

タンパク質接合面の一般化とその構成法

数理情報学第五研究室 修士課程 2年

伊藤 功*

2006年2月6日

概要

2つ以上のタンパク質が空間充填モデルで表現されているときに、その接合面を、power diagram を用いて定めた。また、接合面上に、タンパク質間の相互作用における各残基の寄与度を表すヒエラルキーを構成した。その妥当性は、ホットスポットと呼ばれる部位を予測することによって確かめた。定めた接合面は従来のものの一般化となっている。

1 研究の目的

タンパク質接合面を構成する目的は、タンパク質同士の結合様式を記述する手段、及び可視化の手段を提供することである。タンパク質間の相互作用における各アミノ酸残基の寄与度をシステマティックに測定する手法として、アラニンスキニングという手法がある。これは、ある残基 R をアラニン残基に変異させたときの自由エネルギーの変化 $\Delta\Delta G_R$ を測定するものである。 $\Delta\Delta G_R$ は、

$$\Delta\Delta G_R = RT \ln \frac{(\text{変異後の平衡定数})}{(\text{変異前の平衡定数})} \quad (1)$$

と計算される。ただし、 R は気体定数であり、 T は絶対温度を表す。

Clackson と Wells [2] によれば、「数十個の残基から構成されているタンパク質であっても、結合エネルギーに大きく寄与する残基は数個である。」そのような残基をホットスポットと呼ぶ。Bogan と Thorn [1] によれば、「ホットスポットは周りの残基によって溶媒から隔てられており、O-ring の構造を持っている。」従って、O-ring の構造はタンパク質間の相互作用において重要な意味を持った幾何学的な構造であると考えられる。そこで、タンパク質接合面上に、O-ring の構造を反映したヒエラルキーを構成する。即ち、ポロノイファセットに重みを定め、重要な領域とそうでない領域を、重みの大小によって表現する。

2 定義とアルゴリズム

2.1 空間充填モデル

タンパク質中の全ての原子を球で表し、タンパク質はその和集合で表す。各原子 b_i は結合状態に応じてファンデルワールス半径 r_i が与えられる。タンパク質 j を、 $B_j = \cup\{b_1, \dots, b_n\}$ で表す。タンパク質 $1, 2, \dots, l$ の複合体 B は $B = \cup B_j$ で表す。

2.2 power diagram

球 b (中心 z , 半径 r) と点 x に対して、パワー距離 $\pi_b(x)$ を、 $\pi_b(x) = \|x - z\|^2 - r^2$ と定める。点 x が球 b の内部にあるときは、 $\pi_b(x)$ は負の値をとり、球面上にあるときは、0 となる。

球の集合 $B = \{b_1, b_2, \dots, b_n\}$ に対して、 $V(b_i, B) = \{p \in R^3 \mid \pi_{b_i}(x) \leq \pi_{b_j}(x), j \neq i\}$ で構成される領域 $V(b_i)$ を b_i のポロノイセルと呼ぶ。また、 $\{V(b_1, B), V(b_2, B), \dots, V(b_n, B)\}$ を power diagram と呼ぶ。ポロノイセルの境界は、平面となるから、ポロノイセルは、空または内部を持つ凸多面体となる。

定義 2.1 (タンパク質接合面)

生成元の球を含むタンパク質 B_j であるようなポロノイセルの色を j と定める。ポロノイ面は2つのポロノイセルにより定まるので、1色または2色で色づけされる。ポロノイ辺やポロノイ点は、2色で色づけされたポロノイ面に含まれているときのみ、複数の色で色づけされる。

このような色づけ規則に従い、タンパク質-タンパク質境界面 S を $S = S(B_1, \dots, B_l) = \cup\{\text{複数色で色づけされたポロノイ面, 辺, 点}\}$ と定める。

2.2.1 溶媒分子の通過可能性

球を溶媒分子の半径 r_w だけ膨らませておく。ポロノイ辺 e を定める球を a, b, c , その中心 $p(a), p(b), p(c)$ が定める平面を h とする。 $e \cap h \neq \emptyset$ のときは、 $w_e = \max\{-\pi_a(x), 0\}$

