

創薬のためのドッキング法

平山 令明

Docking Method for Drug Discovery

Noriaki HIRAYAMA

*Basic Medical Science and Molecular Medicine, Tokai University School of Medicine,
143 Shimokasuya, Isehara City 259-1143, Japan*

(Received July 31, 2006)

The effective integration of detailed structural information with computational chemistry, medicinal chemistry, and informatics transforms the dream of virtual screening into reality. One of the most important technologies essential for virtual screening is an effective docking method to find molecules that efficaciously interact with their target molecules. Since an efficient docking method can be a powerful tool for virtual screening, many different approaches to solving docking problems have been proposed. Docking problems have not yet been solved and none of the currently available programs are perfect in predicting all possible scenarios. Despite the limits and imperfections of the methodology, currently available docking methods are very useful for drug discovery. The basic principles and limits of docking methods together with matters for attention in applying the methods are described in this paper.

Key words—docking; protein-drug interaction; drug discovery; *in silico* screening

1. 序 論

多くの疾病は特定の蛋白質の作用と密接に関連している。そうした蛋白質は医薬分子の標的になることから、ここでは標的分子と呼ぶことにする。標的分子の機能を制御するのが医薬分子である。標的分子の立体構造と機能の仕組みが理解できれば、医薬分子を合理的に探索することが可能になる。標的分子の構造と機能を明確に意識して行う医薬分子探索の方法は英語で *structure-based drug design* (略して SBDD) と呼ばれる。適当な日本語訳がないので、以下 SBDD と呼ぶことにする。

ゲノム解析による疾病関連遺伝子の同定が実用レベルになり、かつその遺伝子がコードする標的分子の発現と構造解析が容易になった現在、これまで SBDD の適用を阻んできた大きな障壁は実質的に取り除かれようとしている。SBDD が提唱された 20 年前には、コンピュータの能力とコストがもう 1 つの重大な問題であったが、この問題も現在ではほ

とんど解消されている。研究資金が豊富にある医薬専門企業でのみ SBDD は行えるという時代は既に過去の話になったとも言える。また、医薬分子の研究開発費が年々巨額化する中で、上流の研究の効率化を一段と進める上でも SBDD への期待は非常に大きい。さらに、稀少疾患に関連する疾病遺伝子の解析が、特にアカデミアの中で進む中で、アカデミアを中心とした創薬研究が今正に現実化しようとしているが、資金不足のアカデミアで現実的な創薬研究を展開する上で、SBDD の活用は必須とも言える。

SBDD にはいくつかのアプローチがあるが、疾病と直接的に関係する標的分子と候補医薬分子の結合性を主として分子力場法を用いた理論計算で予測する手法はドッキング法と呼ばれる。日本語に適切な訳がないので、以降このままドッキング法と呼ぶことにする。ドッキング法の適用は、標的分子の立体構造が求められていることが前提条件であり、標的分子の立体構造を強く意識しながら医薬候補分子を探索する方法である。ドッキング法は、標的分子と医薬候補分子（以下リガンド分子と呼ぶ）の相互作用を原子レベルで取り扱う。したがって、その成果を視覚的に理解することが容易であるという大き

東海大学医学部医学科基礎医学系分子生命科学 (〒259-1143 伊勢原市下糟屋 143)

e-mail: hirayama@is.icc.u-tokai.ac.jp

本総説は、日本薬学会第 126 年会シンポジウム S31 で発表したものを中心に記述したものである。

な魅力がある。

本稿では、ドッキング法の概要、適用する上での注意点、実際の適用例そして最後に問題点について述べることにする。

2. ドッキング法の概要

この節では、簡単にドッキング法の概要を述べる。ドッキングはおおむね次の6段階の操作で行うのが一般的である。

第一段階：標的分子構造の原子座標の準備 多くの研究者はX線結晶解析で得られた原子座標を用いる。NMR解析で得られた座標やホモロジー・モデリングなどのモデリング法で得られた座標を用いてもよい。しかし、座標の精度がX線解析で得られたものよりかなり劣るので、結果の信頼性もまた一般に低いと予想される。本稿では、すべてX線結晶解析の結果を用いることを前提とする。

第二段階：標的分子中の水素原子位置の推定 X線解析は精密な構造解析の手段であるが、多くの限界も持っている。一番大きな限界は水素原子の位置がX線解析では求められないことである。低分子有機化合物の場合、結晶が良好であれば、何とか求めることができるが、蛋白質X線解析では全く不可能と考えてよい。しかしドッキング法を適用する場合には、水素原子の位置に関する情報は不可欠である。そこで、まず水素原子の位置を求める。多くの水素原子の位置は隣接する非水素原子の位置に基づいて幾何学的に計算できるが、ヒドロキシル基などの水素原子の位置は、幾何学的な計算では求められないので、まず適当な位置に水素原子を結合させ、その水素原子の最適な位置をあとで述べる構造最適化の方法で求める。分子力場法の限界で、構造最適化は初期に仮定した水素原子の位置に大きく影響されるので、最適な水素原子が求められたかどうかは何らかの手段で確認しておく必要がある。

第三段階：リガンド分子の結合部位の決定 適当なリガンド分子と標的分子の複合体結晶の解析が行われている場合には、結合部位が既に分かっているので、その情報を活用する。もしアポ蛋白質の構造解析しかない場合には、結合部位を推定しなければならない。リガンド分子が結合する部位を予測する手法は色々提案されているが、医薬分子としての適性を具えた化合物の結合部位を高い確率で予測できる方法が筆者らによって最近提案されている。¹⁾

