

生物工学分野における 分子ドッキングシミュレーションの応用

Jasmina Damnjanović*・岩崎 雄吾

はじめに

分子ドッキングシミュレーションは、複合体を形成する2つの生体分子間の相互作用を予測する計算手法である。1970年代から今日まで、ドッキングシミュレーションは、生体高分子とそのリガンドの探索ツールとして医薬開発やタンパク質工学に応用されてきた。現在ではこれに関連する50以上のソフトウェア、20以上のオンラインプログラムが利用されている。これらのプログラムの開発や利用に関する論文数は増大しており、ドッキングシミュレーションの普及と重要性を如実に反映していると言える。本稿では、生物工学分野における分子ドッキングシミュレーションの原理と利用について概説し、これを利用した筆者らの最近の研究例を紹介する。

In silico ドッキングシミュレーションへの展開

分子間相互作用を適切にモデル化するには、分子が互いを認識するメカニズムを正確に理解する必要がある。分子認識メカニズムを初めて述べた鍵と鍵穴説は、酵素の基質特異性を説明するためにFisherによって1894年に提案された¹⁾。その後、1958年にKoshlandはタンパク質の柔軟性の概念を取り入れた誘導適合説を提案し、タンパク質-リガンド相互作用の理解に柔軟性を考慮することの重要性を示した²⁾。さらに1999年に提案された配座選択説は、タンパク質はリガンドと無関係に複数の立体配座間を行き来しており、その中の結合可能な配座をリガンドが選択して結合するというものである³⁾。

周知のように分子間相互作用は、X線結晶解析、NMR、電子顕微鏡および免疫沈降などにより実験的に解析することができる。しかしこれら“wet”の実験は時間と労力を要するため、*in silico*でのシミュレーションは有力な手段となる。ドッキングシミュレーションでは、タンパク質とリガンド間の結合親和性を指標にして最適な結合構造を予測する。その迅速性や簡便性から構造-機能相関や酵素反応機構解析、酵素・抗体の設計、酵素阻害剤の開発、バーチャルスクリーニングとリード化合物の最適化による創薬など、さまざまな場面で利用されている。

ドッキングシミュレーションにおける課題の1つは、分子の柔軟性である。タンパク質とリガンドは結合時には、遊離状態での安定構造から一定程度変形する。シミュレーションではこれを考慮しなければならないが、一方で計算コストの増大を招くことになる。そこで解析精度と計算コストの兼ね合いから、多くの場合にはリガンドの柔軟性は考慮するが、タンパク質構造は固定した剛体とみなすか、タンパク質の結合部位周辺のみを部分的に柔軟にした解析法が使用される。

リガンド分子は構造中に回転可能な結合が増えるに従い、取りうる立体配座の数が増加する。リガンドドッキングとはリガンドの結合部位内での位置、立体配座（形）およびリガンド自身の回転（姿勢）の最適な組合せ（ポーズ）を見つけることである。たとえば、結合部位が一辺10 Åの立方体で、リガンドが2本の回転可能な結合を持つとする。このとき位置に関して0.1 Å刻みで、回転角度に関して5°刻みでリガンドの位置と形と姿勢を少しずつ変化させ、“総当たり方式”で最安定構造を調べるとすると、その組合せは、 $(10/0.1)^3 \times (360/5)^2 \times (360/5)^3 = 1.9 \times 10^{15}$ 通りとなる。原理的にはこのすべてのポーズのエネルギーを計算して最適ポーズを見つければよいが、それには莫大な計算が必要になる。この例で、仮に1秒間に100万ポーズに対するエネルギーを計算できたとしても、完了までには約60年を要する。このような計算コストを低減するため、ドッキングプログラムには計算のための色々な工夫がなされている。

ドッキングプログラムは探索アルゴリズムとスコア関数から構成される。探索アルゴリズムはタンパク質結合部位において多数のリガンドのポーズを発生させるが、計算コストを節約して迅速に最適解に導くための色々な方法が考案されている。たとえば、AutoDockではラマルク型遺伝的アルゴリズム（LGA）により探索が行われる（他のオプションもある）。LGAではランダム発生させた初期世代ポーズ集団（デフォルトで150個）に対してエネルギー計算によるスコア付けを行い、最上位ポーズ（エリート）を次世代集団に含めるとともに、複数のポーズ間での組換え（生物での交配に相当）、一部書き換え（突然変異に相当）、一部の低位ポーズの残留

