

膜タンパク質と低分子有機化合物のドッキング

広川 貴次

Receptor-ligand Docking Simulation for Membrane Proteins

Takatsugu HIROKAWA

Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology,
AIST Tokyo Waterfront Bio-IT Research Building, 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan

(Received August 18, 2006)

Protein structure-based molecular design using the computational techniques of protein structure prediction, ligand docking, and virtual screening is an integral part of drug discovery for limiting the application of the structure-based approach to target proteins such as G-protein-coupled receptors (GPCRs). GPCRs play an important role in living organisms and are of major interest to the pharmaceutical industry. However, structural data on ligand-binding forms for GPCRs from experiments to elucidate structural templates for docking simulations are lacking due to the difficulties associated with crystallization and crystallography. Therefore structural prediction of GPCRs in the ligand-bound state using computational methods has been introduced, but the prediction of ligand conformation onto target GPCRs is still constructed manually by human experts. We developed a molecular modeling technique for the prediction of ligand-receptor binding using comparative ligand-binding analysis (CoLBA) that not only considers interaction energy but also the similarity of interaction profiles among ligands. The advantage of CoLBA is that it can facilitate intuitive and flexible screening based on docking results when protein structures with low resolution (or theoretical models) are targeted. We applied CoLBA to ligand-binding prediction in several GPCRs. The predicted ligand-binding models were evaluated in site-directed mutagenesis experiments in collaborative research, and the enrichment rate of activated ligands was compared with random compounds in virtual screening simulations. We propose that CoLBA can be applied in large-scale modeling of ligand-receptor complexes and virtual screening for GPCRs.

Key words—membrane proteins; G-protein-coupled receptors; molecular modeling; docking simulation; virtual screening

1. はじめに

近年、ゲノム情報や構造ゲノミクスを活用した医薬品開発が盛んに行われている。ゲノム情報は、標的タンパク質の同定や網羅的な探索の基盤となっており、構造ゲノミクスは、創薬標的タンパク質の立体構造解析を通じて、立体構造情報に基づく創薬 (SBDD: structure-based drug design) に大きく貢献している。標的タンパク質の立体構造が実験的手法 (X線, NMR, 電子顕微鏡など) により明らかになれば、標的タンパク質の活性を制御できる化合物を活性部位の形状や相互作用を考慮しながら合理的に設計できる。現存する医薬品の代表的な標的タンパ

ク質には、キナーゼや G タンパク質共役型受容体 (GPCR)、イオンチャネルなどが知られている。^{1,2)} ここで後者の 2 つは内在性膜タンパク質に分類されるのであるが、残念ながら急増する標的タンパク質の構造解析にこれらの膜タンパク質は多く含まれていないのが現状である。

創薬標的タンパク質の代表ともいえる膜タンパク質の立体構造情報が非常に少ないことは、SBDD の大きな問題となっている。確かに GPCR やイオンチャネルを標的とする多くの医薬品は存在しているが、これらの多くはリガンド情報のみに基づく創薬技術によって開発されているものである。リガンド情報だけでは作用機序の理解が不明確であるため、選択性向上を目指した化合物設計には、やはり膜タンパク質の立体構造情報は不可欠となっている。このような背景の中、膜タンパク質の場合、計算機を用いたモデリング技術により構造情報の欠乏

産業技術総合研究所生命情報科学研究センター(〒135-0064 東京都江東区青海 2-42 産総研臨海副都心センター別館(バイオ・IT 融合研究棟))

e-mail: t-hirookawa@aist.go.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。

を補う戦略が注目されている。モデル構造は X 線構造に比べ信頼性は劣るものの、リガンド結合過程を考察し、部位特異的変異体実験を支援することや、結合エネルギーとの構造活性相関に活用できる可能性を持っているからである。実際に GPCR の 1 つである視覚ロドプシンの立体構造が解明されて以降,³⁾ いくつかの GPCR について、視覚ロドプシンの立体構造に基づいたモデリング及びリガンド・ドッキング計算、バーチャルスクリーニングが試行されるようになり、多くの先行研究が報告されてきている。しかし、モデリング技術や、さらにモデル構造が引き起こすリガンド・ドッキング計算の精度の問題など、改善すべき課題が残っている。本論文では、GPCR を対象とした膜タンパク質のモデリング法やリガンド・ドッキング計算の結果を最大限に活用する新しい手法について紹介し、実施例として行ったヒスタミン H₁ 受容体のモデリング及びバーチャルスクリーニングの結果について紹介する。

2. GPCR のモデリング

2-1. GPCR の構造 GPCR は、細胞外領域の N 端部分、7 本の膜貫通ヘリックス、膜貫通ヘリックスをつなぐ細胞外、細胞質内それぞれ 3 つのループ、そして細胞質内の C 端領域で形成される膜タンパク質である。リガンド結合部位は、膜貫通ヘリックスで形成される構造に存在することが多いが、大きな水溶性ドメインを形成する N 端領域や細胞外ループにリガンドが結合するタイプもある。一方、G タンパク質との共役は、細胞内第 3 ループや C 端領域との相互作用を介して行われることが多い。現在、立体構造既知の 7 本膜貫通型 GPCR は、視覚ロドプシンのみであり、³⁾ 視覚ロドプシンは、GPCR 分類において Class A Rhodopsin-like ファミリーに属している。よって Class A Rhodopsin-like ファミリーの GPCR の立体構造は、視覚ロドプシンとの配列比較を利用したホモロジーモデリングにより予測することが多い。本研究でも視覚ロドプシンを用いたホモロジーモデリングを適用している。