しかし、多くのドッキング計算では標的分子とリガンド分子の複合体のX線解析から決定した結合位置を使用している。

第四段階：リガンド分子の立体配座の推定 標的分子に結合するときに、リガンド分子がどのような立体配座を取るかをドッキング前にあらかじめ予測することは困難である。リガンド分子は標的分子に結合したときに非常に不安定な立体配座を取ることができないが、標的分子と結合することである程度安定化できるので、最安定な立体配座を取る必要はない。実際に複合体結晶中でリガンド分子が取る立体配座は最安定な立体配座ではない。筆者らは、後述するようにリガンド分子が取り得る立体配座をあらかじめ発生させ、それをデータベース化する方法を取っている。このときの大きな問題点は、必要な立体配座が本当に発生できているかということである。

第五段階：標的分子とリガンド分子の複合体構造の最適化 標的分子の結合部位にリガンド分子を配置し、結合部位で最適になるリガンド分子の構造を求める。現在、最適化はもっぱら分子力場計算で行われる。例えば、標的分子とリガンド分子間の相互作用エネルギーとリガンド分子のひずみエネルギーを極小化することで、複合体構造の最適化が行われる。分子間相互作用は静電相互作用エネルギー、ファン・デル・ワールス力及び水素結合エネルギーなどで表現される。この段階が文字通りドッキングである。もし、結合部位との相互作用が不利であれば、そのリガンド分子はその結合部位に結合できないことになる。実際に結合部位に収まりきらない分子をドッキングさせると、その分子は結合部位から吐き出されてしまう。この最適化の過程で、リガンド分子だけでなく、標的分子の構造も同時に最適化することも行われる。

第六段階：ドッキング結果の評価 標的分子とリガンド分子の親和性は、これらの分子の結合に伴う自由エネルギー変化を正確に計算できれば容易に求めることができる。しかしエントロピー項の計算が事実上できないこと、分子力場の計算に用いるポテンシャル関数の不正確さ、水や対イオンの影響が見積もれないことなどにより、正確に自由エネルギーを求めることができない。したがって、種々の評価関数（例えば水素結合の本数）などがドッキン

グのアルゴリズム毎に用意されており、ドッキング計算の結果を評価するのに使用される。例えば複数のリガンド分子とのドッキング計算に基づき、その中から最も結合性の高いリガンド分子を選択する場合に、この評価関数で判断する。後述するように、筆者らはもっぱらリガンド分子のひずみも含めた相互作用エネルギーのみで結果を判断している。

ドッキングを行うソフトウェアは種々開発されている。代表的なものに、DOCK,²⁾ FLEXX,³⁾ AutoDock,⁴⁾ GOLD,⁵⁾ GLIDE⁶⁾ 及び Ph4Dock⁷⁾ などがある。それぞれのソフトウェアは各々特徴を持っている。後述するように、ドッキング計算は、使用する標的分子の構造の正確さに大きく依存する。しかし、適切な系に対して十分注意して計算を行うと、十分 *in vitro* 実験に匹敵する良好な結果が得られる。

3. 標的分子構造の評価

ドッキングでは実験的に決定された標的分子の構造を用いる。しかし、X線解析の結果には多くの偶然誤差と系統誤差が入っているため、原子パラメータ（原子座標と後述する原子変位パラメータ及び占有率を含めたもの）の取り扱いには十分な注意が必要である。水素原子の問題については既に触れたので、ここでは割愛する。X線解析は確かに厳密な方法であるが、蛋白質結晶の本質的な問題により、有機低分子結晶の場合に比較して原子パラメータの精度が非常に悪い。

蛋白質結晶には多くの水が含まれている。40 から 60% もの水が蛋白質結晶に含まれるが、回折 X 線により位置が確定できる水分子はごく少数である。つまり残りの水は、明確に回折 X 線に反映せず、漫然とした誤差を全原子の原子パラメータに与えることになる。また分子表面にある蛋白質の原子は大きな運動性を持っていることが多く、蛋白質結晶のすべての単位胞中で同じ立体配座を取っていることはないと推定される。こうした条件は、蛋白質結晶からの X 線回折を劣化させる。具体的には、細かい距離の情報を与えるはずの回折角度の大きい回折 X 線がほとんどゼロになってしまう。これは測定できる回折 X 線の数が劇的に少なくなることを意味する。この 2 つの悪条件の結果、蛋白質結晶から得られる回折 X 線の分解能はせいぜいよくて 1.5 Å ということになる。この数字を文字通り

解釈すると、炭素原子間の単結合と二重結合の判定ができないことになる。実際に、もしアミノ酸やリガンド分子の化学構造が既知でなければ、こうした結合の差を実験的に決定することは非常に困難である。また X 線解析では結晶中の電子密度の分布を求め、その密度の差によって原子種の判定が本来行えるはずであるが、蛋白質 X 線解析で得られる電子密度は誤差を多く含むために、その客観的な解釈は困難である。酸素原子と窒素原子の電子数の差は 1 個しかなく、状況によっては両者を判別することはできない。実際に、多くの解析結果でグルタミンやアスパラギン側鎖の酸素原子と窒素原子の帰属は不確かである。電子密度と結合距離 (C=O と C-N) からは X 線解析でこれらの原子種を確実に帰属することはできない。多くの解析では、周囲の状況を考慮し、水素結合の形成の予測から原子種の帰属をしているが、確証がある訳ではない。

