

—基礎科学から医学・医療を見る—

タンパク質へのリガンド結合に関する計算物理・化学的アプローチ

藤崎 弘士

日本医科大学基礎科学物理学

Physics- and Chemistry-based Computational Approaches to Ligand Binding for Proteins

Hiroshi Fujisaki

Department of Physics, Nippon Medical School

Abstract

We review theoretical and computational approaches to ligand binding, one of the most relevant biomolecular events in a cell. Starting from a kinetic description of ligand binding, which is summarized by the use of the dissociation constant, we discuss simple docking simulations, the Molecular Mechanics/Poisson-Boltzmann Surface Area (Generalized Born Surface Area) approximation for binding free energy (intermediate level of approximation), and more rigorous free energy profile calculations, which will be used in the near future for designing and discovering drugs.

(日本医科大学医学会雑誌 2013; 9: 135-139)

Key words: ligand binding, binding free energy, docking simulations, molecular dynamics, Molecular Mechanics/Poisson-Boltzmann Surface Area (Generalized Born Surface Area), free energy profile

1. リガンド結合とは何か

タンパク質は20種類のアミノ酸がつながり、それぞれが決まった3次元構造をとることによって特定の(ただし一つとは限らない)機能を発揮する。タンパク質の機能としては、大きく分けて、1) 分子認識(あるタンパク質が特定のリガンドや、別のタンパク質、DNAの特定の部位とだけ相互作用する)、2) 触媒作用(タンパク質内の特殊な環境下に分子を置くことで、その化学反応を飛躍的に加速する)、3) スイッチ機能(リガンド結合などによってタンパク質が構造変化し、それが別のタンパク質と結合することで、さらなる構造変化を促し、といった連鎖によってシグナル

伝達を行う)、4) 細胞内の構造を支える(モータータンパク質のレールとしての役割をもつアクチンタンパク質など)といったものがある¹。タンパク質の構造からこういった機能がどのように生じるのか、原子・分子レベルから調べるのが構造生物学、もしくは分子生物学と呼ばれる分野であり、免疫学²、がんの生物学³といった病理に関わる分野においても、これらの基本的な理解は重要である。

しかし、こういったことを分子レベルで調べるためには、実験だけでは様々な限界があり、それを補うために、コンピュータ内で生体分子の運動をモデル化して調べるというアプローチが有効になる。こういった計算物理・化学のターゲットとして、現在ではウイルスのような巨大分子(数百万個以上の原子を含む)、

Correspondence to Hiroshi Fujisaki, Department of Physics, Nippon Medical School, 2-297-2 Kosugi-cho, Nakahara-ku, Kawasaki 211-0063, Japan

E-mail: fujisaki@nms.ac.jp

Journal Website (<http://www.nms.ac.jp/jmanms/>)

非常にゆっくり進行するタンパク質の構造変化(フォールディングを含む), もしくは薬とタンパク質の間の相互作用(その帰結としての理論的なドラッグデザイン)などが精力的に調べられている。

本稿では, これらのトピックスの中から, 機能的にも重要で, 薬学的にも興味のある, リガンド結合について議論する。リガンド結合は上で述べたように, 分子認識やスイッチ機能と関連して非常に重要なものである。また, 触媒作用が起こる際にも, 最初はその基質をタンパク質に結合させる必要があるため, リガンド結合はタンパク質の機能の大部分と関わっているといっている言い過ぎではない。ここで, リガンドとは狭義には, タンパク質と結合する小さな化学物質を意味する(たとえば, 小さなタンパク質ミオグロビンには2原子分子がリガンドとして結合する)が, 以下の議論では, リガンドが別のタンパク質やDNAであったりしても構わない(もちろん計算するのは大変になる)。また, リガンド結合によって, タンパク質に構造変化が引き起こされ, それがシグナル伝達機能や触媒作用にもつながっていくが, 構造変化に関しては稿を改めて論じたい。また, 本稿では著者の専門である, 計算物理・化学的な側面に限って論じるが, できるだけ分かりやすくかみ砕いた説明を心掛けるつもりである。