2-2. モデリングの手順 GPCR のモデリングは、主に、1) 二次構造予測、2) 配列-構造アラインメント、3) 立体構造構築の 3 つの手順で行われる。

2-2-1. 二次構造予測 二次構造予測では、与えられた GPCR 配列のどの部分が膜貫通領域にな

るかを推定するための膜貫通ヘリックス予測と、一定の立体構造を取らないフレキシブルな領域を予測するディスオーダー (disorder) 予測を行う。膜貫通ヘリックス領域は、視覚ロドプシンとの配列-構造アラインメントからも推定できるが、配列情報に基づく膜貫通ヘリックス予測を別途行い、その結果を配列-構造アラインメントの評価や修正に利用することが効果的である。膜貫通ヘリックス予測は、多くのプログラムが存在するが、プログラムの精度の比較調査⁴⁾によると、隠れマルコフモデリングに基づく TMHMM⁵⁾ が高い予測精度を示すことが分かっている。本研究でも TMHMM による膜貫通ヘリックス予測を行っている。一方、ディスオーダー領域については、分子認識や機能に重要な役割を果たすことが分かっており、配列情報に基づく予測法の開発が行われている。^{6,7)} さらに 2005 年の Jaakola らの報告⁸⁾によると、ヒトの GPCR では、N 端、C 端、細胞質内第 3 ループがディスオーダーになる傾向が強いことが知られている。本来、ディスオーダー領域は立体構造予測が難しいため、モデリングを実施する際にあらかじめディスオーダー予測でそのような領域を確認しておくことが望ましい。本研究では、ディスオーダー予測プログラムに DISOPRED⁹⁾ を利用している。

2-2-2. 配列-構造アラインメント ホモロジーモデリングでは、予測の対象となるタンパク質配列と構造既知タンパク質 (鋳型) の配列との対応付けがされた配列比較情報 (アラインメント) に基づいて立体構造が参照される。よってホモロジーモデリングでは、正確な配列-構造アラインメントは高精度のモデリングに反映される。GPCR では、各々の GPCR が異なるリガンドを受容するにも係わらず、位置特異的に保存性の高い残基や配列パターンが存在している。これらの残基は、膜貫通ヘリックス間の相互作用や G タンパク質との共役に重要な役割を果たしていることが多い。

Figure 1 は、視覚ロドプシン構造における第 1, 2, 7 膜貫通ヘリックス間の相互作用に関与する残基とアミン系リガンドを受容する GPCR のいくつかの代表配列間での保存性との対応を示したものである。第 1 膜貫通ヘリックスのアスパラギン残基 (視覚ロドプシンでは、55 番目)、第 2 膜貫通ヘリックスのアスパラギン酸残基 (視覚ロドプシンでは、83

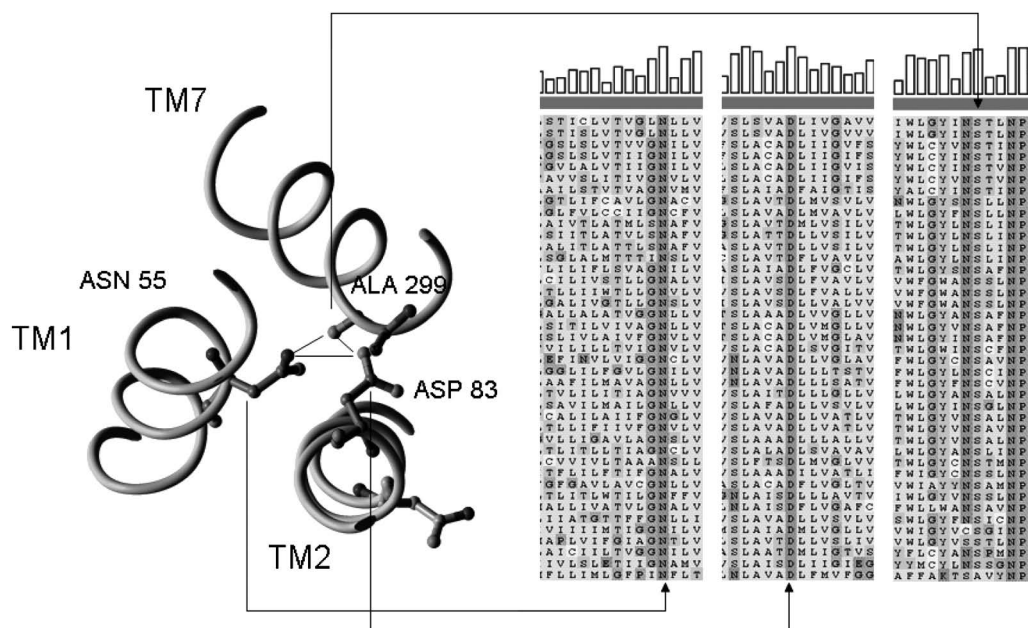


Fig. 1. Conserved Key Residues in Three Transmembrane Helices of the Class A Rhodopsin-like GPCR Family (Right) and Structural Mapping to the Rhodopsin Structure (Left)

番目), 第7膜貫通ヘリックスのアラニン残基(視覚ロドプシンでは, 299番目)に水素結合ネットワークが存在しており, これらの残基は, ファミリーの中でもよく保存されていることが分かる. ほかにファミリー間で保存性の高い膜貫通ヘリックス間水素結合ネットワークが存在しており, これらの残基位置を確認しながら配列-構造アラインメントを構築することがポイントとなる.