結晶中と言えども原子は熱振動している。X 線解析では各原子の座標だけではなく、この運動性も同時に決定する。この運動性は温度とともに大きくなることから通常温度因子と呼ばれるが、原子変位パラメータ (atomic displacement parameter) と呼ぶ方が適切である。本稿ではこれまでの呼び方に従って温度因子と呼ぶ。有機低分子化合物の場合、室温程度で回折 X 線を測定すると、炭素原子の温度因子はおおよそ 5 Å² 程度になる。しかし、良好なデータに基づいて精密化された蛋白質結晶中でも炭素原子の温度因子は 10 Å² 以上になる。その原子がどの結晶単位胞の中でも同じ位置にあれば、その原子の占有率 (occupancy) は 1.0 となるが、原子が複数の位置に乱れて存在する (このことを disorder と呼ぶ) と、原子の占有率は 1.0 以下になる。仮に占有率が 0.5 になることは、酸素原子の場合、電子の数がその位置に 4 個しかないことになり、先に述べたように、回折データに入り込んでくる誤差をあわせて考えると、その原子の帰属がかなり危ういことが分かるであろう。さらに温度因子と占有率は高い相関を持っている。例えば温度因子が 20 Å² で占有率が 0.5 である原子は、占有率を 1.0 にすると温度因子はずっと大きくなる。また 80 Å² など温度因子が極端に大きい場合、仮に占有率が 1.0 と帰属されていてもその原子種と位置の不確定性は非常に高い。現状のドッキングでは原子の熱振動は考慮しな

いが、可能な限り正確な座標を採用するためには、座標だけではなく温度因子と占有率も十分に考慮すべきである。特に結合部位にある原子の温度因子が異常に高いかあるいは原子の占有率が異常に低い場合には要注意である。

蛋白質の X 線解析では非常に多くの原子情報が得られるので、解析者が特に興味を持った部位以外については、それらの情報の 1 つ 1 つについてその内容を詳細に検討することはあまりされない。有機低分子の X 線解析では、各原子座標の標準偏差を正確に算出するので、標準偏差が大きい部分や標準偏差から著しくずれた構造を発見し、考察することができるが、蛋白質 X 線解析では、ほとんどの解析で標準偏差は算出されないで、このような考察は行われぬのが通常である。その代わり、主鎖の立体配座を視覚的に表現した Ramachandran Plot⁸⁾などを用いて、構造全体の妥当性を議論することが多い。しかし、ドッキング計算においては、水素原子の位置まで考慮した構造を扱うため、局所的な構造の妥当性については X 線解析者以上に神経を尖らせて吟味すべきである。

標的分子の選択に幅がある場合には、上記のような点に留意して、最も適切な標的分子構造を用いることが重要である。いずれにしても、X 線解析で得られた構造を盲目的に信じるのではなく、分子間相互作用、立体配座、温度因子及び占有率などを総合的に判断して、その構造の妥当性を吟味すべきである。こうした観点からは、分解能が良好で、R 因子の小さい解析結果を活用することが望ましいが、これらの指数もけっして構造の高質性を 100% 保証するものではないことを明記しておく。

ドッキングでは、リガンド分子と標的分子の相互作用を原子レベルで議論する。したがって、各原子の位置がどの程度正確であるかをあらかじめ知っておくことは極めて重要である。蛋白質結晶の回折能の低さにより、十分な数の観測値を得ることができないので、多くの蛋白質 X 線解析では原子座標の標準偏差（誤差）は算出されない。したがって、多くの蛋白質 X 線結晶解析では各原子の不確定性（誤差）は分からない。これは、特にドッキング法を適用する場合にはかなり深刻な問題になる。

蛋白質結晶解析において、どの程度の誤差を座標は含むものかを初めて検討したのが Cruickshank⁹⁾

で、その式をより実用的な問題に適用できるように変形したのが Blow¹⁰⁾ である。詳細¹¹⁾は省くが、Blow によって提案されている式は次のようなものである。

$$\sigma(r, B_{avg}) = 2.2 N_{atoms}^{1/2} V_a^{1/3} n_{obs}^{-5/6} R_{free}$$

この式で、 B_{avg} は X 線解析から求められた温度因子 B の平均値、 N_{atoms} は規則的な構造を持った溶媒原子を含む占有率が 1.0 の原子の数、 V_a は結晶の非対称単位の体積、 n_{obs} は測定した回折 X 線の数そして R_{free} は結晶解析の確かさを表す指数である。これらの値は Protein Data Bank¹²⁾ に供託された大抵のデータの中に含まれるか、容易にそれらの情報から計算することができる。左辺の $\sigma(r, B_{avg})$ は B_{avg} を持った構造における原子の位置の誤差を示す。この式で求められる原子位置の精度の目安は、Fig. 1 から簡単に知ることができる。分解能が 2.0 Å で R_{free} が 0.20 を与える標準的な解析例では、位置の誤差は 0.2 Å と見積もることができる。つまり、0.20 Å 程度のずれでは誤差範囲ということになる。もちろん、この式で計算される値は、溶媒分子を含んだ占有率が 1.0 の全原子に対する平均的な値であるので、原子の存在する場所（特に分子表面か分子内部か）によって誤差の大きさは変化する。