2. リガンド結合の速度論と結合自由エネルギー

まず, 原子・分子レベルの詳細な説明に入る前に, リガンド結合のマクロな現象論について説明する。これは一見単純に見えるが, 非常に奥深いものを含んでおり, また酵素反応といったほかの現象を統一的にとらえることも可能になる。

試験管に入れたリガンドLとレセプタータンパク質(以下, レセプターと略す)Rが水中に混在している状況を考える。このとき分子レベルではLとRはくっついたり, 離れたりを繰り返しており, またその結合のプロセスは複雑である。それを模式的に, 以下のように表すことが多い。



しかし, ここでは話を単純化して, L, R, RLの濃度だけに興味があるとし, それらを[L], [R], [RL]で表そう。すると, [RL]の時間変化は以下の微分方程式でよく記述できる。

$$\frac{d[RL]}{dt} = k_{on}[R][L] - k_{off}[RL] \quad (2)$$

この右辺の第一項はRとLが会合し, さらに結合

することでRLが作られることを意味しており, 第二項はRLがある割合で解離していくことを意味する。会合する確率は[R][L]に比例すると考えられるので, その比例係数を k_{on} とする。また解離は[RL]に比例すると考えて, その比例係数を k_{off} とする。すると, 平衡状態では濃度のようなマクロな量の時間変化はなくなっているため, (2)式の左辺は0であり, その結果, 以下の解離定数 K_d を定義できる。

$$\frac{[R][L]}{[RL]} = \frac{k_{off}}{k_{on}} = K_d \quad (3)$$

ここで解離定数に関して重要なことは, (a)濃度の次元をもつ, (b)解離定数が小さいほど, 結合しやすいことを意味する, (c)解離定数は平衡状態の量だが, それは k_{on} のような時間変化を記述するrateと関係があるということである。また, 実験的には, K_d は滴定実験によって容易に測ることができるということのもポイントである⁴。レセプターに結合するのが, 阻害剤(inhibitor=I)の場合は, (1)~(3)式のLをIに変えればよい。その場合, (3)式の右辺は阻害定数 K_i と呼ばれる^{5,6}。

解離定数 K_d は結合自由エネルギー ΔG_0^{bind} と呼ばれる熱力学的な結合の強さを表す量と以下の関係にある⁴。

$$\Delta G_0^{bind} = k_B T \log \left(\frac{K_d}{1 \text{ M}} \right) \quad (4)$$

ここで k_B はボルツマン定数, T は絶対温度である。 K_d が濃度の次元をもつために, 1Mという基準の濃度で割っているが, これは本質的なことではない。というのも(4)式が直接測られることはなく, リガンドを変えたときの相対変化を測ることがほとんどだからである。例えば, リガンド a と b の結合自由エネルギーの差は(4)から

$$\Delta \Delta G_{ab}^{bind} = k_B T \log \left(\frac{K_d^{(b)}}{K_d^{(a)}} \right) \quad (5)$$

となる。つまり, リガンド b のほうが a よりレセプターに強く結合する場合, $K_d^{(b)} < K_d^{(a)}$ であるため, $\Delta \Delta G_{ab}^{bind} < 0$ となる。結合自由エネルギーはその名の通り, エネルギーの次元を持ち, kcal/molの単位で測られることが多い。室温(=300K)をエネルギーに換算すると0.6kcal/molほどなので, このエネルギーと結合自由エネルギーを比較することで, 室温での結合の安定性について定量的に議論することが可能となる。

ドラッグデザインするには以下のような状況を考える: レセプターは何らかの機能をもっているタンパク質であり, リガンドはそのレセプターに結合して, その機能を促進する, もしくは阻害する。よって, ド