膜貫通ヘリックス間の相互作用以外にも第3膜貫通ヘリックスの [D/E]R[Y/H] モチーフや第7膜貫通ヘリックスの [N/D]PxxY モチーフなどが重要な配列モチーフとして知られている.¹⁰⁾ 本研究では, 予測対象の GPCR が Class A Rhodopsin-like ファミリーに属する場合, 視覚ロドプシンとの配列-構造アラインメントを行う際に, Class A Rhodopsin-like ファミリーから選択した40の代表配列を GPCRDB¹¹⁾ より取得し, 多重アラインメント(合計42配列)を実施することで, 上記で指摘した配列モチーフの寄与を配列-構造アラインメントに反映させている. 多重アラインメントには, ClustalW¹²⁾ を利用している.

2-2-3. 立体構造構築 立体構造構築では, アラインメントされている配列について, 鋳型タンパク質から主鎖の原子座標が参照される. 続いてアラインメントされている配列でもアミノ酸残基の種類

が一致している部位は, 側鎖の原子座標も主鎖構造と同じく参照される. 変異が起きているアミノ酸残基の側鎖構造の配座については, 側鎖の二面角ライブラリと呼ばれる統計的手法に基づく二面角情報を用いて与えられる. その際, 近隣のアミノ酸残基との立体障害や相互作用エネルギーも考慮されながら最適側鎖構造が構築される. 大抵の立体構造構築プログラムは, 上記のプロトコルに従っているが, 二面角ライブラリや最適化のアルゴリズムの違いでそれぞれ特徴を出している. 本研究では, InsightII/Homology¹³⁾ の機能を用いて初期構造を構築した. この初期構造に対して, 分子動力学計算による焼きなまし法を適用し, 側鎖構造のサンプリングを行った. 計算条件は, 主鎖構造を固定する拘束を与え900 Kから300 Kへの温度計画を行い, 100の候補構造を出力した. 分子動力学計算プログラムには AMBER⁸⁾を用いた. 続いて候補構造に対し, Dastmalchi らによって考案された Lipid Compatibility Score (LCS)¹⁵⁾を計算し, LCSが最大となる構造を最終構造とした. LCSとは, 脂質環境や埋没表面積, 極性環境比率を基に算出されるスコアで, 水溶性タンパク質用評価プログラム Verify3D¹⁶⁾を膜タンパク質評価用に改良されたものである.

2-3. リガンド・ドッキング計算 リガンド・

ドッキング計算は、1) タンパク質の活性ポケット内でのリガンド・ドッキングと、2) ドッキング状態の相互作用エネルギー（若しくはスコア）評価の2つのプロセスからなる。前者は、リガンド自身の持つ配座の自由度と活性ポケットに対する相対位置の自由度を考慮して、無駄なく網羅的にドッキングポーズを探索することが要求される。また後者においては、量子化学計算に基づく精密な評価から分子力学計算、経験ポテンシャルによるスコア化まで、精度と計算処理時間に依っていくつかの方法が存在する。リガンド・ドッキング計算の成功は、これら2つのプロセスの精度に依存するが、ドッキング計算プログラムに係わらず、最も影響を与える大きな要因は、ドッキングの対象となるタンパク質立体構造そのものの解像度である。通常、リガンド・ドッキング計算プログラムの性能評価では、高解像度のX線結晶解析で得られたリガンドを含むタンパク質立体構造データによる再挿入シミュレーション（リガンドを一度取り除き、ドッキング計算プログラムで再びリガンドを挿入させ、結晶構造と比較する）が行われている。タンパク質の立体構造が高解像度のものでなく、ホモロジーモデリングによって得られたモデリング構造の場合には、モデリング構造の精度にもよるが、ドッキング計算プログラムだけでは正解とされるドッキングポーズを推定することは再挿入シミュレーションからも困難とされている。これはモデリング構造を多く扱う膜タンパク質のドッキング計算において非常に深刻な問題で、何らかの処方箋が望まれている。一般にGPCRのモデリング構造に基づくSBDDでは、変異体による実験データを知識として利用した拘束付ドッキングやマニュアルドッキングが広く行われている。^{17,18)} この人的介入により予測構造でも成功率の高いSBDDが実現できることもあるが、特定の熟練者への依存度、多大な時間を要するリガンド・ドッキングモデル構築、実験データが必須とされる点などから新規のGPCRなどには、適用が難しい。

しかし、この問題をリガンド・ドッキング計算のプロセス別で考えると、リガンド・ドッキング自身は、受容体の側鎖構造の自由度を加味すればかなりのドッキングポーズを探索できていることが分かる。つまり、相互作用エネルギーの評価が主な原因でリガンド・ドッキング計算がうまくいっていない

のである。そこで筆者は、相互作用エネルギーの評価にコンセンサス法を導入した新しい評価基準を開発し、柔軟でかつ直感的なリガンド・ドッキングポーズの選択を実現した。本手法は、同一の標的タンパク質に対して同じ作用機序を持つ複数のリガンドのドッキング計算を行い、サンプリングされた候補構造間を比較しながら相互作用の共通性を最大に説明できるドッキングポーズを決定するものである。筆者らはこれを、Comparative Ligand Binding Analysis (CoLBA法)と呼んでいる。

