

残基との相互作用エネルギーに基づく

ドッキングシミュレーション結果の絞り込み

宮津 知美¹, 安尾 信明², 関嶋 政和³東京工業大学 情報工学科¹, 東京工業大学 大学院情報理工学研究科 計算工学専攻²東京工業大学 学術国際情報センター³

1. はじめに

近年、創薬の早期段階において、計算機による薬剤候補化合物のスクリーニングがしばしば行われている。ドッキングシミュレーションはそのための手段のひとつである。ドッキングシミュレーション用のソフトウェアとしては、Glide[1]やGOLD, FlexX, Autodock[2]などが一般的に用いられることが多い。これらのソフトウェアはヒット化合物の提案にある程度有効であるが、ほとんどの場合、これらのソフトウェアでそれほど上位にランク付けされない化合物の中にも活性のある化合物が存在する。

一方で実際に化合物が薬剤として使用されるためには、阻害活性だけではなく、安全性等の問題をクリアする必要がある。特にヒトのタンパク質に類似したものをターゲットとする場合、種特異性に留意した化合物を提案することは、より安全な薬剤の開発につながると考えられる。

本稿では、アミノ酸残基との相互作用エネルギーに基づいてドッキングシミュレーション結果を絞り込むことで、より有望な化合物を薬剤の候補として提案する手法を考える。

2. 使用するソフトウェア

本稿では、ドッキングシミュレーションのソフトウェア、Glide6.2を対象とする。Glideには各リガンドの自由度を反映した pose に対して、それぞれの残基との相互作用を出力する設定が用意されている。本稿はこの設定を用いて出力されたGlideのドッキングシミュレーション結果を、絞り込みの対象とする。

また、相互作用を避けるべき残基の自動選択のため、我々の研究室で開発されたプログラムを使用している。このプログラムでは内部的に、マルチプルアラインメントツールのClustalW、ならびに薬剤結合候補部位を予測するためのツールである Sitemap を使用している。

3. 提案手法の概略とスコア関数

提案手法では、ターゲットとなるタンパク質がヒトのタンパク質と特に強い類似度を有する場合に着目し、共通するアミノ酸残基と強く相互作用する pose を不適と判断する。

pose と残基との相互作用の強さを測る指標として、pose ごとに以下のようなペナルティスコアを定義する。

$$PenaltyScore = \frac{|A| \sum_{j \in S} E_j}{|S| \sum_{i \in A} E_i}$$

E_i i 番目の残基との相互作用エネルギー

S 指定残基の集合

A 出力結果に記載のある全残基の集合

ただし N は Glide が結果として出力した残基の総数を表すため、厳密にはターゲットとなるタンパク質のアミノ酸残基の総数とは異なる場合がある。Glide はデフォルトの設定で、グリッド中心から 12.0 Å 以内にある残基を結果に出力する。

4. テストに使用するデータ

テストにおいては、シャーガス病の原因となる原虫、*Trypanosoma cruzi* が有する酵素、spermidine synthase (TcSRM) (図1)のプトレシン結合部位を対象とする。

指定残基は human spermidine synthase と一致する残基 12 箇所(表1)。

残基番号	種類	残基番号	種類
20	PHE	178	THR
32	MET	209	GLN
33	SER	242	VAL
70	CYS	244	THR
71	ILE	245	TYR
77	ASP	250	ILE

表1: テストデータの指定残基

ナミキ商事の提供する化合物データベース (2013年6月版)のうち、HTS・BB計5,188,150化合物に対し、拘束なし SP モードでのドッキングを行った。残基の出力を除いた各設定はGlideのデフォルトの設定を利用している。

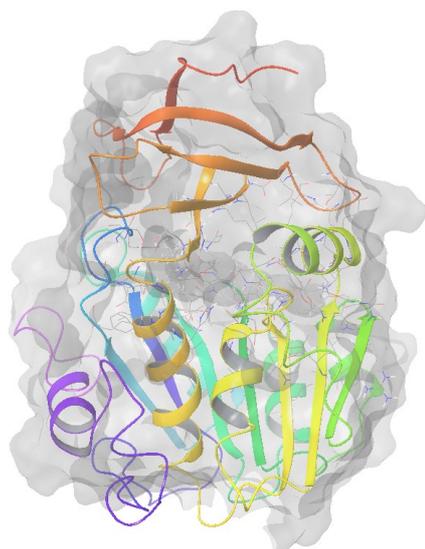


図1 : TcSRM

5. 計算環境

OS ubuntu 14.04 LTS (64bit)
メモリ 5.8GiB
CPU Intel Core i7 CPU 950 @3.07GHz

6. 計算時間

図2に指定残基数ならびに pose 数に対する実行時間のグラフを示す。

ただし、以下は各入力ファイルの読み込みからスコアの計算までの時間を示し、Glideにより出力されたファイルの解凍時間、ならびに結果の出力に要した時間は含まれていない。

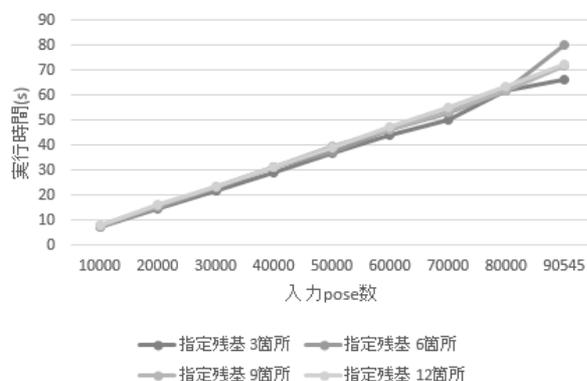


図2 : 縦軸は計算時間(秒)、横軸は pose 数

提案手法の計算時間はドッキング結果の pose 数に対し、ほぼ線形に推移する。指定残基数が増えるにつれ計算時間は微増する傾向があるが、影響はごく小さく、計算時の環境によっては指定残基数が少ないほうが実行に時間がかかる場合もある。

pose 一件あたりの計算時間は 1ms 未満であり、ドッキングシミュレーションの所要時間と比較して高速である。

7. スコア分布

図3にペナルティスコアの分布を示す。

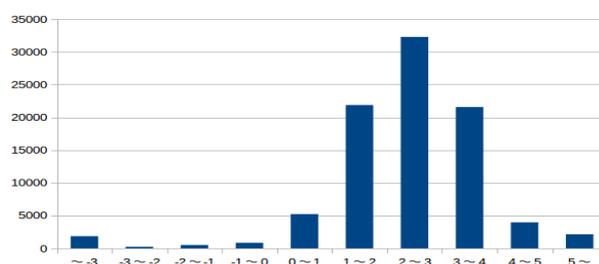


図3 : ドッキングスコア上位 100000 件から、リガンドの重複を除いた 90545 件の pose に対する適用結果

8. 結論

本研究では、特定の残基群との相互作用エネルギー値に着目し、薬剤候補化合物の種特異性を評価する指標を作成した。

この指標はドッキングシミュレーションにかかる時間と比較して、非常に高速に計算することができる。

今後 human spermidine synthase に対するドッキング結果との比較等を通じて指標の有効性を検証していく。

参考文献

- [1] Friesner, R.A.; Murphy, R.B.; Repasky, M.P.; Frye, L.L.; Greenwood, J.R.; Halgren, T.A.; Sanschagrin, P.C.; Mainz, D.T., "Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein-Ligand Complexes," *J. Med. Chem.*, **2006**, *49*, 6177-6196
- [2] Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) Autodock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Computational Chemistry* 2009, 16: 2785-91.