

# 蛋白質表面形状分析手法 PROTEIN を用いたポケット分析

中村 友香理 (指導教員: 伊藤 貴之)

## 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。薬の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。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、我々は蛋白質の分析に着目した。薬は蛋白質に作用する際、蛋白質の窪んでいる部分に入り込んで反応することが多い。一方で、蛋白質と薬成分は表面に接した瞬間から化学反応を起こすわけではなく、一定時間にわたって薬成分が蛋白質周辺の安定的な位置にとどまって反応することが多い。このことから薬成分は、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、ということが考えられる。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本報告ではポケットと呼ぶ。

本研究では、我々が研究を進めている蛋白質表面形状分析プロジェクト PROTEIN[1]の一環として、自動的かつ高速にポケットを抽出する手法を提案する。蛋白質表面形状が三角メッシュとして与えられていることを前提として、本手法ではまずメッシュ単純化を適用し、微小な凹凸を平滑化した概略的な3次元形状を得る。続いて本手法では、単純化された三角メッシュ上で凹部を構成する三角形群を抽出し、これを単純化前の三角メッシュに投影することで、単純化前の三角メッシュにおける凹部を抽出する。こうして抽出された凹部は、ある程度大きく深いものであるといえるため、ポケット候補となる。

## 2. 提案手法

### 2.1 蛋白質の表面形状

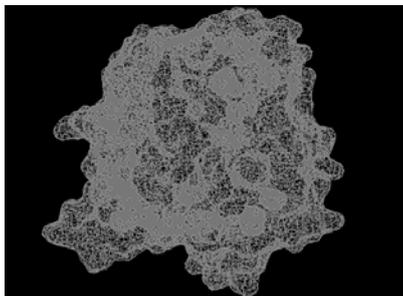


図1 蛋白質表面形状

本手法では、分子構造データベース(eF-site[2])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) に現段階で発見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[3]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。

### 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、我々の実装では、格子構造を適用したスカラ場生成結果に基づくメッシュ再生成手法[4]を適用する。

#### 2.2.1 三角メッシュ頂点にスカラ場を生成

まず表面形状を形成している三角メッシュの各頂点に、式(1)で示されるスカラ場を与える。

$$S = \begin{cases} \frac{r - \text{dist}(C, P)}{r} & (r > \text{dist}(C, P)) \\ 0 & (r < \text{dist}(C, P)) \end{cases} \quad (1)$$

ここで、 $r$  はスカラ場の影響半径、 $C$  は表面を形成する頂点、 $P$  はスカラ場を適用する空間中の任意の点、 $\text{dist}(C, P)$  は  $C$  と  $P$  の距離を表す。

#### 2.2.2 格子点のスカラ値を算出

続いて、図形を包括するような格子を生成し、各格子点にスカラ値を与える。具体的には各格子点について、式(1)の  $P$  に格子点の座標値を代入し、 $C$  に各頂点の座標値を代入し、各々の頂点について算出した  $S$  の最大値を、その格子点のスカラ値とする。結果として、スカラ場の影響半径内にある格子点には正値が割り当てられ、それ以外の格子点においてスカラ値は0となる。

#### 2.2.3 等値面の生成

与えたスカラ値  $S$  を元に、マーチングキューブ[5]という手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に仕組みを示す。稜線で連結された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、稜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、稜線を  $|C-a|:|C-b|$  の比に内分する点においてスカラ値は  $C$  である。1格子を構成する各稜線において、この条件を満たす点を抽出し、これを格子の面上で連結することで、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。

### 2.3 ポケット判定

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、ポケット判定を適用する。

### 2.3.1 頂点属性の決定

三角形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  $(n_{xi}, n_{yi}, n_{zi})$  とする。このとき、式(2)

$$t = n_{xi}(x - x_i) + n_{yi}(y - y_i) + n_{zi}(z - z_i) \quad (2)$$

において、 $v_i$  と稜線を共有するすべての隣接頂点  $(x, y, z)$  に対して  $t$  が正ならば「凸」、負ならば「凹」の属性を  $v_i$  に与える。正である隣接頂点、負である隣接頂点のどちらも含む場合は「該当なし」という属性を与える。図2に示したように、頂点の接平面よりも隣接頂点が内側にあれば「凸」(左)、隣接頂点が外側にあれば「凹」(下)、それ以外は「該当なし」(右)となる。

