

蛋白質表面形状におけるポケット抽出の一手法

理学専攻 情報科学コース 中村 友香理

1 概要

製薬技術の発達により現在、多種多様の薬が大量に開発されている。薬における副作用などの問題を解決するには、現段階よりさらにピンポイントに作用するような薬の開発が重要だと考えられる。そこで、薬の多くは人間の主成分である蛋白質に作用することから、本研究では蛋白質の分析に着目した。近年の研究では、薬成分は蛋白質表面上のある程度深い窪みにおいて蛋白質と反応しやすい、と考えられている。このような、蛋白質表面上にて薬成分が反応しやすい、ある程度大きく深い凹みを、近年の研究ではポケットと呼ぶ。本論文ではそういったポケットを自動的かつ高速に抽出する手法を提案する。

2 提案内容

2.1 蛋白質の表面形状

本手法では、分子構造データベース (eF-site[1]) に記録されている蛋白質の分子表面情報を用いる。eF-site では、PDB(Protein DataBank) に登録されている蛋白質のうち数万種類について、コロニー曲面抽出技術 [2] を適用して表面形状を抽出し、これをデータベース化している。eF-site では、頂点、2 頂点を両端とする稜線、3 稜線によって囲まれた三角形領域、の集合で構成される三角メッシュを XML 形式で記述し、これをオープンに提供している。eF-site からダウンロードして表示した蛋白質表面形状の例を、図 1 に示す。

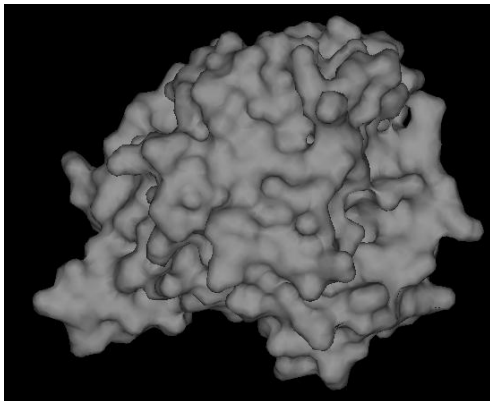


図 1: 蛋白質表面形状の例

2.2 メッシュ単純化

ポケット探索において、浅く小さな凹凸を平滑化し、ある程度大きな凹部だけを確実に抽出する手段として、本手法ではメッシュ単純化を適用する。メッシュ単純化には様々な方法があるが、本論文の実装では、マーチングキューブ法 [3] を適用する。

2.2.1 スカラ値を算出

まず蛋白質表面形状を包含するような格子を生成し、各格子点にスカラ値を与える。各格子点について、蛋白質表面形状を形成する頂点のうち一番距離の小さいものを算出する。その頂点と格子点との距離をスカラ値として付与する。その際、蛋白質表面形状の内側にあるものは負の値とし、外側にあるものは正の値とする。

2.2.2 等値面の生成

与えたスカラ値をもとに、マーチングキューブ法を適用して等値面を生成する。等値面とは、3次元空間中でスカラ値が一定である点の集合として構成される曲面を、三角メッシュなどで近似生成する手法である。以下に等値面生成の簡単な仕組みを示す。格子線を結ぶ隣り合った2格子点のスカラ値を a および b とし、生成すべき等値面のスカラ値を C とする。また、格子線上でスカラ値は線形に変化すると仮定する。このとき3値が $a < C < b$ を満たすとすると、格子線を $|C - a| : |C - b|$ の比に内分する点においてスカラ値は C である。1格子を構成する各格子線において、この条件を満たす点を抽出し、これらを連結することで、格子内部に多角形を生成できる。以上の処理を各格子に適用し、多角形の集合を生成することで、等値面を近似生成できる。以上の方法で蛋白質表面形状である三角メッシュを、 $C = 0$ とし単純化した例を図 2 に示す。

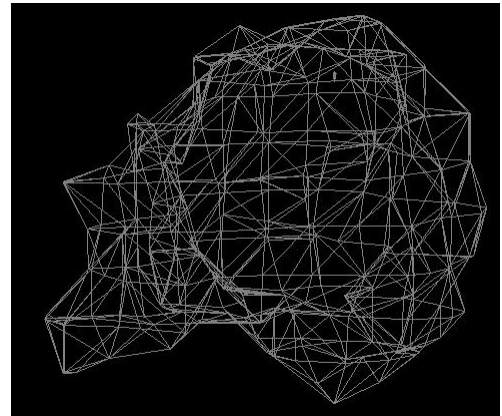


図 2: メッシュ単純化結果の例

2.3 凹部の抽出

続いて、前節に示した手法で生成された単純化メッシュに対して、凹部の抽出を適用する。

2.3.1 頂点属性の決定

三角メッシュを構成する各頂点 v_i について、位置を (x_i, y_i, z_i) 、法線ベクトルを $(n_{x_i}, n_{y_i}, n_{z_i})$ とする。このとき、式 (1) において、 v_i と稜線を共有する全ての隣接頂点 (x, y, z) に対して t が負ならば「凹」という属性を v_i に与える。

$$t = n_{x_i}(x - x_i) + n_{y_i}(y - y_i) + n_{z_i}(z - z_i) \quad (1)$$

2.3.2 三角形属性の決定

2.3.1 節の結果を用いて、「凹」の属性を持つ頂点を 1 個以上含む三角形に「凹」の属性を与える。ここで「凹」の属性を与えられた局所形状をポケット候補とする。図 3 に「凹」属性を与えられた三角形を青く塗った結果を示す。

