

in silico 親和性フィンガープリンティング

(理研 GSC、横浜市大) ○佐藤一朗*、(アクセルリス) Hongwei Huang、

(NTT データ) 黒須雅隆、(理研 GSC、横浜市大) 松尾洋

1. はじめに

近年、遺伝子の配列情報、それからコードされるタンパク質の構造情報は急速に増加している。この情報を利用することによって、標的タンパク質の三次元構造情報を基にしたハイスループットなリード探索や活性部位の構造に基づく合理的な薬剤設計が可能になってきた。その結果、分子標的薬の様に、より効率的で、効果的な医薬品を研究、開発することが出来るようになってきた。

しかし、このように特異性を上げた分子標的薬も含めて、既存薬剤の中には、生体内でどのタンパク質にどのように作用しているかが明確に解明されないままのものが多く存在するのが現状である。分子標的薬であっても、生体内に存在する膨大な種類の「非標的」タンパク質のそれぞれに対してどのような作用を及ぼすかが実際に確認されているわけでは必ずしもない。そこで、我々はプロテオーム規模に対するアプローチが薬剤の作用メカニズムの解明や創薬において重要となると考えた。このようなアプローチについて、構造プロテオミクス研究の進展によって、プロテオーム規模での親和性フィンガープリンティングを *in silico* で行うことが可能になりつつある。これは、薬剤の作用メカニズムを分子レベルで解明する上

で有用であるばかりでなく、新規薬剤開発のために有益な知見ももたらすと考えられる¹⁾。そこで、今回、タンパク質－薬剤間の分子間相互作用予測を行うことで、このような親和性フィンガープリンティングを *in silico* で行った。これらの結果をもとに、新規薬理メカニズム検出における *in silico* プロテオーム規模親和性フィンガープリンティングの可能性と意義について報告する。

2. 材料及び方法

2-1. 化合物データ

MDDR (2002 年 9 月時点) 中の 119, 108 の薬剤から、

- (1) 非水素原子数が 10 以上 35 以下
- (2) phase (開発段階) に関する注釈が phase I/II/III, launched のいずれか

という条件を満たすものを選んだ。なお、MDDR エントリー中には 2 分子以上から構成される錯体も含まれているが、(多くは H₂O, HCl 等との「錯体」)、その場合は、最も分子数の大きい構成分子のみを取り出して、その非水素原子数を計算した。その結果、2140 個の薬剤が抽出された。

* ksato@gsc.riken.go.jp

2-2. タンパク質立体構造データ

NCBI の nrpdb (2003 年 4 月 1 日版) から、p-value が $10e-80$ レベルでのクラスター (配列が類似するクラスター) の代表 7021 個の中から残基数が 100 以上かつ座標未定義原子を含む残基の割合が 5% 以下のヒトタンパク質を選んだ。その結果、561 個のタンパク質が抽出された。

2-3. リガンド結合部位の定義

561 個のタンパク質について、リガンド結合部位を定義する際に、Ceriu2 LigandFit²⁾ を用いた。この LigandFit では、結合予測に用いるタンパク質が化合物との複合体であった場合、“docked ligand” という機能を用いることで、その化合物の形状、及び部位から結合部位をより厳密に定義することができる。また、複合体ではない場合、“protein shape” という機能を用いることで、その表面構造を基に化合物の結合部位を複数予測し、任意に定義することができる。そこで、今回用いた 561 個のタンパク質について、これらの機能を用いて、結合部位を定義した。その際、複合体ではない場合、既知の報告等を参考にすることで適当な結合部位を定義した。

2-4. PC グリッドコンピューティングシステム上での LigandFit による親和性予測

用意したタンパク質と化合物について、親和性を予測するために LigandFit を用いた。これらの計算は PC グリッドコンピューティングシステム (“cell computing”) を用いて、400 台の PC 上で約 6 ヶ月間行った。

2-5. 相互作用既知のデータセットの作成

今回用いた 561 個のタンパク質と 2140 個の化合物の対について、相互作用することが報告されている対があるかをデータベース等で検索した。参照データとして、MDDR に記載されている注釈、The Merck Index、Therapeutic Target Database (TTD)³⁾ のデータベース等を利用した。