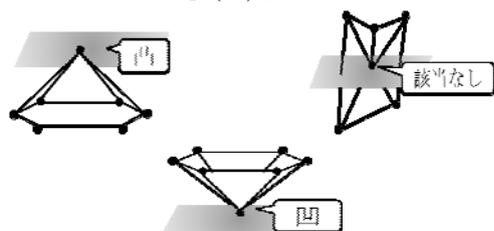


図2 頂点属性

(左)凸頂点 (右)該当なし (下)凹頂点

### 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にメッシュ単純化後、「凹」属性を与えられた三角形を青く塗った結果を示す。

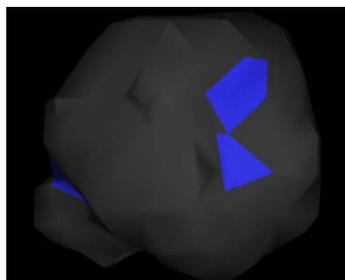


図3 メッシュ単純化後ポケット判定結果の例

## 2.4 ポケット照合

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ために、ポケット照合を適用する。

単純化前の三角形メッシュの各三角形を  $b_i$  とし、単純化後の三角形メッシュの各三角形を  $a_i$  とする。我々の実装では単純に、各々の  $b_i$  の重心に対して、もっとも重心位置に近い  $a_i$  を探索する。そして  $a_i$  の凹凸属性を読み込み、それを  $b_i$  の属性に上書きする。

## 2.5 検証実験

PDB に記録されている蛋白質情報には、結晶化の過程で配合された原子の情報を含むものがある。それらの多くは、蛋白質表面上の化学反応を引き起こしやすい位置に記録されている。そのような位置におけるポケットの

発見は重要である。ポケット照合の結果に、黄色い球で反応しやすい原子の情報を重ね合わせた結果を、図5に示す。これを一つの指標として今後、本手法でポケット候補と判定された部分の検証を進める。

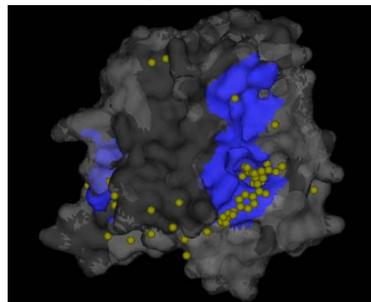


図4 検証結果

## 3. まとめ・今後の課題

本報告では、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照合することで、ポケット候補として抽出した。

本研究はまだ研究の初期段階にあり、薬品開発への貢献を目指すには以下のような課題がある。

[ポケットと薬成分(リガンド)の衝突判定]

どのような形状のポケットに、どのような形状のリガンドが収まりやすいのか、といった形状的観点での判定手法について実装を進めたい。

[ポケット形状の分類]

数万種類もの蛋白質から膨大な量のポケットを抽出した際に、特定のリガンドと形状的に適合する蛋白質を検索することを考えると、全ての蛋白質の全てのポケットに対して衝突判定を実施するのは現実的ではない。ポケット形状の分類によって、その効率化を図りたい。

[ポケット形状と化学的特性の相関性]

化学的性質についても、ポケット分類結果をもとにして分析を試みたい。

## 参考文献

- [1] K. Nishiyama, T. Itoh, PROTEIN: A Visual Interface for Classification of Partial Reliefs of Protein Molecular Surfaces, The Institute of Image Electronics Engineering of Japan, Vol. 37, No. 3, pp. 181-188, 2008.
- [2] eF-site, <http://ef-site.hgc.jp/eF-site/index.jsp>.
- [3] Molecular Surface Package, <http://connolly.best.vwh.net/>
- [4] 小田,村岡,千葉: 仮想粘土の粒子ベース・ビジュアルシミュレーション, 情報処理学会論文誌, Vol. 42, No. 5, pp. 1142-1149, 2001.
- [5] 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.