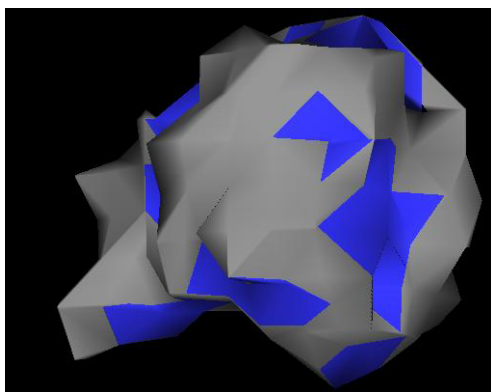


図 3: 凹部の抽出結果の例

2.4 凹部の投影

続いて、2.3 節でポケット候補と判定した局所形状が、単純化前のメッシュにおいてどの部分に対応するかを特定する。単純化後の三角形メッシュの各三角形の重心に対し、蛋白質全体の中心点からのベクトルを算出する。単純化前の三角形メッシュの各三角形の重心も同様にベクトルを算出し、単純化後の三角形メッシュのベクトルとの内積が 0 に最も近いものと同じ属性を投影する。その結果を図 4 に示す。これらがポケットとなる。

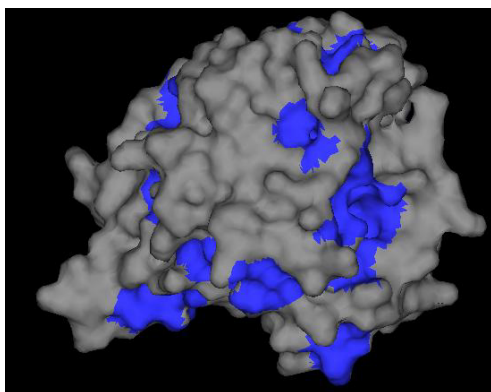


図 4: 凹部の投影結果の例

3 検証

本研究では Halgren[4] の論文にて薬との親和性に関する実験結果が公開されている 60 種類の蛋白質を題材にして、ポケットの各特徴について検証を実施した。

3.1 深さ

まず、ポケットの深さと薬との親和性に関する検証を行った。ポケットの外周を近似する平面をとり、その面からポケットを構成する頂点への距離を算出した。そして、その最大値をそのポケットの深さとした。その結果、薬との親和性が高いとされている蛋白質におけるポケットの深さと、薬との親和性が低いとされている蛋白質のポケットの深さでは大した違いが見られなかった。これは、薬成分が入る入口となる部分の幅が関係しているものと考えられる。今後、ポケットの外周の形から入口部分の幅を特定し、さらなる検証を進める。

3.2 電位と疎水性

続いて、電位・疎水性と薬との親和性に関する検証を行った。eF-site のデータには、表面形状の情報に加え、各頂点における電位と疎水性が記されている。それらを三角形メッシュへ拡張し、ポケットとの関連性を見た。薬が効きにくいとされる蛋白質において、ポケット部位の電位は負であり、疎水性は高いという結果になった。疎水性に関しては、親水性の高いものほど反応をおこしやすいとされているためだと考えられるが、電位についてはまだまだ検討が必要であると考えられる。

4 まとめ

本論文では、蛋白質分子表面形状からポケット形状を抽出する一手法を提案した。本手法では、三角メッシュとして近似された蛋白質表面形状に、まずメッシュ単純化を適用して細かい凹凸形状を平滑化し、その上で大きな凹形状を抽出した。そして、凹形状の位置を単純化前のメッシュに投影することで、ポケット候補として抽出した。さらに、ポケットの特徴と薬との親和性の関連性の検証をいくつか行った。

参考文献

- [1] eF-site
<http://ef-site.hgc.jp/eF-site/index.jsp>
- [2] Molecular Surface Package
<http://connolly.best.vwh.net/>
- [3] W. E. Lorensen. H. E. Cline. Marching Cubes: A High Resolution 3D Surface Construction Algorithm. *ACM SIGGRAPH*, Vol. 21, No. 4, pp. 163-169, 1987.
- [4] T. A. Halgren, Identifying and Characterizing Binding Sites and Assessing Druggability. *Journal of Chemical Information Modeling*, Vol. 49, pp. 377-389, 2009.