

タンパク質内の水分子のダイナミクス解析による リガンド結合部位予測

佐々木 孝章^{†1} 関 嶋 政 和^{†1,†2}

タンパク質とリガンドの結合は、タンパク質の機能発現や創薬に重要である。タンパク 3000 プロジェクトを始めとするタンパク質の大規模な構造解析により、多くのタンパク質の構造情報が得られてきており、タンパク質とリガンドとの結合を予測し、機能発現や創薬への応用が期待されている。従来、リガンド結合部位は立体構造解析と静電場の解析から求められてきており、その結果、*in silico* による創薬では相補的構造を有する分子を中心に探索されている。本研究では、水分子が自由度を得るとエントロピー的に有利に機能することを利用し、タンパク質周囲の水分子のダイナミクス解析を行うことでリガンドの結合部位予測を行った。

Protein ligand binding site prediction by water molecule dynamics

TAKA AKI SASAKI^{†1} and MASAKAZU SEKIJIMA^{†1,†2}

Protein and ligand docking is important for expression of protein function and drug development. Protein 3000 project and other structural analysis project determined much protein structures. And this huge structural information is expected to contribute analysis of expression of protein function and drug development. Previously, ligand binding pocket analysis was based on shape complementarity and electrostatic interaction. We developed and evaluated protein ligand binding site prediction system based on water molecule dynamics around protein.

^{†1} 東京工業大学 計算工学専攻

Department of Computer Science, Tokyo Institute of Technology

^{†2} 東京工業大学 学術国際情報センター

Global Scientific Information and Computing Center, Tokyo Institute of Technology

1. はじめに

近年、日本では薬の特許切れが相次ぐ一方で、新薬の創出能力の低下が問題視されている。医薬品の開発において、ターゲットとなるタンパク質に結合するリガンドは、天然資源からの単離やハイスループット・スクリーニング (HTS : high-throughput screening) が主に使われてきたが、ハイスループット・スクリーニングは、既知の化合物のみに適用可能な方法であること、極めてコストが大きいこと、偽陽性ヒットの割合の大きいことなどが問題になっている。また、HTS では活性が高いが単位原子当たりの有効性が低い化合物から最適化を始めるために、最終化合物は分子量及び構造の複雑性の増大に伴う物性の悪化が懸念される¹⁾。この問題を解決する手法として、低分子量化合物からリード化合物を作り上げて行くフラグメントベース創薬 (FBDD : fragment based drug design) への取組みが進められている²⁾。FBDD では、活性は低い単位原子当たりの有効性が高い化合物に対して、化合物の伸張や結合等の最適化を行う為に、無駄の少ないコンパクトな最終化合物の作成を可能にする。従来のタンパク質立体構造情報に基づく薬剤設計では、リガンド結合部位の立体構造解析と静電場の解析 (鍵穴構造の解析) から、この部位での相補的構造を有する分子 (鍵) をデータベースから検索し、結合を行う³⁾。具体的には、活性部位の空洞を原子を模した球体で埋め、その球体の中心間の距離を求め、原子間距離が近い化合物を分子構造データベースから検索するという方法である。しかし、この手法では設計されるリガンドは立体構造上の相補性に依存することになり、活性が低い母化合物を拡張することで単位原子当たりの有効性が高まるかは保証されない。つまり、現状ではタンパク質とリガンドが結合した構造 (活性構造) を考慮した最適化を行うことが困難である。本研究では、水分子が自由度を得るとエントロピー的に有利に機能することを利用して、タンパク質とリガンドが結合した構造 (活性構造) を考慮したリガンド設計を可能にするために高精度なリガンド結合部位予測を行うことを目的とする。

2. 手 法

2.1 MD シミュレーションによる水分子のダイナミクスの獲得

分子動力学法 (MD : molecular dynamics) を用いることで、対象となるタンパク質の周囲の水のダイナミクスを獲得する。まず、対象となるタンパク質エネルギー最小化を 1000 ステップおこなう。次に、水分子を対象となるタンパク質の周囲に直方体状に配置する。水分子を緩和させる為に、分子動力学法を用いて 30ps かけて 0K から 300K まで圧力を一定

にして温度を上げた後、20ps の等温の MD シミュレーションを実行した。この過程を経た対象となるタンパク質について 2 種類の MD シミュレーションを行う。一つは初期から常に対象となるタンパク質の構造を固定したまま水分子のみが動くことが可能な 20ns の MD シミュレーションをおこなう方法である。もう一つは対象となるタンパク質に対し、MD シミュレーションを 10ns 間実行し、1ns 毎にスナップショットを取ることで熱揺らぎを考慮し、それらのスナップショットを固定して水分子のみが動くことが可能な MD シミュレーションをそれぞれ 10ns 実行する方法である。