従来、こうした座標誤差を見積もるために Luzzati プロット¹³⁾が用いられてきたが、Cruickshank はそれが不適切であることを指摘している。現状で

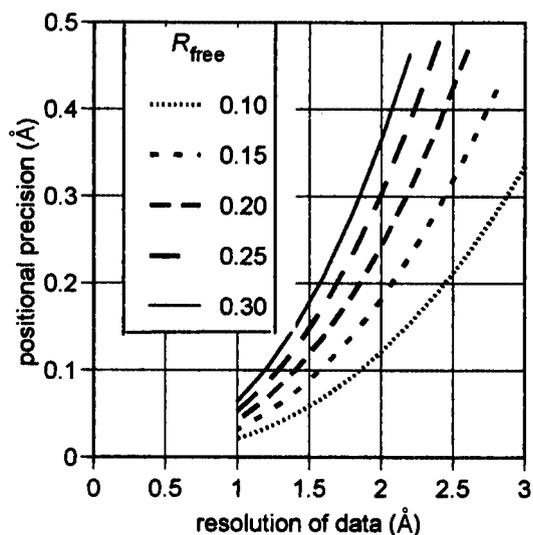


Fig. 1. Rough Guide to Precision of Atomic Position (Reproduced from D. M. Blow (2002) by permission of Oxford University Press.)

は上記の Blow による式が最も簡便かつ適切と考えられる。この誤差は回折データの性質を用いて見積もるため *diffraction component precision index* (DPI) と呼ばれる。これまでドッキングを行う場合、X線解析のデータは一律に扱われることが多かったが、DPI を求めそれに応じてドッキング結果を解釈する方が適切である。またドッキングを行うに当たり、標的分子中の原子位置の不確定性の目安である DPI をあらかじめ知っておくことは重要である。

4. Ph4Dock を用いたドッキング計算の例

筆者が現在使用しているソフトウェア Ph4Dock を用いたドッキング計算の具体例を説明する。この例では、インフルエンザ・ウイルスのノイラミニダーゼに対する阻害剤 BCX-1812 (Fig. 2) という化合物が、どのようにこの酵素を阻害するかをドッキングで推定する。実際には、この酵素と阻害剤の複合体の X 線解析は既に行われているので、どれだけ実験で求められた構造が正しく予想されるかが興味のポイントである。概要のところでも述べたようにドッキングは複数の段階に分けて実行する。Ph4Dock はソフトウェア・システム MOE¹⁴⁾ の下で動くように開発されたものであり、以下のすべての計算は MOE を用いて行った。

第一段階：標的分子構造の準備と検討 標的分子構造には Protein Data Bank の 1L7H¹⁵⁾ という構造を用いた。 R_{free} が 0.177 であることから、この構造の X 線解析は良好に行われていると考えられる。しかし、アスパラギンやグルタミンの酸素原子と窒素原子の帰属及びヒスチジンのイミダザール環の配向とプロトン化状態については、事前に吟味しておく必要がある。筆者らは、この目的に

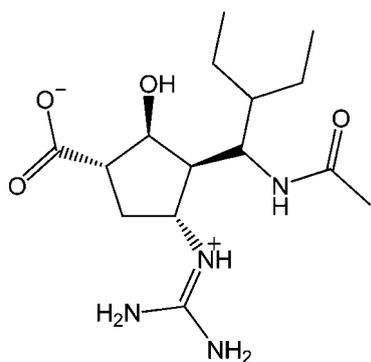


Fig. 2. Chemical Structure of BCX-1812

Reduce¹⁶⁾ というソフトウェアを使用している。このソフトウェアは、X 線解析では実験的に一意的に決定できないこれらの構造に対して、それらのアミノ酸残基周囲の状況を勘案して、妥当な構造を提案してくれる。しかし、これはあくまで 1 つの提案であり、決定的ではない。Reduce の結果は、修正した座標の形で出力されるので、便利である。これまでの筆者の経験では、Reduce による構造修正は比較的良好である。実際に、1L7H に対して、Reduce による検討を行うと、7 個のアスパラギン、2 個のグルタミン及び 1 個のヒスチジンの側鎖を 180° 回転 (フリップ) した方が妥当であると示唆された。これらの残基については Reduce による提案が妥当であるので、採用することにした。さらに 1 個のヒスチジンについてはフリップしても、依然として周囲との相互作用は芳しくないと判断されたので、X 線解析者の解釈を尊重して、元の立体配座を残した。

Reduce では最も妥当な位置に水素原子を発生させるが、一部の水素原子が発生できないので、MOE の分子構築の機能を用いて最終的にすべての水素原子を幾何学的計算で発生させた。発生した水素原子の位置を最適なものにするために、水素原子のみをあとで述べる力場計算で最適化した。これで標的分子構造の準備はできた。

第二段階：阻害剤結合部位の予測と結合部位の特性解析 結合部位が分からない状況では、標的分子の表面にあり、阻害剤の結合が可能な部位をまず探索しなければならない。その目的で、筆者らは MOE の中にある Alpha Site Finder という機能を用いる。このソフトウェアは、標的分子表面の原子から、それらの原子と直接的に相互作用する可能性のある位置を計算し、その位置を小さい球 (これを α 球と呼ぶ) で表現する。またその位置が標的分子から受ける影響により、その球を大きく親水的及び疎水的相互作用を受ける球として表現する。 α 球を近接して複数個置くことができる空間は、標的分子表面のくぼみを意味し、リガンド分子の結合部位になり得る。このような位置を α サイトと呼ぶことにする。

1L7H について求めた 1 つの α サイトの様子を Fig. 3 に示した。通常複数の α サイトが 1 つの標的分子に対して求められるが、この Fig. 3 で表示し

