

蛋白質ポケットとアミノ酸間の距離分布分析

三好 真紀子 (指導教員: 伊藤 貴之)

1. 概要

人体の主成分である蛋白質は、アミノ酸が多数連結して複雑に折りたたまれることで、特有な立体構造をとる。葉は蛋白質表面上の窪み(ポケットと呼ぶ)に入り込みやすいことが知られている。そこで我々は、葉との親和性とポケットがアミノ酸とどのような位置関係にあるかがどのように関係しているか分析している。その一環として本研究では、ポケットとアミノ酸の距離分布分析のための可視化ツールを提案する。

2. 関連研究

2.1 蛋白質ポケット・結合部位 druggability 分析

Perot らは、druggable なポケットおよび結合部位を抽出し、その類似度比較評価結果や公開データベースをサーベイした論文[1]を発表している。この論文で紹介されている手法やシステムは、幾何学的な手法からエネルギー基準に基づく手法まで多種多様である。一例として、水素結合ドナーなどの物理化学の特性を使用した手法である SiteEngine[2]が発表されている。

2.2 低分子結合部位のアミノ酸の傾向

Yamaguchi らは、蛋白質にすでに低分子が結合している位置に対して、どのアミノ酸とよく相互作用しているのかを統計をとり分析している[3]。この研究では、Het-PDB[4]というデータベースを実装し、蛋白質に対してアミノ酸に関する統計をとることで、知られていなかった低分子とアミノ酸の相互作用を発見している。図1は、ATP(左)また GTP(右)と相互作用するアミノ酸を円グラフにより表している。この手法では蛋白質ポケットに分析対象を特定していない。そこで、我々は分析対象を蛋白質ポケットに特定して、アミノ酸との相互作用をポケットとアミノ酸間の距離分布を分析することで調べることとした。

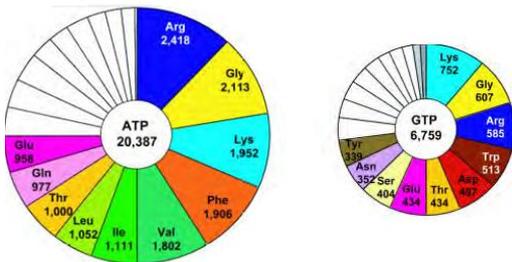


図1: ATP (左), GTP (右) と相互作用するアミノ酸残基 (文献[3]から転載)

3. 提案手法

3.1 蛋白質表面形状データの取得

本手法では、蛋白質表面形状が三角メッシュで近似されることを前提としている。eF-site[5]は蛋白質の分子表面に関するデータベースであり、コノリー曲面抽出技術を用いて分子表面形状を利用する。eF-site では、PDB (Protein DataBank) [6]に登録されている蛋白質を対象とし、頂点、2頂点を両端とする稜線、3稜線によって囲われた三角形領域の集合で構成される三角形メッシュを、XML形式で記述して提供している。

3.2 ポケット抽出

本手法では、まずポケットを蛋白質表面から抽出する。ポケット抽出はNakamura[7]の手法を用いる。この手法は

表面形状データから以下の手順を踏んでポケットを抽出している。図2は蛋白質1ezqに適用した例である。

1. 三角形メッシュ単純化により、表面の細かい凸凹を除去して平滑化された粗い形状を形成する。
2. 平滑化された表面形状から、凹部形状を抽出することで、ペプチドサイズの凹部を抽出する。
3. 抽出された凹部を、単純化前の三角形メッシュを投影する。
4. 投影された凹部形状の不要な部位を除去する。

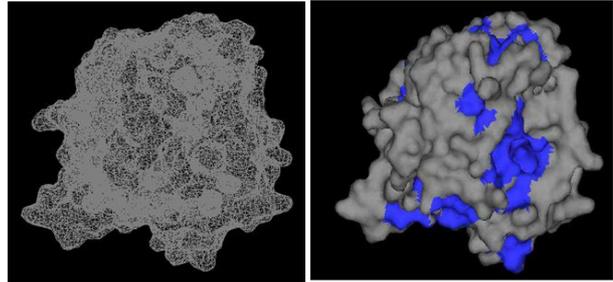


図2: 蛋白質表面形状 (左) とポケット (右) の例
この手法ではメッシュ単純化により、非常に小さい計算量で的確にポケットを抽出する。

3.3 ポケットとアミノ酸間の距離算出

続いて、蛋白質を構成する各アミノ酸とポケットの距離を算出する。現時点での我々の実装では、最初の試みとしてポケットとアミノ酸間の距離を「ポケットの最深部から、アミノ酸を構成する原子のうち最も近い原子の中心まで」と定義している。ただし、この定義はあくまでも一例であり、他の定義にしたがった距離算出も今後の課題として取り組みたい。