2.2 水分子のダイナミクスの解析

MD シミュレーション結果から PDB ファイルを出力し、解析を行なう。対象となるタンパク質から 10ps 毎の原子の各座標を PDB ファイルとして出力する。次に、対象となるタンパク質と水からなる系を一辺 1.5\AA の立方体に細分化する。MD シミュレーションの結果毎に PDB ファイルから水がどの立方体内に存在するか数えあげ、水の向きを調べる。水の向きは酸素原子の座標と、二つの水素原子の midpoint を結ぶベクトルとして定義する。観測されたベクトルは、立方体毎に和をとる。ベクトルの和を調べる理由はカウントされた数によって正規化を行なう事ができ、正規化の結果、その大きさが 0 に近いと向きが揃っていない事になり、1 に近づくほど立方体内で同じ向きをしている事がわかる。水分子が同じ向きをしている領域では水素結合ネットワーク等が存在する事が想定できる。各シミュレーションから出力されたすべての PDB からデータを得られたら、シミュレーション時間の半分以上において水が観測された領域と水の向きの平均を解析する。

3. 実験結果

3.1 実験の条件

本研究では、対象となるタンパク質として FKBP (FK506 Binding Protein, 図 1⁴⁾)、水分子は TIP3P モデルを用いた。分子動力学シミュレーションには、AMBER10 を用いて、クーロン静電相互作用境界条件として PME 法を用いて、カットオフ距離は 10\AA に設定した。

3.2 固定した構造からの解析

20ns 間固定したシミュレーションでは、水分子が全ステップ数の半分 (1000) 以上観測された領域は 27 か所存在した (図 2 左)。その大部分はタンパク質の表面に点在しているが、リガンド結合部位の内部に留まっている領域はなかった。閾値を 800 ステップとして

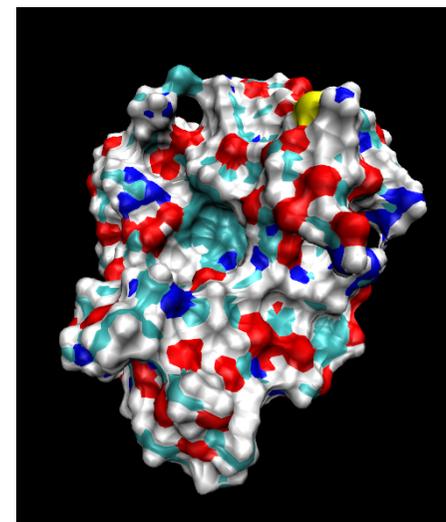


図 1 FKBP
Fig. 1 FKBP

観測された領域を調べたところ、該当する領域は 67 か所見つかるといえる。そのうち 2 か所がリガンド結合部位の内部に見つかった。(図 2 右)。なお水分子が少なくとも 1 回観測された領域における全体の平均観測ステップ数は 105 であることから、800 ステップ以上観測された領域は他の領域と有意な差があるといえる。

水の向きベクトルは、1000 ステップ以上観測された領域の平均が、0.828 であり、領域ごとに同じ向きをしている割合が多い事がわかる。また 800 ステップ以上では 0.772 という結果になった。1000 ステップ未満の領域における平均は、0.157 である。水分子の偏在する領域はでは水の向きについても特徴がある事がわかる。

3.3 熱揺らぎを考慮した解析

次にスナップショットからの解析について述べる。それぞれのスナップショットから MD シミュレーションを行ない、全ステップの半分 (500 ステップ) 以上、水分子を観測した領域は最少で 23 か所、最多で 44 か所となり幅がある (表 1)。しかしその位置は 3.2 と同様に、どの場合でもほぼすべてがタンパク質表面に点在している。そしてリガンド結合部位の

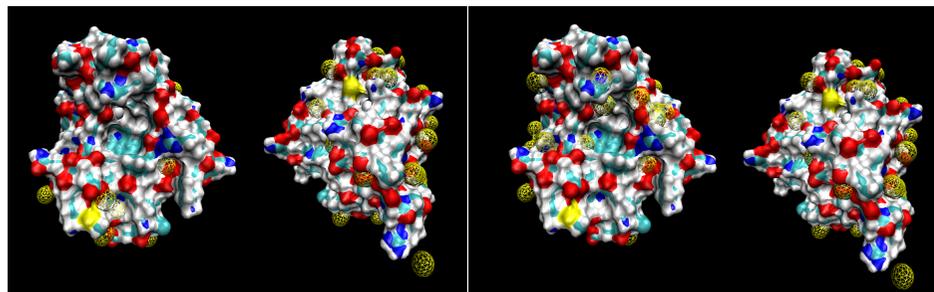


図 2 固定した構造からの解析結果. 黄色の wireframe が水分子が多く観測された領域を表す. 左側が 1000, 右側は 800 をカウントの閾値としたもの. それぞれ表と裏を示す.