集められたデータから今回の予測に関係する 84 件の相互作用既知データセットを作成した。

3. 結果

3-1. 親和性予測精度の検証

今回行った予測の精度を確認するため、GOLD の検証実験⁴⁾ に用いられた 100 個の既知複合体構造のうち、座標未定義原子などを含まない 93 個の構造を用いて検証を行った。予測結果の中から化合物の構造を比較し、RMSD の値が一番小さいものをそれぞれの結果として選んだ。選んだ結果の RMSD とその結果に対する pKi の予測値 (LigScore2) をプロットしたところ、結果の 80% は RMSD の値が 3.0 以下だった (図 1、上図)。

今回、561 個のタンパク質と 2140 個の化合物について結合予測を行ったが、1 つのタンパク質と化合物の対について、結合予測の結果の上位から 5 つまでの予測結果が得られるように設定されていた。そこで、pKi の予測値 (LigScore2) を指標として値が一番高いものを選び、プロットしたところ、結果の 60% が RMSD の値が 3.0 以下だった (図 1、下図)。

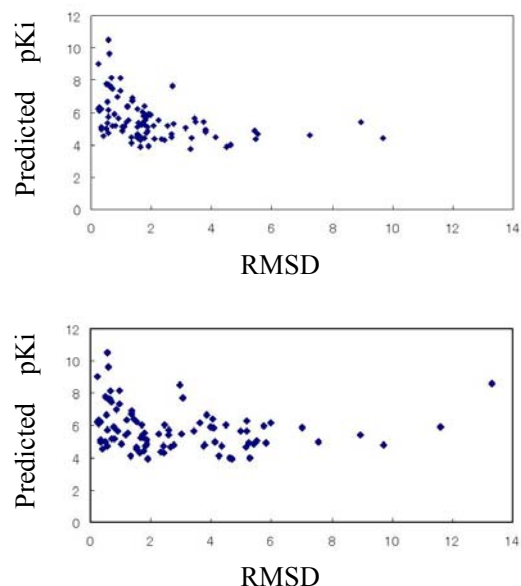


図 1 RMSD をもとに選んだコンフォメーション (上) と LigScore2 をもとに選んだコンフォメーション (下) についての結果

3-2. 親和性既知のデータセットを用いた予測精度の検証

今回行った予測結果の信頼性を検討するために、既知の相互作用についての予測結果と全予測結果をまとめた。それぞれの結果は図2のようになった(全体;mean=4.52, s. d. = 0.67, n = 1420640, 高親和性既知; mean = 5.52, s. d. = 0.77, n = 84, Goldのテストセット; mean = 5.66, s. d. = 1.31, n = 93)。

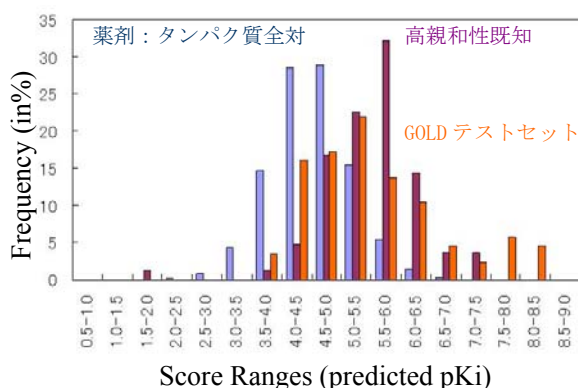


図2 結合予測値 (LigScore2) を用いた検証

3-3. タンパク質-薬剤間の新規相互作用の予測

タンパク質-薬剤間の結合予測値から親和性フィンガープリントを作成した。このフィンガープリントにカットオフ値を設定し、それ以上の値であれば高親和性相互作用候補とみなし、タンパク質と薬剤の相互作用について予測を行った。その結果、カットオフのpKi値を6.4以上とした場合、5827個の高親和性相互作用予測候補が見つかった。この結果について、さらに構造を確認していたところ、いくつかの新規高親和性相互作用予測候補となるようなものが見つかった。その中の一例として、マトリックスメタロプロテアーゼ-9 (MMP-9) と FK633 があつた。これは、予測された親和性の値が高く (LigScore2=6.46)、既知の複合体構造のペプチド鎖と良く重なるような結果であつた。