ラッグデザインとは結合自由エネルギーができるだけ小さくなる(あるリガンドを基準にして、できるだけマイナスになる)リガンドを見つけるということにほかならない。実験的には様々なリガンドを用意し、 K_d を測定することで、その中から最も解離定数の小さいリガンドを薬の候補として拾い出す。その際は、薬とタンパク質の間の分子間相互作用(水素結合、ファンデアワールス力、疎水性相互作用など)がもちろん重要である。しかし、解離定数は熱力学的な(マクロな)量であり、リガンド結合の分子論的な(ミクロな)理解を K_d のみから行うことは困難である。そこで、X線回析からの結晶構造などの情報と、分子動力学などの計算的なツールを組み合わせることで、リガンド結合のより進んだ理解や、よりよいドラッグデザインの可能性が期待されている。一部に、タンパク質などの結晶構造「だけ」が分かれば薬の機序が分かるという誤解がある。しかし、実際は結合自由エネルギーというマクロな量を概算する必要があり、そのためには分子動力学のような、ミクロとマクロをつなぐ方法論(実際は分子動力学と統計力学を併用する)が必要不可欠である。

その際に重要なポイントは、二つある。一つはリガンドがレセプターのどこにどのように結合するか(それを結合様式 binding mode と呼ぶ)知るということであり、もう一つはその結合自由エネルギーを(できるだけ正確に)見積もる、ということである。こういった計算をコンピュータ上で行う場合は、生体分子が複雑なこともあり、用途に合わせて様々なレベルの近似が用いられる。もしリガンドの分子構造のみが分かっている、レセプターの分子構造が分からない場合は、QSAR と呼ばれる統計学的方法がとられることが多い⁷。以下ではリガンドとレセプター両方の分子構造が(水素まで含めて)分かっている場合のアプローチに関して述べる。こういった、分子構造に基づいて薬を探す手法は一般に structure-based drug design (SBDD) と呼ばれている。

3. 結合様式や結合自由エネルギーに対する 計算物理・化学的アプローチ

3.1 ドッキング・シミュレーション

リガンド結合のような分子認識の際に最も基本となるのは、Emil Fischer が最初に唱えた「鍵と鍵穴(lock and key)」の概念である。これは2つの分子が特異的に結合する際には、その結合部位において、2つの分子が相補的な形になっているはずだという考えであ

り、多くの場合この考えは成り立つ(もちろん例外もある)。そこでリガンドとレセプターの構造が分かっているときに、これらが相補的になるような配置を探すというアプローチが自然に考えられるが、これを実現したものがドッキング・シミュレーションと呼ばれるものである。この計算を実行した結果、結合様式の候補が得られる。また、近似的に結合自由エネルギー(に対応するもの、スコア関数と呼ばれる)を計算することで、これらの候補の中の最善のものを取り出すこともできる。

このタイプの計算では、リガンドはフレキシブルではあるが(といっても動きに制限を加えることも多い)、レセプターのほうは分子構造が固定されていることが多く、リガンドとレセプターが結合するときのレセプターの変化をとらえることは難しい。また、ドッキング・シミュレーションで計算される結合自由エネルギー(スコア関数)はあくまで近似(系統的な近似でない場合が多い)であり、定量性は乏しい。しかし、以下の定量的な計算に比べると計算コストは著しく小さく、大量の化学物質の中から薬の候補を選び出す(スクリーニングする)ときや、まず最初に当たりをつけるときにこの手法を用いることが多い。

ドッキング計算のためのソフトウェアとしては、DOCK⁸、AutoDOCK⁹などがよく使われており、またドッキング・ソフトウェアをほかのソフトウェアと組み込んで販売しているものとして Discovery Studio¹⁰、MOE¹¹、GOLD¹²などがある。