2-4. CoLBA法 CoLBA法のプロトコルは次の通りとなる。1) 標的タンパク質に対し、同一の作用機序（作用・拮抗）が想定される複数のリガンドを既存のドッキング計算アルゴリズム等で個別にドッキング、2) 各リガンドについてドッキングスコア順に上位複数の候補ポーズを保存、3) 候補ポーズをリガンド間で交差比較し相互作用類似尺度及び既に得られているドッキングスコアをエネルギー順位スコアに変換した平均値を計算、4) 相互作用類似尺度とエネルギー順位スコアからなる散布図解析により統計的に有意と表現される候補ポーズの組み合わせを選択し、各リガンドの標的タンパク質に対する結合モデルとする。

2-4-1. リガンド・ドッキング計算 リガンド結合部位におけるリガンド結合のポーズ候補の探索をMOE¹⁹⁾/AS-DOCK²⁰⁾を用いて行った。AS-DOCKでは、リガンドの配座の自由度とリガンド結合部位周辺残基の側鎖構造の自由度も加味している。リガンド結合部位は、MOE/Site-Finderによって偽原子で記述されており、リガンドの配座は、この偽原子との重なりを考慮しながら効率的に行われる。配座探索は、2000回実施され、配座が決定されたのち、周辺側鎖の自由度を考慮にいたった構造最適化計算が行われる。最適化計算の力場には、MMFF94²¹⁾を利用した。最適化計算を終えたそれぞれのリガンド結合ポーズは、リガンド-タンパク質間のファン・デル・ワールス相互作用と静電相互作用、リガンドの歪エネルギーに基づく相互作用エネルギーで順位付けされる。

2-4-2. 相互作用類似尺度 2つのリガンドA、Bについてドッキング計算から得られたそれぞれ任意の候補ポーズ*a*、*b*間の相互作用類似尺度*S_{ab}*は、リガンド-タンパク質間の水素結合数*HB*と

疎水性接触数 HC の類似性 $S_{ab}^{HB, HC}$ により計算される。

$$S_{ab}^{HB, HC} = \frac{\sum_{i=1}^L P(i)_a^{HB, HC} P(i)_b^{HB, HC}}{\sum_{i=1}^L \{P(i)_a^{HB, HC}\}^2 + \sum_{i=1}^L \{P(i)_b^{HB, HC}\}^2 - \sum_{i=1}^L P(i)_a^{HB, HC} P(i)_b^{HB, HC}} \quad (1)$$

ここで、 $S_{ab}^{HB, HC}$ は、Tanimoto 係数で表現される相互作用類似性スコアで、水素結合数類似性 S_{ab}^{HB} と疎水性接触数類似性 S_{ab}^{HC} から構成される。 $P_{a,b}^{HB, HC}$ は、水素結合数や疎水性接触数をそれに関与するタンパク質残基の L 次元のベクトルとして表現 (Fig. 2) されるものである。 L は、タンパク質残基長となる。水素結合数及び疎水性接触数は、LIGPLOT プログラム²²⁾により同定した。

最終的に相互作用類似尺度 S_{ab} は S_{ab}^{HB} 、 S_{ab}^{HC} に重み w と最大水素結合数で規格化した重み C を考慮した、次式で計算される。

$$S_{ab} = w \cdot C \cdot S_{ab}^{HBond} + (1-w) S_{ab}^{HC} \quad (2)$$

$$C = \frac{a_{ab}}{N_{A,B}^{max}} \quad (3)$$

$N_{A,B}^{max}$ は、リガンド A、B のすべての候補ポーズの中に存在する水素結合数の最大値で、 a_{ab} は、候補ポーズ a, b で共通に存在する水素結合数である。

2-4-3. エネルギー順位スコア 比較するリガンド A、B は、それぞれ分子量や原子組成が異なることから、ドッキング計算スコアをそのまま利用せず、それぞれのリガンドについてドッキング計算で 1 位のドッキングスコアからの差を任意の候補ポーズ a, b の差分スコアとして定義した (1 位の候補ポーズは、0 となる)。

$$\Delta E_{a,b} = E_{A,B}^{lowest} - E_{a,b} \quad (4)$$

そしてリガンド A、B の任意の候補ポーズ a, b 間のエネルギー順位スコア ES_{ab} は、それぞれの差分スコアの平均で評価した。

$$ES_{ab} = \frac{|\Delta E_a - \Delta E_b|}{2} \quad (5)$$

2-4-4. 散布図解析 リガンド A、B の各候補ポーズをリガンド間で総当たりの交差比較し、 S_{ab} 及び ES_{ab} の値を計算後、それぞれ標準偏差で規格化し散布図を作成する。各々の候補ポーズペアの特