4. 考察

今回行った予測結果について、GOLDのテストセットや相互作用既知のデータセットと比較した。その結果、LigandFitによって行った今回の結合予測によって真の相互作用を判別できている可能性が期待された。そこで、作成した親和性フィンガープリントを用いることでいくつかの新規高親和性相互作用予測候補を検出することが出来た。その内、既知の複合体構造と良く重なるような結果であつたMMP-9とFK633相互作用候補について、さらに薬理メカニズムについて検討した。

FK633という薬剤は”RGD”ペプチド配列を改変して作られたgpIIb/IIIaのアンタゴニストである。gpIIb/IIIaというのは血小板、巨核球系に特異的に発現しており、機能としてフィブリノーゲンやフィブロネクチン、von Willebrand因子のレセプターとして働いている。FK633のようなgpIIb/IIIaのアンタゴニストは血小板の凝集や血小板粘着を阻害するため、抗血栓の治療等に应用されている。マトリックスメタロプロテアーゼは血管新生、癌の浸潤・転移などのさまざまな現象に関わっており、細胞外マトリックス (ECM) の分解に関わっている。MMP-9はMMP-2とともにゼラチナーゼに分類される。gpIIb/IIIaは血小板の凝集だけでなく、血管内皮細胞上のインテグリンである $\alpha_v\beta_3$ との結合により、CD40L-CD40の結合とそれに続くtPA, uPA, MMP-1, MMP-2, MMP-9の発現誘導をもたらすシグナル伝達経路を起動する⁵⁾。従って、FK633等のgpIIb/IIIa阻害剤は血小板凝集の阻害だけでなく、このシグナル伝達経路の起動をブロックすると考えられている。

今回の予測はFK633が上のシグナル伝達経路の下流でMMP-9を直接阻害する可能性を示唆している。FK633はMMP-9を直接的に阻害することによって、gpIIb/IIIaだけを阻害するよりも、より効果的にシグナル伝達経路を阻害している可能性が考えられる。今のところFK633がMMP-9に対して直接的に阻害作用を示すことは報告されていない。しかし、”RGD”ペプチド配列を改変して作られた

gpIIb/IIIa のアンタゴニストである FC-336 については、濃度依存的に MMP-9 を直接阻害することがゼラチンザイモグラフィーを用いた研究によって示唆されている⁶⁾。

5. 結論

今回はテストケースということもあり、タンパク質の数は制限されていたが、PC グリッドコンピューティングを用いて *in silico* プロテオーム規模親和性フィンガープリントを作成することが出来た。それによって、*in silico* 親和性フィンガープリンティングの有効性を確認することが出来た。また、MMP-9 阻害作用などを含む新規の薬剤-タンパク質相互作用候補を検出することが出来た。

今回行った以外にも親和性フィンガープリントを用いることによって、化合物のクラスタリング、ファーマコフォアの探索など様々な応用も考えられ、現在検討中である。今後、親和性フィンガープリントを用いることによって、さらに創薬における新しい知見が得られると期待している。

参考文献

- [1] Greenbaum, D.C., *et al.*, *Chemistry & Biology* **9**: 1085-1094 (2002)
- [2] Venkatachalam, C.M., *et al.*, *J. Mol. Graph. Model.* **21**: 289-307 (2003)
- [3] Chen, X., *et al.*, *Nucleic Acids Res.* **30**: 412-5 (2002)
- [4] Jones, G., *et al.*, *J. Mol. Biol* **26**: 727-748 (1997)
- [5] May, A.E., *et al.*, *Circulation* **106**: 2111-2117 (2002)
- [6] Y. Tsuchiya, *et. al.*, *Int. J. Oncol.* **20(2)**: 319-24 (2002)