3.2 MM-PBSA (GBSA) 近似による結合自由エネルギーの計算

熱力学の基本原則から、自由エネルギー変化 ΔG はエンタルピー変化 ΔH とエントロピー変化 ΔS を使って、

$$\Delta G = \Delta H - T\Delta S \quad (6)$$

と分解できるので、このそれぞれの項を求めて、全体の自由エネルギーを計算するということが考えられる。MM-PBSA (GBSA) 近似と呼ばれるアプローチではそれを実行する。ここで ΔH はさらにリガンド・レセプター間の相互作用から生じる項と、リガンド、レセプターと溶媒分子(水分子であることが多い)との相互作用から生じる項に分かれる。その前者を計算するのに、molecular mechanics (MM) と呼ばれる経験的なポテンシャル関数を用い、後者を計算するのに、溶媒を均一なものと近似して、Poisson-Boltzmann (PB) 方程式、もしくは Generalized Born (GB) 方程式を用いる。また、分子の表面(surface area=SA)

から計算される寄与はSAで表されており、以上のことからMM-PBSA (GBSA)という略語がついている。ただし、エントロピー変化 ΔS の部分は計算が難しいため、近似的に計算する(具体的には基準振動解析を用いて調和近似のもとで計算する)か、リガンドを変えたときのエントロピー変化「の変化」 $\Delta\Delta S$ は小さいと思って無視するが多い。

このMM-PBSA (GBSA)の計算は分子動力学(molecular dynamics=MD)計算と並行して行うことが多く、そのためにMD計算用のソフトウェアに実装されていることが多い(例えば、Amber¹³など)。また、分子レベルの詳細な情報を保ったまま計算できるので、薬のどの部位が結合において重要であるか、レセプターの方のどの残基が重要であるかといった細かな情報も、結合自由エネルギーを分割することで調べることができる。

ただし、(計算コストの低い)ドッキング・シミュレーションより必ずしも精度がいいわけでもなく、結合自由エネルギーの値は数十kcal/molの誤差が含まれる場合もある。実際の実験では K_d で考えたときに、10倍程度の変化を説明したいことがあるが、これは式(5)を使うと、1.4kcal/molの変化を正確に計算しなければならぬことになり、これはMM-PBSA (GBSA)近似の精度では難しい。また、値そのものを信用しないとして、実験から得られる結合自由エネルギーとの相関をとってみても、それが非常に低くなることもある¹⁴。よって、この計算だけからドラッグデザインを行うことはできず(ただし、スクリーニングには威力を発揮する¹⁵)、ほかのもっと洗練された方法や実験結果と組み合わせることが必要になるだろう。

MM-PBSA (GBSA)近似の利点であり、同時に問題点となるのは、自由エネルギーを分割してしまうということである。エンタルピー差やエントロピー差の計算に誤差があると、その合計の自由エネルギー差に大きな誤差が含まれることになってしまう。例えばエンタルピー差ではPB、GBの近似を行っているが、これは溶媒から来る自由エネルギーの寄与を正確には見積もっていない。また、エントロピー差も非常に近似的な計算になっている。それ以外にも、MMで計算されるエネルギーの精度や、結合様式がきちんと求められているのか、といったMM-PBSA (GBSA)近似とは直接関係のない誤差の問題もある。

そこで、以下ではより精密に結合様式や結合自由エネルギーを計算し、成功を収めているいくつかのアプローチについて紹介する。

3.3 より精密な計算手法：最近の発展

ドッキング・シミュレーションではレセプターを剛体として取り扱ったり、リガンドとレセプターの間の相互作用をコストの観点から近似的に取り扱ったり、といった様々な近似が入っていた。しかし、レセプターも運動しており、MM力場を使って種々の相互作用をきちんと記述するという、最も直接的なアプローチを考えることができる。これはリガンドとレセプターの複合体(溶媒も含む)のMD計算をすることにほかならない。原理的には、非常に長時間のMD計算を行えば、その中でリガンドとレセプターが安定に結合する配置を見出すことは可能である。ただし、その時間が非常に長くなる(1ミリ秒以上かかることもある)ことが予想されるのと同時に、1回きりのMD計算では統計的に有意なことが言えないという問題がある。現在は高速で安価な並列型の計算機が専門家でもなくても使えるような状況になってきており、将来的にはこういう直接的なアプローチを行う(つまり、何千何万という初期状態からMDを始めて、それをできるだけ長時間行い、その結果を統計処理する)ことも可能になるだろうが、現在はその問題を回避するために、統計力学的な考えに基づく効率的なアプローチがとられることが多い。