*著者紹介 名古屋大学大学院生命農学研究科応用生命科学専攻（助教） E-mail: jasmina@agr.nagoya-u.ac.jp

表1. 代表的なドッキングプログラムとその開発元

プログラム	開発元
AutoDock	The Scripps Research Institute
AutoDock Vina	The Scripps Research Institute
DOCK	Univ. of California, San Francisco
FRED	OpenEye Scientific
GOLD*	Univ. of Sheffield/GlaxoSmithKline/ Cambridge Crystallographic Data Centre
Glide*	Schrödinger
MOE-Dock*	Chemical Computing Group
FlexX*	BioSolveIT

*は商用プログラム

および局所探索によるポーズ最適化（ラマルク型遺伝）を行って次世代ポーズ集団を作り、再度スコア付けを行う。この操作を設定した世代数（デフォルトで27000世代）または設定したエネルギー計算回数（デフォルトで250万回）に達するまで繰り返し、1セットのドッキングランの結果としてベストポーズを出力する。通常はこのドッキングランを数十セット行い、各セットで得られたベストポーズについて評価・考察することになる。

スコアリング関数はポーズの結合親和性を評価するための計算式である。AutoDockでは結合の自由エネルギーをファンデルワールス力、水素結合、静電相互作用、脱溶媒和効果から算出している。実際にはスコアリング関数は計算効率を上げるためにさまざまな仮定と近似を行っている^{4,5}。表1に代表的なドッキングプログラムとその開発元を示す。

タンパク質工学における応用例

ホスホリパーゼD (PLD) は、リン脂質のホスホジエステル結合の加水分解および系内に共存するアルコールへのホスファチジル基転移反応を触媒する。放線菌由来PLDは高い転移活性を有しており、さまざまな有用リン脂質の酵素合成に利用されてきた⁶。

野生型PLDは嵩高いアルコールを基質にできないという欠点があった。筆者らのグループは、タンパク質工学による半合理的デザインにより、ホスファチジルイノシトール(PI)、1-ホスファチジル-β-グルコース(1-PGlc)やホスファチジル-L-トレオニン(P-Thr)などの嵩高い極性頭部を持つリン脂質を合成可能な改変PLDの開発を行ってきた⁷⁻⁹。一連の研究では、目的とする基質特異性を獲得した変異PLDをスクリーニングによって取得して酵素科学的な解析をするとともに、ドッキング

シミュレーションを行って変異酵素の基質認識機構を解析した。具体的な内容について以下に述べる。

PLDの構造モデルは、放線菌PLD (PDB: 2ZE4) およびその変異体 (H168A)・リン脂質複合体 (PDB: 2ZE9) の立体構造データを使用した。これらの立体構造から本酵素の反応機構に従って共有結合型反応中間体であるホスファチジル化酵素中間体のモデル構造をYASARA Structure¹⁰で作製し、エネルギー最小化を行い使用した。低分子リガンドはPubChemもしくはPDBファイルを基に作製した。ドッキングにはYASARAに実装されているAutoDockを用いて50から100セットのドッキングランを行った。各セットでのドッキングポーズを結合エネルギーで順位づけするとともに、リン脂質のリン原子、基質(リガンド)の反応点である水酸基、および触媒残基であるHis442の位置と配向に基づいて、そのポーズがプロダクティブかどうかを評価した。ここでプロダクティブなポーズとは、反応点となるリガンドの水酸基が触媒残基His442によって脱プロトン化(活性化)され、リン原子を求核攻撃できる位置に配向する、つまり、触媒反応が進行するとみなせる状態を意味する。

野生型PLDを改変した変異型酵素NYR (W187N/Y191Y/Y385R) は、ホスファチジルコリン(PC)とmyo-イノシトール(myo-Ins)からのPIの合成において、myo-Insの1位水酸基に選択性を示し、天然型PIである1-PIを優先的に生成する。一方、GYR (W187G/Y191Y/Y385R) はmyo-Insの3位に選択性を示し、非天然型PIである3-PIを優先的に生成する。この187位残基による位置選択性の相違を検証するため、NYRおよびGYRのホスファチジル化酵素中間体モデルとmyo-Insとのドッキング解析を行った。

