

蛋白質ポケット=アミノ酸間距離と druggability の相関性の可視化

三好真紀子^{†1} 伊藤貴之^{†1} 由良敬^{†1}

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、固有の立体構造を形成する。薬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、薬との親和性と、アミノ酸のポケット内で位置関係がどのように関係しているか分析している。その一環として本報告では、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。本手法では、蛋白質表面形状から抽出された各ポケットに対して最深部を求め、アミノ酸を構成する各原子との距離を算出し、その分布を可視化する。本報告では可視化結果から得られた傾向を示し、その生物学的解釈の可能性について推察する。

Visualization for Correlation between Druggability and Distance of Amino Acids in Protein Pockets

MAKIKO MIYOSHI^{†1} TAKAYUKI ITOH^{†1}
KEI YURA^{†1}

Protein is the major component of the organism. It has a unique three-dimensional structure determined by its amino acid sequence. A hole (pocket) on the surface of a protein is known to be the best target for a drug to react. We have started analyzing how "druggability" of proteins relates to the locations of amino acids in the pocket. For starter of the study, this paper presents a visualization tool for distance distribution analysis between two types of amino acids in the protein pocket. Providing protein surfaces by triangular meshes, this tool first identifies a pocket from the protein surface, specifies the deepest point in the pocket, and calculates distances between atoms of an amino acids and the deepest point of the pocket. The tool then visualizes the distribution of the distances by scatterplots. This paper presents a biological interpretation of the visualization results.

1. はじめに

蛋白質は生体を構成する主成分であり、20種類のアミノ酸が多数連結し複雑に折りたたまれることで球状に近い特有な立体構造をもつ。蛋白質の表面には凸凹の形状があり、その中でも深くて大きな凹みはポケットと定義される。ポケットに薬成分が一定時間入り込むことで、薬が蛋白質に作用することが多いとされている。ここで薬との親和性の高さを druggability といい、親和性の高いポケットを druggable なポケット、そうではないポケットを undruggable なポケットという。druggable なポケットの発見は、創薬過程において開発費用の低減につながることから、これまでも数多くの研究が報告されてきた。

我々も先行研究として、表面形状の凹部を高速抽出する手法を提案した 1)。次に、凹部の形状特徴量の比較、あるいは化学的情報である疎水性、電位などを含める多変量を解析するための可視化ツールを発表した 2)。

一方で、薬を含む低分子とアミノ酸との結合しやすさには統計的傾向があることも知られている 3)。蛋白質に対してアミノ酸種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。

これらのことから、ポケットとアミノ酸との距離も druggability を診断するための有効な情報となり得ることが示唆される。そこで本研究では、ポケット内のアミノ酸間距離の分布を視覚的に分析することを試みた。これにより、アミノ酸と druggability の相関性を明らかにすることを目標にする。

2. 処理手順

本章では、本手法の処理手順をのべる。

2.1 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-site4)は蛋白質の分子表面に関するデータベースであり、本手法は eF-site4)に登録されている分子表面形状を利用する。

2.2 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。ポケット抽出手法は Nakamura1)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。

1. 三角形メッシュ単純化により、平滑化された粗い形状を形成する。
2. 凹部形状を抽出することで、ペプチドサイズの

^{†1} お茶の水女子大学
Ochanomizu University

凹部を抽出する。

- 抽出された凹部を，単純化前の三角形メッシュに投影する。
- 投影された凹部形状の不要な部分を除去する。

2.3 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，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まで」と定義している。

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，ポケットの外周を近似する平面の方程式 $pv + k = 0$ を定義する。次に，式(1)を適用してポケットを構成している各頂点 p_0 について平面との距離 d を算出する。そしてこの距離が最も大きい頂点を最深部とする。

$$d = \frac{p_0v + k}{|v|} \quad (1)$$

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とのユークリッド距離を算出する。ここで我々の実装では，Protein DataBank 5)にて取得したPDBファイルを開き，各原子の座標値，およびどのアミノ酸に属するか，などの情報をここから参照する。

2.4 視覚的分析

特定のアミノ酸の組み合わせの近傍にどんなポケットが多いかを調べた。我々は2.3節までの処理を，金子ら2)が提案した可視化ツール上で実装し，視覚的に分析している。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。ここで散布図の各プロットはポケットを表している。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。

3. 実行結果

本章では Halgren が文献 6)にて druggable, difficult, undruggable のいずれかの結果を付与した 60 個の蛋白質を対象とし，Nakamura らの手法 1)で抽出した 274 個のポケットについて距離分布を可視化した結果を示す。

3.1 分析

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 の相関性を観察した結果を示す。

図 1 は X 軸にポケット最深部とアスパラギンとの距離，Y 軸にチロシンとの距離を割り当てた結果である。この結果をみると，二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わかった。

3.2 分析まとめ

前節の分析の全体を通じた知見を得るために，図 2 を作成した。格子状に色分けされた領域には，縦方向，横方向ともに同じ順番で 20 種類のアミノ酸種が割り当てられており，各格子領域が 2 種類のアミノ酸種の組み合わせを示している。図 2 において各格子領域の色は，二つのアミノ

酸との距離が 5 Å 未満の時の druggable の個数を全ポケット数で割った値によって変化する。この値を以後 ds とする。 ds が 0.85 未満の時に青色，0.85 以上の時に赤色で色付けする。また $|ds - 0.85|$ と全ポケット数それぞれの値が大きい程，色の彩度が大きくなる。

図 2 から，druggability 予測ができる結果を示すものは特定のアミノ酸が関わっていることがわかり，これらの一部のアミノ酸が druggable に影響しているのではないかと考えられる。

4. おわりに

本報告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ら 2)の可視化ツールで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。

参考文献

- Y. Nakamura, T. Itoh, Fast Detection and Visualization of Protein Surface Pockets For Drug Design, The 11th Asian Symposium on Visualization 2011.
- 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.
- 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.
- K. Kinoshita, eF-site, <http://ef-site.hgc.jp/ef-site/index.jsp>
- "Protein DataBank", <http://www.rcsb.org/pdb/home/home.do/>
- T. A. Halgren, Identifying and Characterizing Binding Sites and Assessing Druggability, Journal of Chemical Information and Modeling, 49(2), 227-289, 2009.

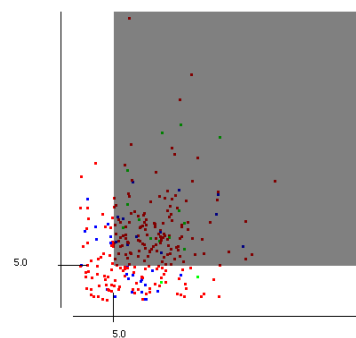


図 1 散布図での可視化結果

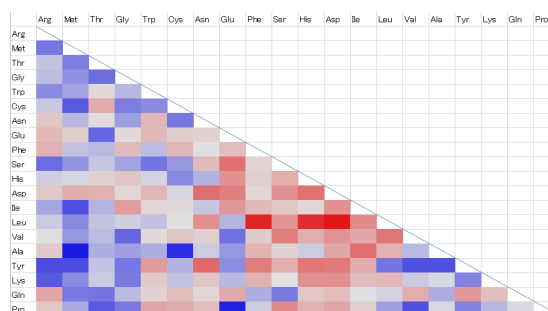


図 2 分析まとめ結果