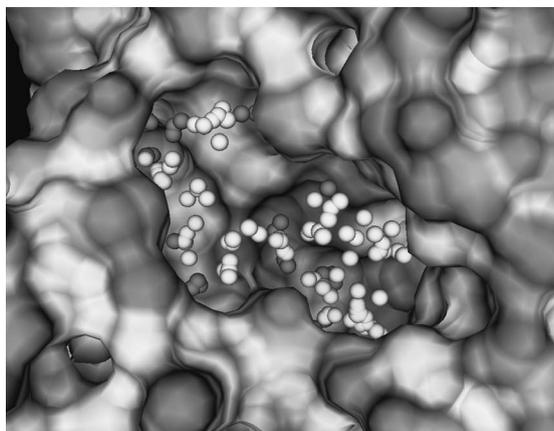


Fig. 3. α Spheres Located in a Cavity of Neuraminidase

Grey and white α spheres represent hydrophilic and hydrophobic characters, respectively. The enzyme is shown by a solvent accessible surface.

たのはその中で最も大きいものである。この標的分子の場合、他のサイトは小さすぎてリガンド分子は結合できないと判断されたので、事実上1つのサイトが求められた。このサイトにある α 球は標的分子と強く相互作用するリガンド分子の原子がくる位置と考えられる。したがって、 α 球を頼りにリガンド分子をドッキングしてもよいが、現実的にはリガンド分子の特定の原子団が α サイトの特定領域に結合すると考えられる。したがって α サイトを各 α 球周囲にある標的分子からの影響により、いくつかの領域(ファーマコフォア)に分けることができる。例えば、疎水的な領域、正の電荷を帯びた領域、負の電荷を帯びた領域などである。このような領域に分けておくと、リガンド分子がこの結合部位にどのような配向で結合するかをあらかじめ予測できるので、ドッキングの効率が向上する。Ph4Dockではこの方法を採用している。Figure 4には、 α 球をより大きな球であるファーマコフォアで表現した様子を示す。このFig. 4をみると、どの領域にリガンド分子のどのような原子団がくるかを想像できるので、医薬分子設計者にとっては非常に便利である。

第三段階：リガンド分子の可能な立体配座の発生
第二段階で同定された結合部位に結合できるリガンド分子がどのような立体配座を取るかは、ドッキング前には分からない。また最も安定な立体配座でリガンド分子は標的分子に結合する訳ではない。そこで、可能な限り多くの立体配座を発生させることが必要になる。この問題は意外に難しい。Ph4Dock

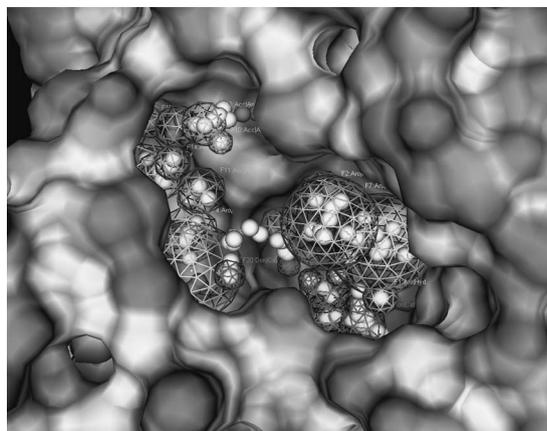


Fig. 4. Pharmacophores are Expressed by Chicken-wire Networks

Small spheres are α spheres.

には立体配座を発生させる複数の方法が用意されているが、適切な試行回数を設定すれば確率的な手法で、十分に立体配座空間を探索することが可能である。必要とされる試行回数は分子内の自由回転可能な結合の数におおよそ比例すると考えてよい。この例では250個の互いに類似度の低い立体配座を発生させた。

第四段階：ドッキング計算とその結果の検討

第二段階で求めた結合部位のファーマコフォアと第三段階で求めたリガンド分子の複数の立体配座を用いて、ドッキング計算を行う。大きく分けてこの段階は2つに分けられる。第一はファーマコフォアと立体配座の一致度を確認することであり、これにパスしたものが第二段階で分子力場法により最適化される。種々の分子力場が提案されているが、Ph4DockではMMFF94x¹⁷⁾を採用している。

第二段階で求めた結合部位はかなり広く、その中で比較的小さなBCX-1812がどの位置に結合するかは興味深い。通常ドッキング計算では、複数の結合様式が計算されるので、答えは唯一ではない。したがって、それらの中からどれが妥当であるのかを判断する必要がある。そのため、ドッキング結果を評価するために種々の評価関数が提案されている。筆者らは基本的に次式で表される U_{total} で結果を判断している。本来は自由エネルギー変化を求めたいが、エントロピー項の算出が難しいことなどから、 U_{total} で判断することになっている。

$$U_{total} = U_{ele} + U_{vdw} + U_{ligand}$$

Uele 及び Uvdw はそれぞれ標的分子とリガンド分子間の静電相互作用及びファン・デル・ワールス相互作用を示す。Uligand はリガンド分子内のひずみのエネルギーである。この計算結果では、最低の U_{total} は -18.93 kcal/mol で、このエネルギーを与えるドッキング結果が最も妥当な結果と考えられる。ドッキング計算で予測された BCX-1812 の結合部位での位置と構造を Fig. 5 に示した。幅の広い結合部位の片側奥に BCX-1812 は結合することが予想された。

