

タンパク質の立体構造を評価するための 統計的粗視化4体ポテンシャルの開発

○石飛秀斗¹, 孫 林玉¹, 増田尚之¹, 相田拓洋², 西垣功一², 後藤仁志¹

¹豊橋技術科学大学 大学院工学研究科 (〒441-8580 愛知県豊橋市天伯町雲雀ヶ丘 1-1)

²埼玉大学 大学院理工学研究科 (〒338-8570 埼玉県さいたま市桜区下大久保 255)

【諸言】

タンパク質の生化学機能を利用, 制御しようとする創薬やタンパク質工学において, 観測に頼らずより高精度でより高速にその立体構造を予測できる手法の開発が強く望まれている. 既に様々な手法が開発されている中で, ノースカロライナ大 Tropsha 等は立体構造が既知のタンパク質構造データセットの統計処理によって, タンパク質内で空間的に近接する4つのアミノ酸残基の組み合わせ出現確率に基づく4体ポテンシャルを導出し, これをタンパク質立体構造の評価に用いる方法を提案した[1]. 我々はこのユニークで興味深い Tropsha 法に着目し, 4体ポテンシャルを柔軟に開発できる独自システムを開発している. 本研究では, インフルエンザウイルスのシアル酸認識タンパク質である Neuraminidase の天然型 (PDB ID: 2hu4) とタミフル耐性型 (PDB ID: 3cl0) の立体構造から4体ポテンシャルを定義し, これを用いてそれぞれの立体構造をクロス評価する. また, より高精度な評価を目指し, 水素結合を考慮した4体ポテンシャルの改良を試みる.

【方法】

本研究で開発したプログラムの概要を以下に示す.

Tropsha 法による立体構造評価システム: 各アミノ酸残基の重心を代表点として粗視化し, Delaunay 分割により4つの代表点を頂点とする四面体の集合体としてタンパク質の立体構造を記述する. そして, 四面体の頂点にある残基種の重複組み合わせ (8,855 種類) と, 辺を構成する残基間結合パターン (5 種類) に従って四面体を 44,275 種類のクラスに分類し (実際に全てのクラスが存在するわけではない), それぞれに与えられた4体ポテンシャルを積算したスコアからタンパク質立体構造を評価する. スコアが高いほど既知の立体構造で出現する頻度の高い四面体で構成されていることになり, より存在する可能性の高い立体構造であると評価できる.

4体ポテンシャル開発システム: 1つ, または複数のタンパク質立体構造の参照データセットに対して, Delaunay 四面体分割し, 分類された各種四面体の出現確率から参照データセット固有の4体ポテンシャルを導出する.

水素結合による分類クラスの細分化: 一次配列に従った残基間結合パターンに水素結合パターン加える[2]. 2から3分岐までの水素結合パターンを考慮すると, 結合パターンは11種類になる. また, 遠距離残基間に水素結合が見つかった場合には, 新たに四面体を追加する.

水素結合を評価するため, 水素原子の座標が定義されていないタンパク質構造は, CONFLEX を用いて水素を付加し, 水素位置のみの最適化を行う.

【結果】

Tropsha オリジナルの4体ポテンシャルに加え, 天然型と耐性型, およびそれら両方を参照データとして, 水素結合を考慮した場合としない場合のそれぞれで4体ポテンシャル (6種類) を導出した. それぞれのポテンシャルを用いて天然型と耐性型の立体構造を評価した結果を Table 1 に示す.

Tropsh オリジナルの4体ポテンシャルでは耐性体構造が高く評価されるが, それぞれの構造を参照データとして導いた本研究による4体ポテンシャルは, 参照構造が高い積算スコアを示した. ポスター発表で議論を深める.

[1] B. Krishnamoorthy, A. Tropsha, *Bioinformatics*, 19, 1540-1543 (2003).

[2] 増田尚之, 孫林玉, 相田拓洋, 西垣功一, 後藤仁志, SCCJ2011 春季年会, 1006 (2011).

Table 1. Total scores of some four-body potentials for nature and drug tolerant Neuraminidase structures

four-body potential	Tropsha Method			This work (including HB)		
	Natural	Tolerant	Δ	Natural	Tolerant	Δ
Tropsha	1417.7	1430.7	-13.0	-	-	-
This work						
Natural	2480.8	2136.1	344.7	3080.2	2228.2	852.0
Tolerant	2138.1	2464.1	-326.0	2237.9	3049.8	-811.9
Average	2389.4	2371.9	17.5	2878.1	2852.1	26.0