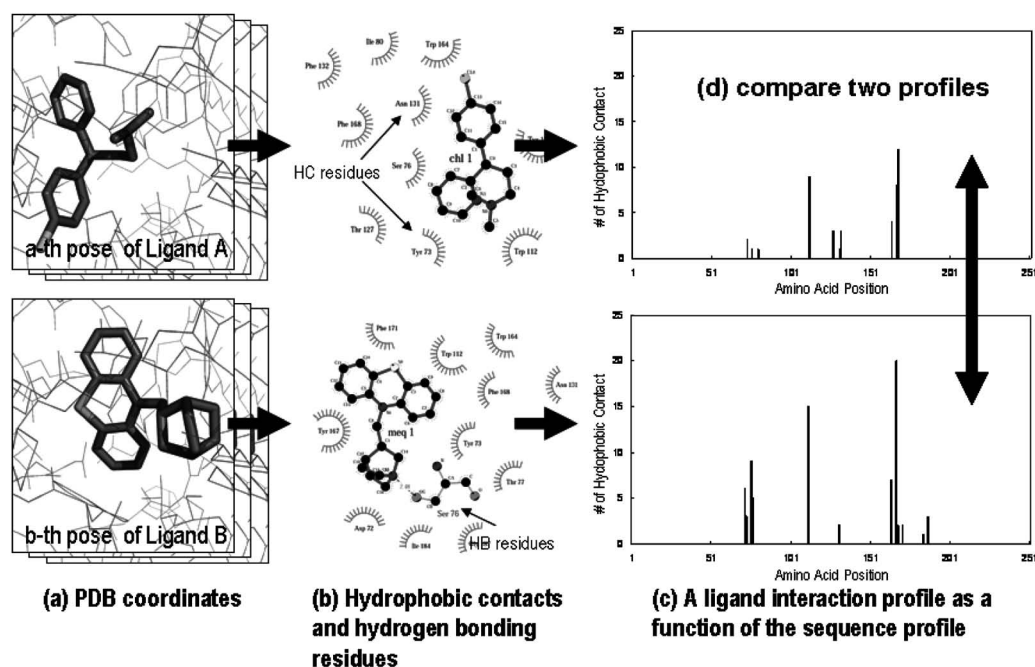


Fig. 2. Characterization of Ligand-binding Profile in CoLBA

(a) atomic coordinates of the candidate poses for each ligand derived from the docking program, (b) enumeration of the hydrophobic contacts and hydrogen bonds between the receptor and ligand using the LIGPLOT program, (c) ligand interaction profile by plotting the number of contacts as a function of the receptor-sequence profile, and (d) comparison of the two profiles.

性は、散布図の原点を通る直線 $ES_{ab}=S_{ab}$ からの垂直線の距離 (Z-score) で評価する。

$$\text{Z-score} = \text{sign}(S_{ab}, ES_{ab}) \times \sqrt{\left(\frac{S_{ab}-ES_{ab}}{2}\right)^2 + \left(\frac{ES_{ab}-S_{ab}}{2}\right)^2} \quad (6)$$

$$\text{sign}(S_{ab}, ES_{ab}) = \begin{cases} -1: ES_{ab} > S_{ab} \\ 1: \text{otherwise} \end{cases} \quad (7)$$

分布から大きく右下に位置するプロット (ドッキングスコアでも上位で、かつ、リガンド間でタンパク質に対する相互作用様式が似ている) が理想的な候補ポーズペアとなる。

3. 結果と考察

本研究では、ヒスタミン H_1 受容体 (ヒト) とリガンドの結合を実施例として、受容体構造のモデリング、CoLBA 法を用いたリガンド結合予測を行った。ヒスタミン H_1 受容体は、末梢神経系や中枢神経系に発現し、炎症やアレルギー反応の調節に関与している GPCR で、抗ヒスタミン剤開発の標的タンパク質である。ヒスタミン H_1 受容体の拮抗薬には、クロルフェニラミンに代表される第一世代とエピナスチンやメキタジンに代表される第二世代とに分類される。第一世代型拮抗薬は、中枢神経系に存在するヒスタミン H_1 受容体にも働きかけることから、眠気などを引き起こす副作用を有している。これに対して第二世代型拮抗薬は、中枢神経系への働きかけが少ない拮抗薬であるが、時折、hERG (human ether-a-go-go related gene) の阻害機能を有するものがあり、副作用として心電図における QT 延長を引き起こしてしまう。ヒスタミン H_1 受容体と拮抗薬結合のメカニズムをモデリングにより明らかにすることは、副作用を改善した第三世代型のヒスタミン H_1 受容体拮抗薬の開発においても重要である。よって本研究では、リガンド結合モデルを利用して、試薬ライブラリからのバーチャルスクリーニングも試みた。また、バーチャルスクリーニングで選ばれた化合物は、実際に結合アッセイ試験により結合活性を評価した。

3-1. ヒスタミン H_1 受容体構造モデリング
ヒスタミン H_1 受容体の立体構造構築部分は、リガンド結合部位が存在する膜貫通ヘリックス領域のみ行った。特にヒスタミン H_1 受容体は、細胞質内第 3 ループに大きなドメインを持つが、ディスオー

ダー予測によりこのループがディスオーダーである可能性が示唆されたため、構造構築の対象から除外した。構造評価の指標となる LCS は、初期構造では、0.81 であったが、分子動力学計算による側鎖構造の極小化により理想値 1.0 に近い 0.97 へ向上した。