現時点での我々の実装ではまず、抽出した各ポケットに対して最深部を算出する。最初に、外周を近似する平面の方程式 $pv+k=0$ を定義する。ここで p は平面上の点であり v は平面の法線ベクトルである。次に、式(1)を適用してポケットを構成している各頂点 p について平面との距離 d を算出する。そして距離が最も大きい頂点を最深部として算出する。

$$d = \frac{pv+k}{|v|} \quad (1)$$

続いて、最深部である頂点とアミノ酸を構成している各原子間のユークリッド距離を算出する。我々の実装では、Protein DataBank[6]にて取得したPDBファイルを開き、蛋白質を構成する各原子とその座標値、および各原子がどのアミノ酸に属するかといった情報を参照する。

3.4 視覚的分析

我々は以上の処理を、金子ら[8]が提案した可視化ツール上で実装して、ポケットとアミノ酸間の距離に関する傾向を視覚的に分析している。現時点での我々が利用している可視化ツールでは、druggable なポケットの基準を発見するために、各ポケットが持つ特徴量と druggability の関係性を散布図として可視化している。そこで我々は現在、個々のポケットと2種類のアミノ酸との距離を2変数として散布図表示を試みている。

図3は、可視化ツールの表示例である。表示画面右側は散布図の軸選択領域となっている。選択された二項目についての散布図を表示画面の左側に表示される。ここで散布図の各プロットはポケットを表しており、赤いプロ

ットがdruggableなポケット, 青いプロットがundruggableなポケットである。散布図の中の各プロットをクリックすると画面中央にポケットの形状が表示されるようになっていく。

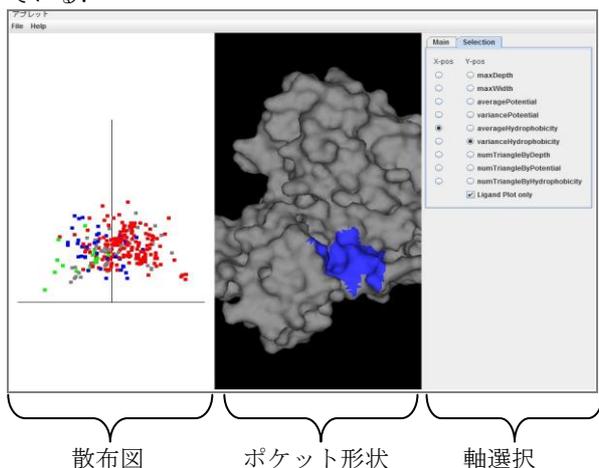


図3: 可視化ツール

現時点での我々の実装では, 20種類のアミノ酸から2種類を選択し, その2種類のアミノ酸とポケット間の距離を可視化できるようにしたが, この実装では3種類以上のアミノ酸との距離を同時に視覚的分析するのは困難であるので, その他の可視化手法の適用を今後の課題にしたい。

4. 実行結果と考察

前章で述べた可視化ツールを用いて, ポケットとアミノ酸間の距離と druggability の関係性を観察した結果を示す。本章では Halgren らが文献[9]にてテストデータとして使用した 60 個の蛋白質を対象とし, Nakamura らの手法[7]で抽出した 274 個のポケットについて可視化した。また本章での実行結果に際して, ポケットから距離の近いアミノ酸に着目するために, 2 種類のアミノ酸の両方と距離が遠い部分 (5 Å以上) に陰を作ることで, アミノ酸との距離が近いものだけを注目できるようにした。

図4は, X軸にポケットとアラニン間の距離, Y軸にポケットとチロシン間の距離を割り当てた結果である。この結果からこの2つのアミノ酸との距離が近くても druggable なポケットか undruggable なポケットどちらであるか推測をたてることは出来ない。

図5はX軸にポケットとヒスチジン間の距離, Y軸にポケットとリシン間の距離を割り当てた結果である。この結果からこの2つのアミノ酸に近いポケットは druggable なポケットが多いことが分かった。

5. まとめと今後の課題

本研究では蛋白質ポケットの druggability 分析の一環として, ポケットとアミノ酸間の距離分布を金子[8]の可視化ツールを使用して表示した結果を示した。その結果として, druggable なポケットには特定のアミノ酸との距離が近い傾向があることが示唆された。

