

タンパク質セグメントの構造特性と フォールドの多様性への考察

池田 和由

分子設計チーム 産総研特別研究員



タンパク質の構造（フォールド）は、約一千種類あると言われていて、実に多種多様なものが存在しています。しかし、それらフォールドを構成する部分構造（セグメント構造）は、全く異なるフォールド間でも類似した構造を共有していることが知られています（図1）。つまり、多様で複雑なタンパク質のフォールドは、かなり限られた数のセグメント構造の組み合わせで出来ているのです。最近の研究^{*1}によって、新規のフォールドでさえもセグメントレベル（10残基）で比較すると、そのほとんどが既知構造に相当することが示唆されました。これらの知見は、タンパク質の構造予測にとってセグメント構造情報を利用することの重要性を示しています。

このような背景のもと、我々はタンパク質セグメントの構造特性を明らかにする目的で研究を行ってきました。通常、タンパク質に代表される生体高分子の立体構造は、数残基のセグメントでさえも複雑で、そのままでは多次元空間で表現する必要があります。我々は、多変量解析法の1つであるPCA法を用いることで、あらゆる種類のフォールドから切り出されたセグメントを低次元空間に射影し、その分布を詳細に解析することを可能にしました。今までに、20残基以下のセグメント構造空間を構築し、その構造軸の意味を解釈することで、短いセグメントの主要な構造変量を同定しました。また、本手法を構造モチーフの分類へ応用することで、タンパク質相互作用に関与が示唆される新規モチーフも同定しました。^{*2}

現在、より長いセグメント（～50残基長）に対して本手法を応用し、今まであまり研究されなかった長いセグメントの構造空間を解析しています。解析の結果、構造分布の統計的性質はセグメントの長さに対して大きく変化することが分かりました。また、二次構造の含量で定義される構造クラス（All- α 、All- β 、 α/β 、 $\alpha+\beta$ ）の構造空間の性質を比較することで、多様なタンパク質において保存されているセグメントの構造変量を同定することができました。さらに、セグメント構造空間の構造軸には、各構造クラスに固有なものが存在していることが分かりました。

本研究で得られた結果は、構造予測においては構造サンプリング法などで出力された多様な構造の分類や天然構造に近いモデルのフィルタリングへの利用が期待できます。また、タンパク質の構造多様性の原理を理解する上でのヒントが隠されていると期待しており、現在さらに研究を進めています。

Reference

- *1 Du, P., Andrec, M. and Levy, RM.: "Have we seen all structures corresponding to short protein fragments in the Protein Data Bank? An update.", *Protein Eng.*, **16**, pp.407-414 (2003).
- *2 Ikeda, K., Tomii, K., Yokomizo, T., Mitomo, D., Maruyama, K., Suzuki, S. and Higo, J.: "Visualization of conformational distribution of short to medium size segments in globular proteins and identification of local structural motifs", *Protein Sci.*, **14**, pp.1253-1265 (2005).

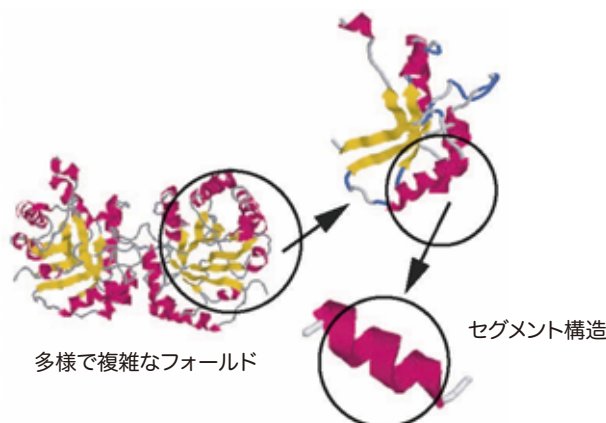


図1 タンパク質セグメントの構造

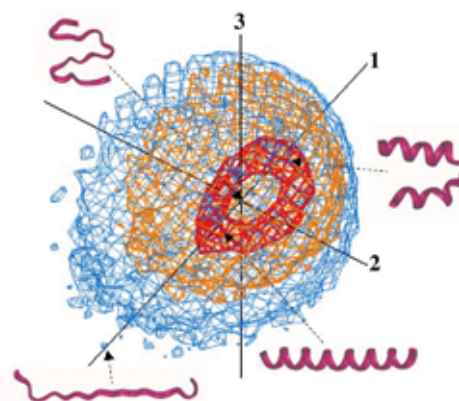


図2 可視化されたセグメントの構造空間