3-2. CoLBA 法によるリガンド結合予測
ヒスタミン H_1 受容体の構造予測を行ったのち、クロルフェニラミン、エピナスチン及びメキタジンのドッキング計算を行った結果、それぞれ 25、25 及び 19 の候補構造が得られた (Fig. 3(a))。これらの候補構造について、11875 通りの組み合わせで CoLBA 法による解析を行った。3 種のリガンドの比較では、相互作用類似尺度、エネルギー順位スコアは、3 通りのペア比較の平均により算出した。Figure 3 (b) は、結果を散布図により表現したものである。この散布図から分かるように、分布から右下に離れたところにある組み合わせが存在していることが分かる。これは、クロルフェニラミン (1 位)、エピナスチン (2 位)、メキタジン (4 位) の候補ポーズの組み合わせによるものである。Z-score は、8.41 と有意なものであることが確認できた。Figure 3 (c) は、CoLBA 法で選択された 3 つのリガンドの候補ポーズの重ね合わせである。各リガンドには、カチオン性窒素が含まれているが、それが 3 番目の膜貫通ヘリックス (TM3) に存在する 107 番目のアスパラギン酸に共通して結合し、疎水性の官能基は、TM5—6 に対してほぼ同一の空間に存在していることが分かる。このようなファルマコフォアと呼ばれる機能に関する官能基の空間的特性から、ドッキング計算のみで決定される最安定構造よりも CoLBA 法により選定された候補ポーズが作用機序を説明する上でも適切なことが分かる。ヒスタミン H_1 受容体について、このファルマコフォアが情報伝達を拮抗するモデルとして理想的であることが、アミノ酸残基の変異実験等からも確認されている。²³⁾

3-3. 濃縮係数評価
CoLBA 法によって同定されたリガンド結合モデルを評価するために、ランダムに選んだ化合物との結合比較による濃縮係数評価を行った。約 280 万件の試薬ライブラリ (ナミキ商事カタログライブラリ) から Lipinski ルール²⁴⁾ を満たす化合物をランダムに 990 個選択し、これを

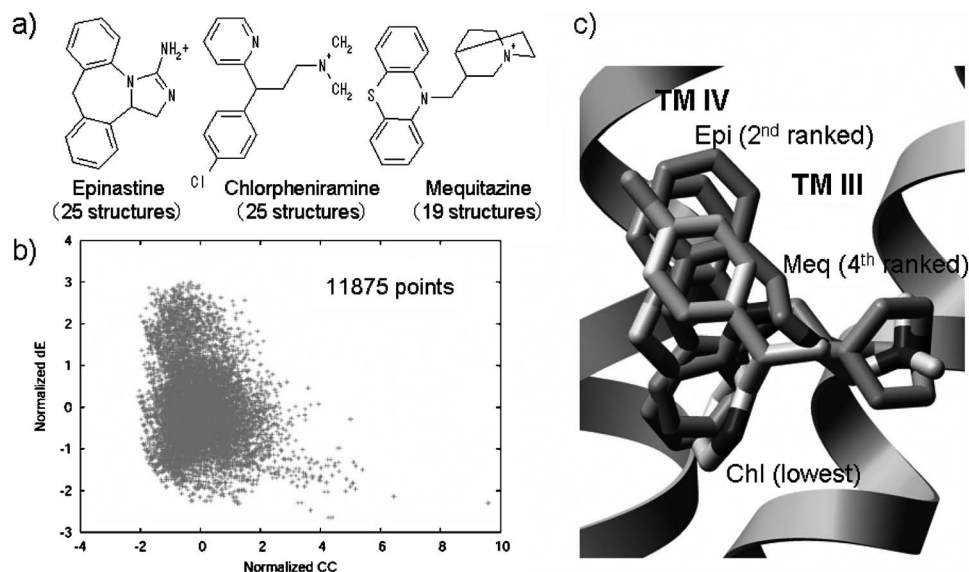


Fig. 3. Results Produced Using CoLBA for Three Histamine H_1 -receptor Antagonists

(a) chemical structures of epinastine, chlorpheniramine, and mequitazine, (b) dispersion diagram in CoLBA, and (c) proposed binding mode for three antagonists determined using CoLBA.

ヒスタミン H_1 受容体リガンドと比較するランダム化合物セットとした。ヒスタミン H_1 受容体のリガンドには、クロルフェニラミン、メキタジン、エピナスチンを除いた 10 個の拮抗薬を用意した。化合物の三次構造構築、極小化には MOE を用いて行った。990 個のランダム化合物及び 10 個の拮抗薬、合計 1000 個の化合物を CoLBA 法で同定したメキタジン結合状態のヒスタミン H_1 受容体に対してドッキング計算を行った。ドッキング計算には、GLIDE (SP-mode)²⁵⁾を利用した。

Figure 4 は、X 軸に順位、Y 軸に任意の順位内に存在する拮抗薬の数をプロットしたグラフである。破線は、拮抗薬がランダムに存在する場合を示したものである。よって、グラフ左上に近接するように放物線状にプロットされる場合は、拮抗薬がランダム化合物からよく識別されていることを意味する。■線は CoLBA 法によって同定されたメキタジンの結合モデル（相互作用エネルギーのみで評価した場合は、第 4 位の候補）、▲線は、相互作用エネルギーのみで同定したメキタジンの結合モデルを用いて 1000 個の化合物をドッキングした結果を示している。このグラフから、CoLBA 法によって同定した結合モデルが、既存の拮抗薬とランダム化合物をよく識別しており、バーチャルスクリーニングが効率よく実施できるモデルであることが分かる。濃縮係数は、ランダムな識別との比率で計算されるが、