*e-mail: isao@simplex.t.u-toyko.ac.jp

とする。そうでないときは、 e の端点を s, t として、 $w_e = \max\{-\pi_a(s), -\pi_a(t), 0\}$ とする。

2.2.2 flow relation

regular 三角形分割の四面体の集合 T について、flow relation ' \prec ' $\subset T \times T$ を次のように定める。 $\tau \prec \sigma$ とは、以下の二つの条件を満たすことである。

- (i) τ と σ は三角形 ϕ を共有する。
- (ii) $\text{int } \tau$ と z_τ は平面 $\text{aff } \phi$ について反対側にある。

2.3 ボトルネック最短路問題

有向グラフ $G = (V, E)$ 及び出発点 $s \in V$ 、辺から実数へのラベル l が与えられたとする。1つの経路に対して、経路上の辺ラベルの最大値を、その経路のボトルネック長と呼ぶ。頂点 s から頂点 v までのすべての経路のボトルネック長の最小値をボトルネック距離と呼び、 $b_s(v)$ と記す。また、その最小値を達成する経路をボトルネック最短路と呼ぶ。ボトルネック最短路問題とは、全ての頂点 v について $b_s(v)$ を求めるという問題である。頂点 s から頂点 v までの経路の集合を P_v と書くとき、

$$b_s(v) = \min_{p \in P_v} \max_{e \in p} \{l(e)\},$$

と書くことができる。ボトルネック最短路問題は、ダイクストラ法において、演算+の代わりに演算 \max を用いて解くことができる。

2.4 アルゴリズム

入力: タンパク質複合体 B 、仮想球の半径 r_w 。

出力: 接合面 $S(B, r_w)$ 。

1. B の各原子の半径を r_w だけ増加させ、その結果を B' とする。
2. power diagram $V(B')$ を構成し、そのグラフを $G = (V, E, F)$ とする。
3. 全ての辺 $e \in E$ に対して、重み w_e を定める。
4. ボトルネック最短路問題を解いて、各頂点 v の重み w_v を定める。
5. foreach フェイス $f \in F$
6. $w_f = \min_{v_1, \dots, v_n} \{w_{v_i}\}$ 。
7. $w_f > 0$ であれば、 f に「到達不可能」のラベルをつける。

8. endfor

9. 「到達不可能」のラベルのついた f のうち、複数色で色づけされたポロノイ面の集合 $S(B, r_w)$ を出力して終了する。

ステップ2の power diagram の構成は、原子数 n に対して $O(n \log n)$ の計算量で実行できる。ステップ4の計算量は $O(|E| \log |V|)$ であり、ステップ5以降は $O(|F|)$ の計算量で実行できる。タンパク質の power diagram において、ポロノイ頂点、ポロノイ辺、ポロノイ面の数は $O(n)$ であるから、全体の計算量は $O(n \log n)$ である。

3 ヒエラルキーによるホットスポット予測

接合面 $S(B, r_w)$ に属するポロノイ面 p_0, p_1, \dots, p_k が、残基 R の側鎖によって構成されたとする。このとき、関数 $h(R)$ を、

$$h(R) = \sum_{i=0}^k (p_i \text{ の面積}) \frac{(R \text{ の面積})_{w_i}}{(S \text{ の面積})_{w_i}}, \quad (2)$$

と定義する。ただし、 w_i はポロノイ面 p_i に対して定めた重みとし、 $(R \text{ の面積})_{w_i}$ と $(S \text{ の面積})_{w_i}$ は、溶媒分子が到達不可能な接合面 $S(B, w_i)$ における面積を表すものとする。

予測の結果を先行研究と比較したものが表1である。先行研究と比較して、提案手法は単純な方法はあるが、遜色のない精度が得られた。また、埋没表面法よりも精度良く予測できたことから、生物学的に意味のある領域を表すことができたと考えられる。

表 1. ホットスポット予測の精度。

手法	データ	ホットスポット	中性残基
埋没表面法	233	64.6 %	64.6 %
Kortemme et al.	233	51.4 %	87.6 %
Ban et al.	233	72.7 %	72.5 %
Kortemme et al.	193	44.6 %	92.1 %
提案手法	193	73.8 %	73.5 %

参考文献

- [1] A. A. Bogan and K. S. Thorn: Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology*, Vol. 280, pp. 1–9, 1998.
- [2] T. Clackson and J. A. Wells: A hot spot of binding energy in a hormone-receptor interface. *Science*, Vol. 267, pp. 383–386, 1995.