統計力学的には、安定な配置というのは結合自由エネルギーが最も小さくなる状態である。そこで、MD計算を自由エネルギーを計算するためのツールだと割り切ると、自分が見たい領域の自由エネルギープロファイル(平均力ポテンシャルとも呼ばれる)を計算することができる。これはある適当な座標 ξ (見たい現象を見るために必要な座標であり、反応座標と呼ばれる)に沿って得られる自由エネルギー $F(\xi)$ のことである(その数学的な定義については、文献4などを参照)。例えば、リガンドとレセプター間の距離をこの ξ として定義すれば、自由エネルギープロファイルが最小になる距離のところに、一番安定にリガンドは存在することになる。

しかし、これを単純なMD計算で効率よく行うことはできず、実際は系に拘束をかけて計算し、それを後で補正するという統計力学に基づくアプローチが取られる。例えば、最近の計算ではレプリカ交換アンブレラサンプリング法と呼ばれる方法を用いて計算し、いくつかのリガンドとタンパク質に対して結合様式を計算したところ、実験との非常によい一致が得られている¹⁶。また、結合様式が分かっているときの、結合自由エネルギーそのものの計算に関しても、最近は非常に精度のよい計算が可能になっており、その誤差は

1 kcal/mol 程度と見積もられている^{17,18}。現在はスーパーコンピューターや GPGPU といった超並列型の計算機やデバイスが開発されてきており、数年前と比べるとこういった精密な計算も容易になりつつある（もちろん計算量は多く、並列計算機を長時間使わなければならないが）。よって、近い将来に基礎研究の段階を終え、ドラッグデザインに上で述べたような計算物理・化学の手法が本格的に利用されることになるのではないかと思われる。

文 献

1. Petsko GA, Ringe D: タンパク質の構造と機能. 2005; メディカル・サイエンス・インターナショナル.
2. Janeway's 免疫生物学 (第7版). 2010; 南江堂.
3. Weinberg RA: がんの生物学. 2008; 南江堂.
4. Zuckerman DM: Statistical Physics of Biomolecules: An Introduction. 2010; CRC Press.
5. Kikuchi H, Fujisaki H, Furuta T, Okamoto K, Leimkühler S, Nishino T: Scientific Reports 2012; 2: 331; DOI: 10.1038/srep00331.
6. 藤崎弘士, 古田忠臣, 岡本 研, 菊地浩人: 日本医科大学医学会雑誌 2012; 8: 222-227.
7. Gramatica P: QSAR & Comb Sci 2007; 26: 694-701.
8. <http://dock.compbio.ucsf.edu/>
9. <http://autodock.scripps.edu/>
10. <http://accelrys.co.jp/products/discovery-studio/>
11. <http://www.rsi.co.jp/kagaku/cs/ccg/index.html>
12. http://www.ccdc.cam.ac.uk/products/life_sciences/gold/
13. <http://ambermd.org/>
14. Hou T, Wang J, Li Y, Wang W: J Chem Inf Model 2011; 51: 69-82.
15. Okimoto N, Futatsugi N, Fuji H, et al: PLoS Comput Biol 2009; 5: e1000528.
16. Kokubo H, Tanaka T, Okamoto Y: J Comput Chem 2011; 32: 2810-2820.
17. Woo H-J, Roux B: Proc Natl Acad Sci (USA) 2005; 102: 6825-6830.
18. Fujitani H, Tanida Y, Ito M, et al.: J Chem Phys 2006; 123: 084108.

(受付: 2012年11月30日)

(受理: 2013年1月25日)