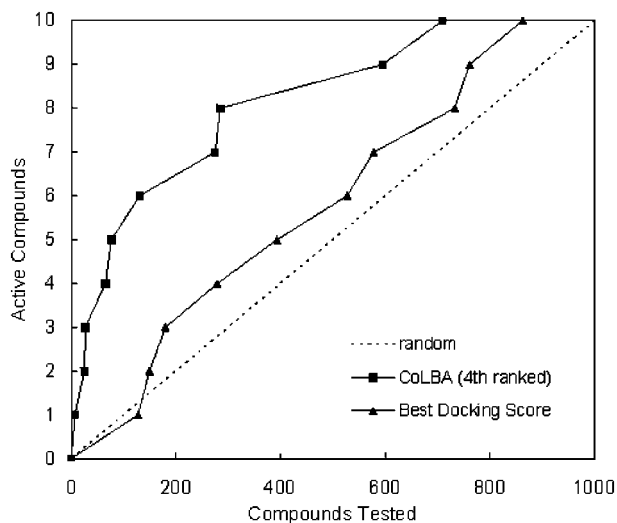


Fig. 4. Results of an Enrichment Graph Illustrating How the Ligand-binding Model Determined Using CoLBA Can Be Useful for Virtual Screening More Effectively than Using the Best Docking Score and Random Selection

今回の結果では、50%の拮抗薬を識別する効率（濃縮係数）は、6.5(倍)であった。

3-4. バーチャルスクリーニングの実施と評価
ランダムな化合物を利用した濃縮係数による結果から、本研究で予測したヒスタミン H_1 受容体とリガンド結合モデルが妥当であることが評価された。よって、本研究では、さらに大規模な化合物データベースを利用してバーチャルスクリーニングを実施した。スクリーニング化合物は、ランダム化合物を

選択したときと同様の試薬ライブラリを基本とし、化合物の重複や Lipinski ルールによるフィルタリングより、約 140 万件に絞り込んだものを用いた。バーチャルスクリーニングは、GLIDE (SP-mode) によるドッキング計算で行い、Glide Score によって順位付けを行った。そして上位 140 化合物を購入し、結合アッセイ試験による評価を行った。

結合アッセイ試験は、PerkinElmer 社の協力により行った。CHO 細胞に発現したヒトヒスタミン H_1 受容体に対して、140 化合物の阻害率をラジオリガンド及びリファレンスリガンドを用いた測定法で評価した。測定の濃度は、 $10 \mu\text{M}$ の単一濃度で、2 回測定を行った。Figure 5 は、140 化合物の阻害率 (2 回測定の平均) の棒グラフである。グラフが示すように、140 化合物中、60%以上の阻害率を示すものを 14 化合物見つけ出すことができ、一次スクリーニングとしては良好な結果を得ることができた。このことから、本研究で予測したヒスタミン H_1 受容体構造及びリガンド結合モデルが妥当であることが示唆された。

4. まとめ

CoLBA 法による複数のリガンド・ドッキング計算のコンセンサス解析は、非常に単純なプロトコルであるが、創薬現場の専門家からみて直感的によい結果が得られ易い (ここでは紹介できないが、ほか

の実施例でも良好な結果が得られている)。ドッキング計算結果をコンセンサスにより選び出すアプローチ自体は、既に創薬現場で行われているが、多くは、候補ポーズ間のリガンド構造に着目した類似尺度 (リガンド構造のアラインメントスコア) でコンセンサス評価を行っている。この場合、結果がリガンドの構造類似性の信頼性に依存することになり、一見、リガンド間の構造骨格や官能基組成に共通性がない場合は困難となる。これに対し、CoLBA 法は、候補ポーズ間の類似性をリガンド自身の構造特性 (配置・配向) ではなく、相互作用情報のプロファイルとしてタンパク質配列プロファイルに変換することで、情報科学的処理による比較計算を容易にし、リガンド同士の類似性に依存しない比較法を実現した。

もう 1 つの長所としては、今回のように標的タンパク質が実測された構造ではなく解像度の低下が伴う計算機モデリングでも適用範囲が広がる点が挙げられる。一般のドッキング計算の結合エネルギー計算は原子座標にとっても敏感で、タンパク質側鎖の向きが少しでも変わっただけで数 10 kcal/mol の変動が生じる。CoLBA 法では、ドッキング計算の結合エネルギーだけではなく、水素結合数や疎水性接触数に基づいた相互作用類似尺度を評価軸に取り入れたことで、標的タンパク質構造精度に対して柔軟な

