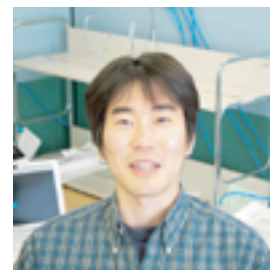


タンパク質の立体構造予測

分子設計チーム
(研究員)

トミイ ケンタロウ
富井 健太郎



学生時代、タンパク質の巻き戻り実験をしていた私は、あたかも形状を記憶しているかのように構造を自己復元できるこの非常に小さな物質に大層興味をひかれ、そのような物質を部品として生物が出来ていることに、ある種の感動を受けたことを覚えています。現在では、当時とは比較にならない量のゲノム配列データの解読が進み、生物を構成する部品のカタログは充実する一方です。それが、計算機によるタンパク質の構造／機能推定の必要性を増大させるだけでなく、実は、計算機を用いた配列情報からの立体構造予測に恩恵をもたらしています。

予測したいタンパク質と構造既知のタンパク質との配列類似性の検証は、立体構造推定の根幹をなす作業です。強力な類似配列検索手法であるPSI-BLASTは、局所配列比較法を予測配列についてのプロファイルと逐次的に組み合わせることで配列情報を有効に活用し、ペアワイズの配列比較法をはるかに凌ぐ構造予測を可能にしました。ここでプロファイルとは、相同であると推測される類似配列タンパク質群での各残基位置における20種類のアミノ酸出現頻度をスコア化したものです。一般に、配列データの増大につれてその信頼度も上昇し、予測結果の質、量も共に向上するものと考えられます。ここ数年、より効果的な配列情報活用のために、予測したいタンパク質と構造既知タンパク質の両方についてのプロファイルを作成し、比較を行うプロファイル比較と呼ばれる手法が盛んになってきました。我々は、立体構造類似性の認識感度上昇による予測可能範囲の更なる拡大とアラインメント精度の一層の改良を目的とし、新たなプロファイル間類似性尺度を導入することで、より強力なプロファイル比較法FORTE1 (FORTEは、FOld Recognition TEchniqueの略)を開発しました。このシステムでは、予測配列、構造既知タンパク質の配列の両方について、PSI-BLASTを利用して作成される配列プロファイルが用いられます。現在6千を超える既知構造ライブラリの配列プロファイルは、CBRCの大規模PCクラスタを利用して迅速に準備されています。

2002年、我々はこのシステムを用いて、世界規模のタンパク質立体構造予測コンテストCASP (Critical Assessment of techniques for protein Structure Prediction)に参加しました。そこでは、FORTE1が新規構造部門での構造類似性尺度による機械的な評価で165予測チーム中30位内に入り、構造認識部門で構造ドメイン分割した評価において配列情報のみを利用したサーバでは最高得点を得るなど、本手法の有効性を示すことができました。現在FORTE1は、計算機による自動的立体構造予測コンテストへ引き続き参加しており、また、一般ユーザ向けウェブサービスFORTEも開始します。^{*1} 今後は、これまで参加したコンテストで得た反省点を生かし、更に優れた予測法の構築を行おうと考えています。

*1 : 予定URL <http://www.cbrc.jp/forte/>

参考文献 : K. Tomii and Y. Akiyama: Bioinformatics (accepted)



図1 FORTE1のwebインターフェイス

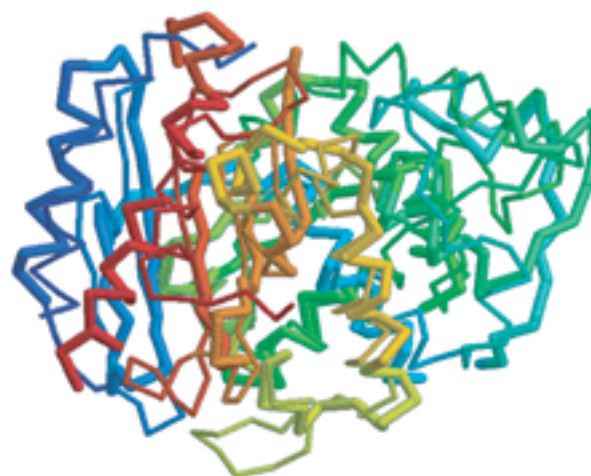


図2 FORTE1によるCASPでの予測の一例(T0168)
(PDB code:1MKI)
太線は予測構造、細線は実際の構造。