ドッキングで得られたリガンド分子の結合位置と構造が、X 線解析で実験的に得られたものをどの程度再現しているかを知るためによく計算される量が、対応する非水素原子の「ずれ」の二乗平均の平方根 (root mean square deviation 略して rmsd) である。rmsd が 0 であれば、結晶解析で求められた非水素原子と完全に一致する構造がドッキングで求められたことになる。rmsd が大きくなればなるほど、ドッキング結果と実験結果の一致度は低くなる。Ph4Dock で得られた BCX-1812 では、rmsd は 0.397 Å となり、結合位置だけでなく、阻害剤の立体構造まで正確に予測できたことが分かる。Figure 6 に X 線解析で求められた BCX-1812 とドッキングで得られた構造を重ね合わせたものを示す。両者は極めて良好に一致している。この例は、Ph4Dock を使えば、X 線解析の結果にかなり近い構造をドッキングで得ることができることを示す。

5. ドッキング計算はどの程度正確か

ドッキングに関する多くの論文では蛋白質 X 線

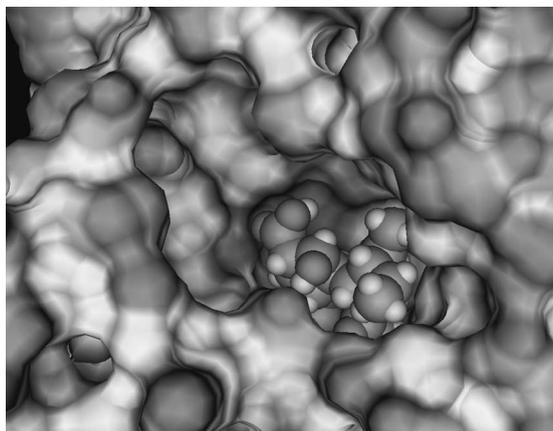


Fig. 5. The Binding Site of BCX-1812
BCX-1812 is depicted by van der Waals radii.

解析の結果とドッキングの結果の rmsd が 2.0 Å 以下になったときに、ドッキングは成功したと判断する。しかし、ドッキングの結果はドッキングに用いる標的分子の X 線解析の精度に依存するで、一律に 2.0 Å という基準は望ましくない。上述したように蛋白質 X 線解析では原子の座標の誤差が求められない。そこで筆者らは DPI に基づいて、rmsd の誤差に相当する量を計算し、それをもってドッキングの結果を評価することを提案した。⁷⁾ rmsd の誤差は $2\sqrt{2} DPI$ と見積もることができ、この値を rmsd の標準偏差と考えれば、ドッキングの結果を評価できる。先の 1L7H の場合、PDB にあるデータに基づけば、DPI は 0.171 Å と計算できる。したがって、ドッキングによって得られる構造も本質的に同じ誤差を持つと仮定すれば、rmsd の誤差は 0.484 Å となる。 U_{total} が最低であった構造に対する rmsd が 0.397 Å であったので、この rmsd はその標準偏差以下ということになる。つまりドッキングで得られたリガンド分子の構造は本質的に X 線解析で得られた構造と同じであることを示す。このように DPI を用いれば、標的分子毎にドッキングで得られた結果をある程度定量的に評価することが可能である。Figure 6 をみても分かるように、Ph4Dock で BCX-1812 をノイラミニダーゼにドッキングさせると、ほとんど X 線解析で得た構造と同じ構造が得られるが、誤差のレベルからも実験に匹敵する正確さで構造が求められている。

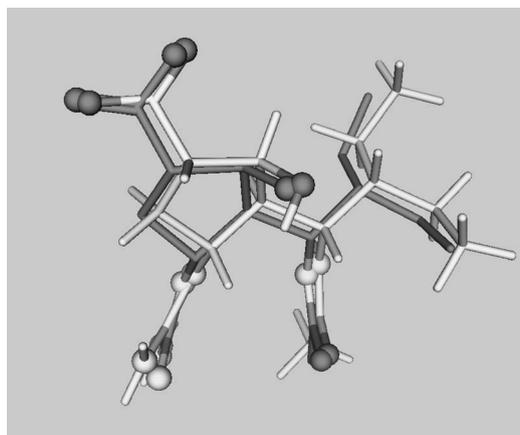


Fig. 6. A Superposition of the Predicted and Experimental Structures of BCX-1812

Oxygen and nitrogen atoms are depicted by small balls. The molecule with dark bonds is the X-ray structure and another molecule is the predicted structure by Ph4Dock.

1L7Hのように、既にX線解析で複合体の構造が分かっている構造を用いて、その構造をドッキングで再現できるかどうかを再ドッキング (redocking) によって確認することは、使用しているソフトウェアの性能や特性をチェックしたり、現在使用している標的分子の特徴を理解する上で極めて重要である。筆者らは、高質のX線解析より得られた標的分子と医薬分子らしい (drug-like) 化合物との複合体構造 43 種類を用いて、Ph4Dock による再ドッキング計算を行った。⁷⁾ その結果を Table 1 に示す。計算の条件の詳細は省くが、おおよそ前節で用いた手続きに沿って行った。ドッキング結果のうち、rmsd の値が 3.0σ (DPI) より小さい値になった構造は 25 あり、全体の 60% である。すなわち 60% の構造は X 線解析の実験誤差以内で、リガンド分子の構造と結合位置を予測できたことを意味する。また一般的な基準より厳格な $\text{rmsd} < 1.5 \text{ \AA}$ の構造を与えた結果は 34 構造あり、全体の 80% になる。これら 43 種類の複合体を構成する標的分子とリガンド分子は多様であるが、それらの系に対して Ph4Dock は良好な予測結果を出した。この結果はドッキング計算の有効性と限界をともに示しているが、正確な標的分子の立体構造があれば、少なくとも再ドッキングによるリガンド結合性の予測は十分可能なことを示唆する。