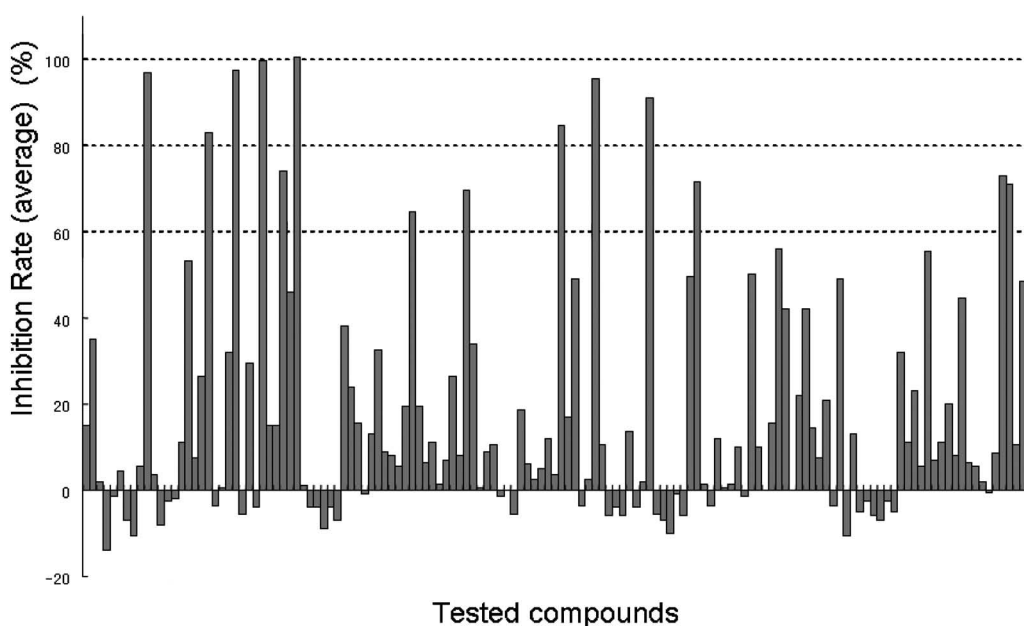


Fig. 5. Average Inhibition Rate (%) of 140 Screened Compounds at One Point Concentration ($10 \mu\text{M}$) in Duplicate

手法になっていると考察できる。

5. おわりに

ヒスタミン H₁ 受容体を始め、現在、構造決定が困難とされている GPCR などの標的タンパク質についても、将来的には解像度の高い結晶構造が得られ、さらにドッキング計算技術そのものも進化すれば、CoLBA 法のような補正技術は不必要になるであろう。CoLBA 法は、そのような時代が到来するまでの橋渡しの役割に過ぎないのかもしれないが、一刻でも早く新薬創製が望まれる創薬分野において、本手法が創薬現場の専門家の人的介入に代わる技術の 1 つとなり、SBDD を加速することにつながれば幸いだと思っている。

謝辞 ヒスタミン H₁ 受容体予測構造及びリガンド結合モデルについて貴重な助言を賜った徳島大学薬学部福井裕行教授に心より感謝申し上げます。

REFERENCES

- Drews J., *Science*, **287**, 1960–1964 (2000).
- Hopkins A. L., Groom C. R., *Nat. Rev. Drug Discov.*, **1**, 727–730 (2002).
- Palczewski K., Kumasaka T., Hori T., Behnke C. A., Motoshima H., Fox B. A., Le Trong I., Teller D. C., Okada T., Stenkamp R. E., Yamamoto M., Miyano M., *Science*, **289**, 739–745 (2000).
- Moller S., Croning M. D., Apweiler R., *Bioinformatics*, **17**, 646–653 (2001).
- Krogh A., Larsson B., von Heijne G., Sonnhammer E. L., *J. Mol. Biol.*, **305**, 567–580 (2001).
- Romero P., Obradovic Z., Li X., Garner E. C., Brown C. J., Dunker A. K., *Proteins*, **42**, 38–48 (2001).
- Ward J. J., Sodhi J. S., McGuffin L. J., Buxton B. F., Jones D. T., *J. Mol. Biol.*, **337**, 635–645 (2004).
- Jaakola V. P., Prilusky J., Sussman J. L., Goldman A., *Protein Eng. Des. Sel.*, **18**, 103–110 (2005).
- Ward J. J., McGuffin L. J., Bryson K., Buxton B. F., Jones D. T., *Bioinformatics*, **20**, 2138–2139 (2004).
- Bissantz C., Logean A., Rognan D., *J. Chem. Inf. Comput. Sci.*, **44**, 1162–1176 (2004).
- Horn F., Bettler E., Oliveira L., Campagne F., Cohen F. E., Vriend G., *Nucleic Acids Res.*, **31**, 294–297 (2003).
- Thompson J. D., Higgins D. G., Gibson T. J., *Nucleic Acids Res.*, **22**, 4673–4680 (1994).
- <http://www.accelrys.com/jp/>
- Case D. A., Cheatham 3rd T. E., Darden T., Gohlke H., Luo R., Merz Jr. K. M., Onufriev A., Simmerling C., Wang B., Woods R. J., *J. Comput. Chem.*, **26**, 1668–1688 (2005).
- Dastmalchi S., Morris M. B., Church W. B., *Protein Sci.*, **10**, 1529–1538 (2001).
- Eisenberg D., Luthy R., Bowie J. U., *Methods Enzymol.*, **277**, 396–404 (1997).
- Bissantz C., Bernard P., Hibert M., Rognan D., *Proteins*, **50**, 5–25 (2003).
- Shacham S., Marantz Y., Bar-Haim S., Kalid O., Warshaviak D., Avisar N., Inbal B., Heifetz A., Fichman M., Topf M., Naor Z., Noiman S., Becker O. M., *Proteins*, **57**, 51–86 (2004).
- <http://www.chemcomp.com/>
- Kataoka R., Goto J., “Posutoshihkuensu Tanpakushitsu Zikkenn-hou 1 (in Japanese),” Tokyo Kagaku Dojin, Tokyo, 2002, pp. 191–201.
- Halgren T. A., *J. Comput. Chem.*, **17**, 490–512 (1996).
- Wallace A. C., Laskowski R. A., Thornton J. M., *Protein Eng.*, **8**, 127–134 (1995).
- Bruysters M., Pertz H. H., Teunissen A., Bakker R. A., Gillard M., Chatelain P., Schunack W., Timmerman H., Leurs R., *Eur. J. Pharmacol.*, **487**, 55–63 (2004).
- Lipinski C. A., Lombardo F., Dominy B. W., Feeney P. J., *Adv. Drug Deliv. Rev.*, **46**, 3–26 (2001).
- <http://www.schrodinger.com/>