図1A, Bは、NYRおよびGYRを用いて得られた高親和性のプロダクティブなドッキングポーズを示す。NYRでは、1位水酸基を活性化するポーズが3位を活性化するものと比較してより高い親和性を示した。GYRではその逆の結果を示し、3位を活性化するポーズは1位を活性化するものより強い親和性を示した。一連の解析により、PI合成型変異PLDの位置選択性はmyo-Insが特定の配向で結合し、特定の水酸基が選択的に活性化されるためと説明できた。また、187位の違いがmyo-Insと結合部位の他の残基との相互作用にも影響を与えることから、立体効果も関与することもわかった。

KWY (W187K/Y191W/Y385Y) 型変異PLDは、PCとグルコース(Glc)からの1-PGlc合成反応を触媒する。通常、野生型PLDはGlcの6位水酸基などの1級水酸基を好むが、アクセスしにくい1位の2級水酸基と反応で

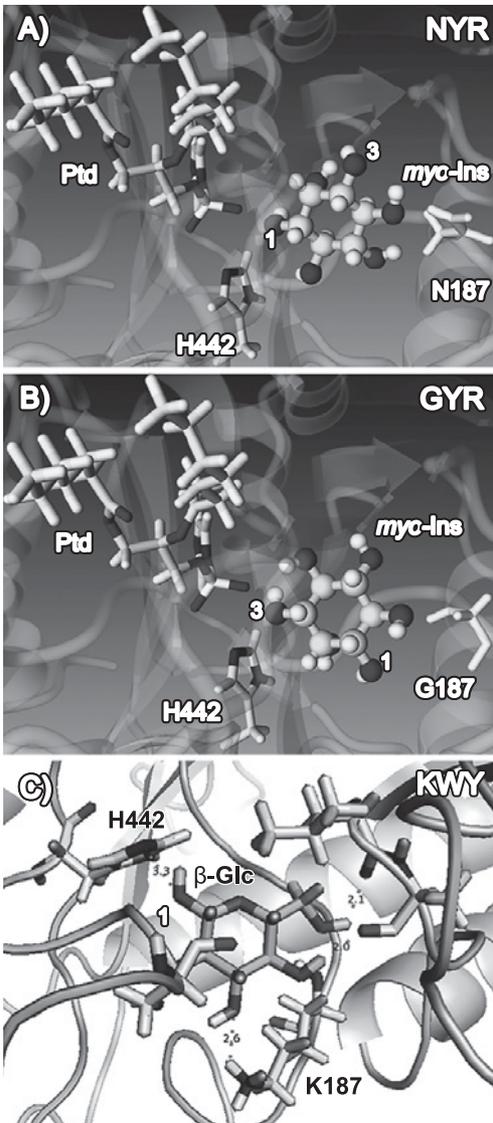


図1. A) NYR-およびB) GYR-PLDにドッキングしたmyo-Inosのプロダクティブなポーズ. C) KQY-PLDにドッキングした β -Glcのポーズ. 数字は水酸基の位置番号を示す. 文献7および8より許可を得て転載.

きるこの変異酵素は大変ユニークである. また, この酵素によって生成した1-PGlcのグリコシド結合は完全に β 配置となる. KQYがどのようにGlcの1位水酸基を認識し, かつ生成物の β -配置を維持しているか, その機構を理解するため, KQY変異体のホスファチジル化中間体モデルに α -および β -Glcをドッキングした. α -Glcの場合, 1-PGlc形成を導くプロダクティブなポーズはまったく得られなかったが, β -Glcでは10%のポーズがプロダクティブであった. この結果は, 反応系には両者が存在するにも関わらず, アキシアル (α -Glc) よりもエカトリアル (β -Glc) を認識するというKQYの特性と合致する. もっとも高い親和性を有する β -Glcとのドッキ