6. ドッキング法を応用する上での問題点と今後の課題

筆者らはドッキング法を用いて特定の標的分子に対する親和性を持った化合物の探索を現在行っている。この場合、多くの化合物を含むデータベースを用いて、それらの化合物のどれが標的分子との結合性が高いかを予測する。探索の対象とする化合物は実際に購入可能な化合物でも、仮想的な化合物でも構わない。ドッキング法はいったん手続きを確立すれば、多くの可能性を短時間に検証できるという極めて大きな長所を持っているので、それほど能力の高いコンピュータでなくとも、1週間に数万化合物の探索を行うことも可能である。

ドッキング法を適用する上での大きな問題点は 4 つほどある。

第一の問題点は標的分子の立体構造である。まず既に述べたように蛋白質 X 線結晶解析の限界による曖昧性があり、この問題はすぐには解決されない

Table 1. Results of Docking Predictions on 43 Complexes^{a)}

PDB code	Resolution (Å)	R_{free}	$2\sqrt{2}\sigma$ (Å) ^{a)}	rmsd (Å)
1A28	1.80	0.228	0.191	0.580
1A15	2.36	0.222	1.137	1.238
1AQW	1.80	0.261	0.447	1.165
1B58	1.80	0.223	0.583	1.344
1B9V	2.35	0.271	1.321	>2.00
1BCU	2.00	0.212	0.673	0.784
1BXO	0.95	0.125	0.045	1.215
1BYG	2.40	0.287	1.510	1.307
1C1E	1.90	0.294	0.981	>2.00
1C5C	1.61	0.253	0.600	0.568
1C5X	1.75	0.244	1.244	0.678
1C83	1.80	0.231	0.501	0.793
1CBS	1.80	0.237	0.617	1.019
1CKP	2.05	0.260	0.939	1.167
1CVU	2.40	0.235	0.896	1.000
1D0L	1.97	0.200	0.602	1.770
1D3D	2.04	0.223	0.814	>2.00
1D3H	1.80	0.185	0.385	1.956
1D3P	2.10	0.214	0.834	>2.00
1D4P	2.07	0.231	0.877	0.713
1DD7	2.25	0.284	1.307	1.209
1DG5	2.00	0.243	0.735	1.125
1E11	2.30	0.266	1.394	1.381
1EJN	1.80	0.240	0.696	0.905
1F0R	2.10	0.263	1.024	1.405
1F0S	2.10	0.263	1.007	1.533
1F3D	1.87	0.218	0.464	0.596
1FL3	2.45	0.262	1.137	1.113
1KEL	1.90	0.258	0.902	1.351
1LIC	1.60	0.225	0.436	1.480
1NGP	2.40	0.250	1.162	0.837
1QCF	2.00	0.257	0.834	0.556
1QPE	2.00	0.254	0.888	1.013
1QPQ	2.45	0.253	0.817	0.712
1WAP	1.80	0.225	0.218	1.251
1YEE	2.20	0.260	0.399	>2.0
25C8	2.00	0.292	0.922	>2.0
2ACK	2.40	0.257	1.007	0.918
2PCP	2.20	0.290	1.301	0.713
3ERD	2.03	0.248	0.905	0.898
3ERT	1.90	0.262	0.690	0.774
4LBD	2.40	0.285	1.926	0.906
2RTD	1.65	0.236	0.696	1.142
				0.577

a) $\sigma = \sigma(r, B_{avg})$

が、本稿で述べたように化学的に妥当な構造を考えることで、この問題はある程度回避できる。

第二の問題点は、標的分子とリガンド分子の相互作用の評価である。ほとんどすべてのドッキング法

が現状では分子力場を用いている。現在使われている分子力場はかなり改良されており、先の例にみるように、多様な複合体構造を再現することができる。したがって、どの分子力場でもかなり多くの系に適用が可能であることを意味する。しかし、標的分子の構造を再現するという意味では、現状の分子力場は十分であるが、多様な化学構造を持つリガンド分子に対してはかならずしも十分ではない。

第三の問題点は、リガンド分子結合による標的分子構造の変化の見積もりである。標的分子は誘導適合 (induced fit) により、リガンド分子に依存して構造変化を起こすと考えられている。しかし、具体的な構造変化をどのように起こすかについて詳細に考察した研究例はほとんどない。筆者らは HIV-1 プロテアーゼとその阻害剤 6 種の複合体構造を再ドッキングして、この系での誘導適合の考察を行った。¹⁸⁾ これらの阻害剤は分子量で 2 倍近い幅を持っている。したがって複合体の結合部位も少なからず誘導適合を受けている。しかしドッキング計算では、6 種の阻害剤の結合様式を良好に再現することができた。このことが一般的に成り立つかどうかは定かではないが、このような酵素の場合、誘導適合の影響は一般に予想されるより少ないことが考えられる。蛋白質 X 線解析のルーチン化に伴い、多くの医薬分子誘導体と標的分子の複合体結晶の構造解析がなされていると予想されるが、結果が公表されているものは非常に少ない。誘導適合に関する研究をさらに進展させる上で、これらのデータ公開が望まれる。