今後の課題として, 以下について検討したい。

- ポケットの最深部以外の部位とのアミノ酸の距離分布を視覚的分析する。
- 3種類以上のアミノ酸との距離分布を同時に分析するために, 散布図以外の可視化手法を導入する。
- アミノ酸との距離分布と, その他の化学的特性(電位や疎水性など)を同時に視覚的分析する。
- 可視化ツールだけでなく判別分析などの統計的手法を導入する。
- Halgren[10]が示した 60 種類の蛋白質だけでなく, もっと多数の蛋白質を対象として距離分布の可視化を試みる。

参考文献

- [1] S. Perot, O. Sperandio, M. A. Miteva, A. -C. Camproux, B. O. Villoutreix, Druggable Pockets and Binding Site Centric Chemical Space: A Paradigm Shift in Drug Discovery, Drug Discovery Today, 15(15-16), 656-667, 2010.
- [2] "SiteEngine", <http://bioinfo3d.cs.tau.ac.il/SiteEngine/>
- [3] A. Yamaguchi, K. Iida, N. Matsui, S. Tomoda, K. Yura, M. Go, Het-PDB Navi. :A Database for Protein-Small Molecule Interactions, J. Biochem, 135, pp. 79-84, 2004.
- [4] "Het-PDBNavi.", <http://hetpdbnavi.nagahama-i-bio.ac.jp/index.php>
- [5] K. Kinoshita, eF-site, <http://ef-site.hgc.jp/eF-site/index.jsp>
- [6] "ProteinDataBank", <http://www.rcsb.org/pdb/home/home.do/>
- [7] Y. Nakamura, T Itoh, Fast Detection and Visualization of Protein Surface Pockets for Drug Design, The 11th Asian Symposium on Visualization, 2011.
- [8] 金子彩香, 中村友香理, 伊藤貴之, "蛋白質ポケットと druggability 分析のための可視化ツール", NICOGRAPH 2012, 1-8.
- [9] T. A. Halgren, Identifying and Characterizing Binding Sites and Assessing Druggability, Journal of Chemical Information and Modeling, 49(2), 377-389, 2009.

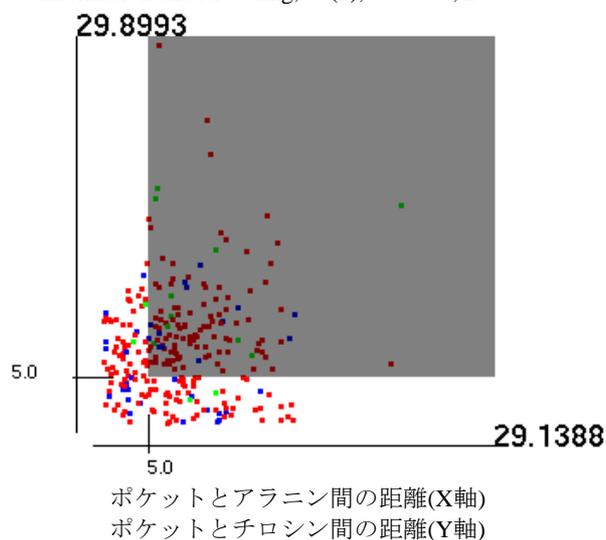


図4: 散布図の観察結果1

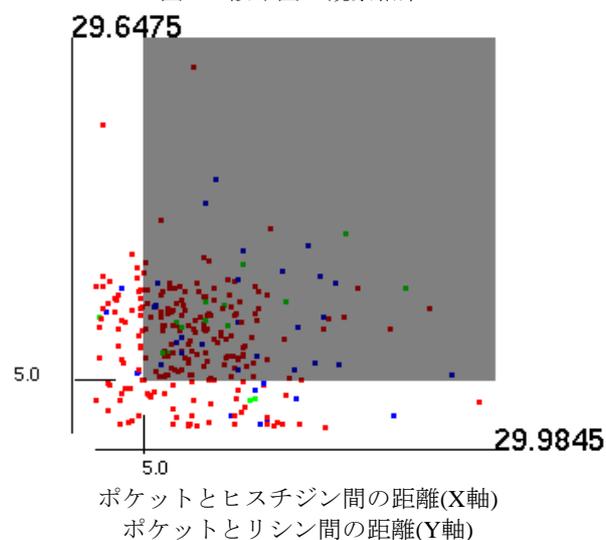


図5: 散布図の観察結果2