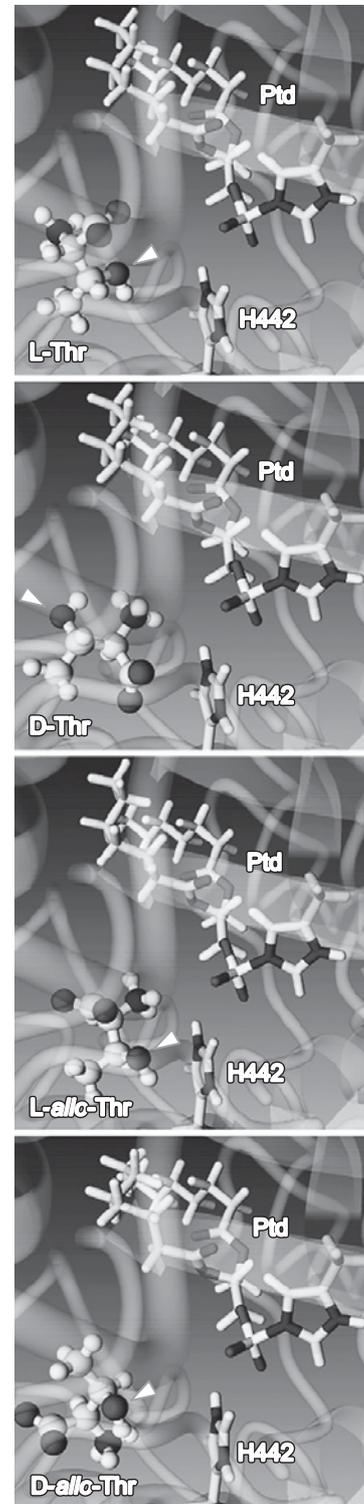


図2. L-, D-, L-allo-, およびD-allo-ThrとFYL-PLDとのドッキング. 矢印はトレオニンの水酸基を示す. 文献9より許可を得て転載.

ングポーズを図1Cに示した. 変異残基であるLys187が β -Glcの3位水酸基と相互作用し, β -Glcのプロダクティブな結合を安定化していることが示唆された.

FYL (W187F/Y191Y/Y385L) 型変異PLDは、2級水酸基をもつL-ThrとPCからP-Thrを合成することができる。2つの不斉炭素をもつThrにはL-, D-, L-*allo* およびD-*allo*体の4種の立体異性体がある。興味深いことに、FYL型PLDはこの4種の立体異性体のうち、L-, L-*allo*- およびD-*allo*- 体に作用して対応するP-Thrを与えるが、D-体にはまったく活性を示さない。

そこで、FYL型PLDのThr立体異性体に対する認識機構を解析するため、FYL変異体のホスファチジル化中間体モデルとThr立体異性体をドッキングさせた(図2)。50回のドッキング試行の結果、予想通りにL-, L-*allo*- およびD-*allo*- 体を用いた場合は、プロダクティブなドッキングポーズが高頻度で得られた一方で、D-体はプロダクティブなポーズの頻度が低かった。さらに、最安定構造にランクされたD-Thrとのドッキングポーズでは、水酸基が触媒残基His442から離れてアミノ基で遮蔽されており、水酸基の脱プロトン化が困難であることがわかった。D-Thrは6つの疎水結合と3つの水素結合により活性部位に非プロダクティブな配向で強く結合することがわかった。このようにドッキングシミュレーションからFYL型変異PLDのThr立体異性体に対する反応性の違いを説明することができた。

おわりに

以上、タンパク質工学で得た変異酵素の特性を*in silico*のドッキング解析により説明づけることができた。

このようにwetでの酵素改変実験と*in silico*での解析を結びつけることは有効で、さらに有用な酵素開発に大きな威力を発揮するだろう。

ドッキングプログラムの改良とコンピュータの高速化により、*in silico*相互作用解析は精度も速度も向上しており、今後もさらに高性能化するだろう。

酵素研究者の筆者としては、それを酵素の解析やリデザインに有効に活用していきたい。また、筆者のような計算科学を専門としない研究者も手軽に使えるインターフェイス開発にも期待したい。

文 献

- 1) Fischer, E.: *Ber. Dtsch. Chem. Ges.*, **27**, 2985 (1894).
- 2) Koshland, D. E.: *Proc. Natl. Acad. Sci. USA*, **44**, 98 (1958).
- 3) Ma, B. *et al.*: *Protein Eng. Des. Sel.*, **12**, 713 (1999).
- 4) Kitchen, D. B. *et al.*: *Nat. Rev. Drug Discov.*, **3**, 935 (2004).
- 5) Guedes, I. A. *et al.*: *Biophys. Rev.*, **6**, 75 (2013).
- 6) Damjanović, J. *et al.*: *J. Biosci. Bioeng.*, **116**, 271 (2013).
- 7) Damjanović, J. *et al.*: *Biotechnol. Bioeng.*, **113**, 62 (2016).
- 8) Inoue, A. *et al.*: *ChemistrySelect.*, **1**, 4121 (2016).
- 9) Damjanović, J. *et al.*: *Eur. J. Lipid Sci. Technol.*, **120**, 1800089 (2018).
- 10) Krieger, E. and Vriend, G.: *Bioinformatics*, **30**, 2981 (2014).