Fig.2 The result is analyzing from fixed structure. yellow wireframe shows site which much of water molecule is observed.

内部には 500 ステップ以上観測された領域が少なくとも 1 か所以上存在している.

水の向きの平均については, どのスナップショットの場合でも 0.8 前後でほとんど一定である (図 3). 全体の平均は約 0.16 であり, この結果に関しても 3.2 と同じように, タンパク質が偏在する部分においては, 他の領域と比べて向きもほぼ決まっているという事を示している.

表 1 全ステップのうち半分以上水分子が観測された領域の数
Table 1 The number of site that observed water.

シミュレーション	A	B-1	B-2	B-3	B-4	B-5	B-6	B-7	B-8	B-9	B-10
結合部位の内部	0	1	2	2	1	1	3	1	1	3	2
その他の領域	27	22	40	38	34	33	31	28	37	36	42
計	27	23	42	40	35	34	34	29	38	39	44

A:20ns 固定の MD シミュレーション B-n:固定せずに n ps の MD シミュレーション

熱揺らぎを考慮した解析では, 水の偏在する領域がそれぞれ異なっている. この現象は, 各スナップショットにおいて偶然その部分に水分子の偏りができたために現れたと解釈する事ができる. 逆にいくつもの結果で同じ領域に, 水が多く観測されることがわかれば, その領域はなんらかの意味を持つと考えられる. 図 4 は, 6ns~8ns のスナップショットを利用した実験の結果である. 熱の揺らぎによって構造の変化や移動が起きている事と同時に, 水分子の観測される領域が変化している. 例えば, 図 4 の下段は左 (6ns) では中央に水分

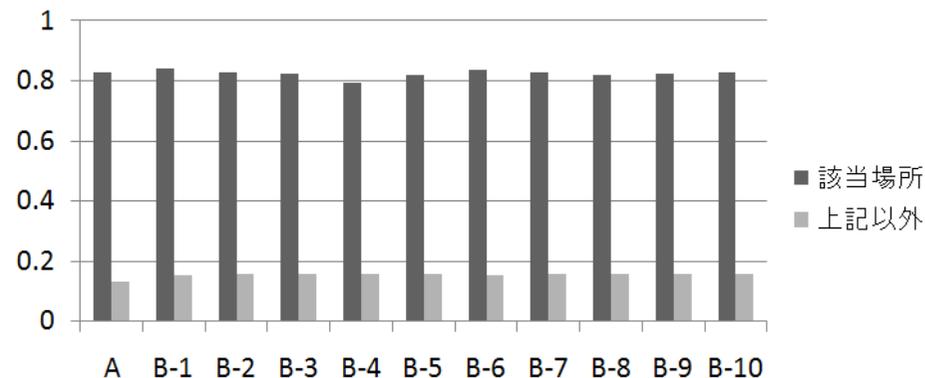


図 3 水の向きの平均を表したグラフ.

Fig.3 The graph shows average of water direction.

子の偏りが見られるが, 中央 (7ns) ではそのような領域が無くなっている. 同じく下段を見ると, タンパク質右上にくぼみがあり, 右 (8ps) では水が多く観測されたが, 6,7ps の時は閾値を超えるほど多く観測されなかった. 全スナップショットのシミュレーションを解析したところ, 1, 2 回のみ結果に現れる領域が数か所見つかった. 主な領域として, 対象のタンパク質の先端部分や, 周辺の残基が疎水性を示す領域が挙げられる. 逆に水分子が常に多く観測される領域としては親水性の残基が存在する周辺に多く見られる.

4. ま と め

タンパク質とリガンドが結合する時に, 活性部位に水和している水が外に排出されることによってエネルギー的な安定をはかる⁴⁾. この排出は結合自由エネルギーにエントロピー的にもエンタルピー的にも寄与する. なぜなら結合前では, 水和水は配向と位置の制約からエントロピー的に不利であり, 疎水性クローージャに水和した水は, 水素結合の補完が形成できないためエンタルピー的に不利となるからである.