第四の問題点は、水やイオンの影響である。蛋白質結晶の中には多くの水が含まれていることは既に述べた。しかし大部分の水やイオンの運動性は高く、その位置を X 線解析では確定できない。一方で X 線解析で明らかにできる、標的分子との相互作用が強い一部の水やイオンの位置をドッキングの際にどのように考慮に入れるべきかは難しい問題である。いくつかの X 線解析により、特定の水分子が低分子化合物の結合により、排除されることが分かっており、これらの水分子は固定的に存在するものではない。X 線解析で求められた水分子について、筆者らはそれを含んだ計算と含まない計算を同時に行い、それらを比較することを行っている。最近の筆者らの研究では、少なくとも再ドッキングに

については結晶水を考慮した方がずっと成績がよい。イオンの扱いは難しいが、幸いにも医薬分子が結合する部位において対イオンが結合性に重大な影響を与える系に筆者らは遭遇していない。

6. 結 語

本稿ではドッキング法の概要、実際そして問題点について概観した。多くの問題点はあるものの、高質の蛋白質 X 線構造がかなり容易に得られるようになってきたこと、Ph4Dock のような強力なアルゴリズムが登場してきたこと、そしてドッキング計算を実際に行うコンピュータ能力が向上してきたことにより、ドッキングの実用的な応用は既に可能になったと言っても過言ではない。医薬分子の研究開発対象とする標的分子の構造的な特徴と低分子化合物との相互作用の様式を十分に理解して、上記の問題点を念頭におきながらドッキング法を活用すれば、HTS に匹敵する効果を生み出すことが可能であると筆者は思う。またドッキング法を応用した研究は、生体分子間の分子認識の基礎を理解する上でも非常に重要である。ドッキング法に関する研究が、ともすれば実用的な側面だけを強調され過ぎ、その基礎科学的な側面が見失われているのは残念である。ドッキングの成功率を上げることは、われわれが分子間相互作用をより正しく評価できることを意味する。

筆者のドッキング法に関する見方はいささか楽観的過ぎるかもしれない。しかし、今後の周辺科学の進展により、この方法は確実に向上し、創薬研究を強力に推進すると筆者は信じている。分子間相互作用を原子レベルで把握し、その相互作用を予測する、というのは分子科学に携わる研究者の大きな目標であり、夢である。

謝辞 Ph4Dock の開発研究は菱化システム(株)科学技術計算部の片岡良一博士及び後藤純一氏との共同研究で行われたものである。両氏に深く感謝する。またこれらの研究を遂行するに当たり、研究費の援助を得た東海大学総合研究機構及び新エネルギー・産業技術総合開発機構に深く感謝する。

REFERENCES

- 1) Soga S., Shirai H., Kobori M., Hirayama N., Abstracts of papers, the 6th Annual Meeting

- of the Protein Science Society of Japan, Kyoto, April 2006, p. 57.
- 2) Ewing T. J. A., Makino S., Skillman A. G., Kuntz I. D., *J. Comput.-Aided Mol. Des.*, **15**, 411–428 (2001).
 - 3) Rarey M., Kramer B., Lengauer T., Klebe G., *J. Mol. Biol.*, **261**, 470–489 (1996).
 - 4) Morris G. M., Goodsell D. S., Halliday R. S., Hey R., Hart W. E., Belew R. K., Olson A. J., *J. Comput. Chem.*, **19**, 1639–1662 (1998).
 - 5) Jones G., Willett P., Glen R. C., Leach A. R., Taylor R., *J. Mol. Biol.*, **267**, 727–748 (1997).
 - 6) Friesner R. A., Banks J. L., Murphy R. B., Halgren T. A., Klicic J. J., Mainz D. T., Repasky M. P., Knoll E. H., Shelley M., Perry J. K., Shaw D. E., Francis P., Shenkin P. S., *J. Med. Chem.*, **47**, 1739–1765 (2004).
 - 7) Goto J., Kataoka R., Hirayama N., *J. Med. Chem.*, **47**, 6804–6811 (2004).
 - 8) Ramachandran G. N., *J. Mol. Biol.*, **7**, 95–99 (1963).
 - 9) Cruickshank D. W. J., *Acta Cryst.*, **D44**, 583–601 (1999).
 - 10) Blow D. M., *Acta Cryst.*, **D58**, 792–797 (2002).
 - 11) Blow D. M., “Outline of Crystallography for Biologists,” Oxford University Press, 2002.
 - 12) Bernstein F. C., Koetzle T. F., Williams G. J. B., Meyer F., Bryce M. D., Rogers J. R., Kennard O., Shimanouchi T., Tasumi M., *J. Mol. Biol.*, **112**, 535–542 (1997).
 - 13) Luzzati V., *Acta Cryst.*, **5**, 802–810 (1952).
 - 14) MOE (Molecular Operating Environment), Version 2004.04, Chemical Computing Group Inc., Montreal, Quebec, Canada, 2004.
 - 15) Smith B. J., Mckimm-Breshkin J. L., Mcdonald M., Fernley R. T., Varghese J. N., Colman P. M., *J. Med. Chem.*, **45**, 2207–2215 (2002).
 - 16) Word J. M., Lovell S. C., LaBean T. H., Taylor H. C., Zalis M. E., Presley B. K., Richardson J. S., Richardson D. C., *J. Mol. Biol.*, **285**, 1711–1733 (1999).
 - 17) Halgren T. A., *J. Comp. Chem.*, **17**, 490–519 (1996).
 - 18) Muta H., Hirayama N., Abstracts of papers, the 126th Annual Meeting of the Pharmaceutical Society of Japan, Sendai, March 2006, P28 [R] pm-137.