摘されてきた。つまり、分子認識を理解するためには分子認識に伴う蛋白質のダイナミクスを考えることが必要となる。この問題に対してこれまで、分子動力学などの演繹的な手法が多く用いられてきた。しかしながら、これらの方法では実際の蛋白質内で起こっている構造変化の時間オーダーに相当する計算を行うことは非常に困難であり、ましてや数多くの蛋白質について系統的、網羅的に行うことは不可能である。そこで我々は、(1) X線結晶解析により立体構造が解かれているすべての蛋白質の中から蛋白質単体、リガンド分子または他の蛋白質との複合体がともに解かれている PDB のエントリを収集してデータベースを作成し、(2) 単体と複合体の構造の違いを比較することで、この問題にアプローチしていることを前回報告した。今回は、露出表面積との関連について報告する。また、構造変化を考えるときにドメイン単位で動いているケースが多いことから、様々なグループが動的ドメインの決定法を模索している。今回我々は、Delaunay 四面体を用いた動的ドメイン決定法を開発した。詳細については、当日報告する。

M.Kato, H.Wako, K.Kinoshita and A.Kidera : A study on molecular recognition mechanism by analyzing structural changes between monomer and complexed proteins

B107

Superfamily を超えた蛋白質構造の進化的類縁関係 -Fold-based superfamily-

○松田 敬子¹、川端 猛¹、木下 賢吾²、西岡 孝明³、郷 信広¹ (¹奈良先端大・情報、²横浜市大・院・総合理、³京大・院・農)

蛋白質のスーパーファミリーの分類はアミノ酸配列、グローバルフォールド、部分構造または機能などの特徴を総合的に判断することでなされている。我々はグローバルフォールドと機能のラフな類似性を考慮することで、既存のスーパーファミリーをより拡大定義できるかどうか調べた。はじめに、それぞれの蛋白質のドメインの β シートのトポロジーを“P2314”等の簡潔な記号で表現した。進化の過程において β シートの両端にストランドが1本付加・削除されると考え、 β シート間の類縁関係 (“kinship”) を構造が変化するのに必要な進化的イベントの数として定義した。この kinship に基づき、ドメインの類縁関係のネットワークを構築した。このネットワークと分子機能の関係を調べるため、ATP 結合蛋白質について解析を行なった。ATP 結合蛋白質の機能をリン酸基またはアデニル基の転移と結合エネルギーの変化に注目して5つの“core chemical function”に分類し、ネットワーク上にマップした。解析の結果、ネットワーク上に分布したドメインはフォールド・core chemical function とともに分離した20個程度のクラスターを形成した。これはATP 結合能力を持つ蛋白質は進化の過程において20個程度独立に出現したことを示唆する。

K.Matsuda, T.Kawabata, K.Kinoshita, T.Nishioka, N.Go : Finding evolutionary relations beyond superfamilies: Fold-based superfamilies

B106

蛋白質・リガンド相互作用データベース、ProLINT、の開発と QSAR 解析

○北島 浩司¹、Ahmad Shanda r¹、久保寺 英夫²、砂田 真二²、皿井 明倫¹ (¹九工大・情報工学科、²三菱ウエルファーマ株式会社)

蛋白質の機能にとって重要なリガンドとの相互作用のメカニズムを理解するため、蛋白質とリガンドの相互作用に関するデータベース開発と、それらのデータを用いた構造・活性相関 (QSAR) の解析をすすめている。これまでに、医薬品開発の標的に数多く取り上げられる蛋白質群である kinase 及び protease について、相互作用データとして2万件余りを文献から収集した。データベースシステムは、文献情報データベースシステム、3Dリガンドデータベースシステム、複合検索システムから成り、検索は、文献情報、キーワード、リガンドの部分化学構造などから行うことができる。現在、簡単な Web 検索インターフェイスをとおして一部のデータをインターネット上に公開している。ProLINT のデータを用いた QSAR 解析ではまず、リガンドの分子量、Surface Area、logP、疎水性などの記述子についてプロファイル解析を行い、データベース中での分布と IC50 などのリガンドの結合活性との相関を調べた。また、これらの構造記述子を用いてニューラルネットを使いリガンドの結合活性を予測することを試みている。さらに、ProLINT 中のリガンド分子のフラグメントとそれに対応するさまざまなファミリーに属する kinase の相関を調べたところ、フラグメントの頻度はファミリーごとに特徴的なパターンを示した。これは、フラグメントパターンが kinase ファミリーのリガンド特異性を特徴づけるフィンガープリントとして利用できる可能性を示唆した。また、フラグメントを構造記述子として加えることにより、ニューラルネットによる結合活性予測の精度が大きく上がることが確認された。

K. Kitajima, S. Ahmad, H. Kubodera, S. Sunada and A. Sarai : Protein Ligand Interaction Database ProLINT and Its Application to QSAR Analysis

B108

タンパク質の de novo 立体構造予測：ポテンシャル関数の最適化

○藤田 佳見¹、千見寺 淨慈²、高田 彰二^{1,2,3} (¹神戸大院・自然科学、²神戸大・理学部、³科技団さきがけ)

タンパク質立体構造の予測には、構造空間の十分なサンプリングと天然構造がエネルギー的に最も低いと認識できるポテンシャル関数とが必要とされる。現在の計算機で多くのタンパク質が折り畳みに要する時間のシミュレーションを行うことは未だ困難である。そこで、我々は、アミノ酸の側鎖を一つの球で近似した粗視化モデルを構築し、構造予測シミュレーションを行ってきた。予測に失敗した予測構造と天然構造とを比較した結果、ポテンシャル関数のうち、特に局所相互作用及び水素結合を改良する必要があることが分かった。本研究ではこの問題に、物理化学的知見と既知構造データベースを利用した情報科学的手法を組み合わせることで取り組む。

局所相互作用のポテンシャルに関しては、関数形を構造データベースの2面角の頻度から導いた。さらに、アミノ酸ごとの2次構造形成傾向を表すために、非天然構造を生成し、天然構造が最安定となるようにパラメータを最適化した。同様に、水素結合ポテンシャルに関しても、水の効果を取り入れた環境依存性項を構造データベースから決定した。次に、各エネルギー項の全エネルギーに対するバランスを最適化した。まず、PDB 全エントリの中で、互いに25%以上配列相同性のないタンパク質の非冗長セットを training set とし、それぞれのタンパク質について非天然構造を用意した。最適化は training set の全てにおいて、非天然構造アンサンブルに対する天然構造のエネルギー偏差値が大きくなるように行った。この方法から導かれたポテンシャルの精度を調べるため、構造予測ベンチマークテストを行った。構造探索法にはフラグメントアセンブリ法を用いた。その結果、80残基以下の幾つかのタンパク質で、トポロジー予測に成功した。

Y.Fujitsuka, G.Chikenji and S.Takada : De novo protein structure prediction: Optimization of potential function