タンパク質のリガンド結合部位の予測をするために, FKBP を用いた MD シミュレーションを行ない, タンパク質内の水分子のダイナミクスについて解析を行なった. その結果, タンパク質の表面では水分子が通常よりも多く観測される領域が存在し, リガンド結合部位で

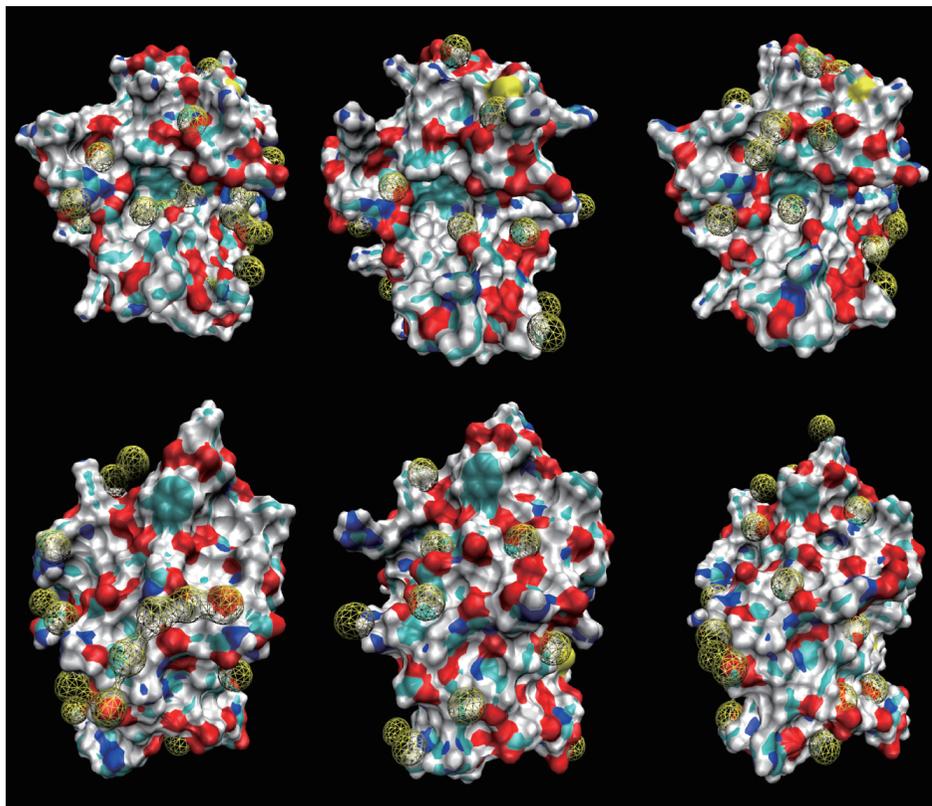


図 4 スナップショットを使った結果. 左から 6ns, 7ns, 8ns を示している.
上段はリガンド結合部位が手前側, 下段はその裏側

Fig. 4 From left to right, the results of snapshot 6ns, 7ns and 8ns are shown.
The uppers show active site side. The lower show the back.

も同様の傾向が見られた。またそのような領域の水分子はある一定の方向を向いている事も解析からわかる。一定の方向を向く事は水素結合などの分子間の相互作用により束縛されていると予想される。この領域の水はそこから離れる事によって自由度を得ることができる事から活性になんらかの影響を及ぼしていると考えられる。スナップショットを用いることで熱揺らぎを考慮したシミュレーションでは、タンパク質表面のある部分に偶然留まるような場合をある程度取り除いて、偏在しやすい領域を考える事ができる。その除去をおこなって

も、リガンド結合部位以外において偏在する領域がある場合、そのような部分は、他のリガンドやタンパク質との結合、フォールディングに重要な役割を果たす領域と考えられる。熱揺らぎを考慮した実験では、1nsのあいだに構造が大きく変わっている事もあり、結果の評価が難しい部分もある。今後は相対的な座標を考慮できるような予測を可能にする必要がある。

参考文献

- 1) Oprea, T.I., Davis, A.M., Teague, S.J. and Leeson, P.D.: Is There a Difference between Leads and Drugs? A Historical Perspective, *J. Chem. Inf. Comput. Sci.*, Vol.41, No.5, pp.1308–1315 (2001).
- 2) Shuker, S.B., Hajduk, P.J., Meadows, R.P. and Fesik, S.W.: Discovering High-Affinity Ligands for Proteins: SAR by NMR, *Science* 29 November 1996, Vol.274 No.5292 pp.1531–1534 (1996).
- 3) Kuntz, I.D.: Structure-Based Strategies for Drug Design and Discovery *Science* 21 August 1992, Vol.257 No.5073 pp.1078–1082 (1992).
- 4) Van Duynne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., Clardy, J.: Atomic Structure of FKBP-FK506, an Immunophilin-Immunosuppressant Complex *Science* 10 May 1991, Vol.252 No.5007 pp.839–842 (1991).
- 5) Abel, R., Young, T., Farid, R., Berne, B.J. and Fresner, R.A.: Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding *J. Am. Chem. Soc.*, Vol.130, No.9, pp.2817–2831 (2008).