

【研究の目的】

分子の立体構造に基づく薬物設計を成功裡に行うためには、標的蛋白質の厳密な立体構造情報に加え、候補化合物とのドッキング計算で精度の高い相互作用エネルギー値を求める必要がある。しかし、現実にはそれらの要求を完全に満たすことは難しく、従来は専門家が自身の経験に基づいて計算結果を吟味し有望な化合物を選択していた。そこでの「知識」を明示的に示すことは難しく、これまで実現されていない。本研究では、対話型進化計算法により人間と計算機がやり取りを繰り返す過程で、人間の経験に基づく主観や直感という漠とした知識を計算機が学習する量子生命科学シミュレータの開発を提案する。これにより人間の暗黙知を共有し、人と計算機が協調した薬物設計を可能とする。

【研究の内容、成果】

1. 標的タンパク質の設定

分子の立体構造に基づく薬物設計（SBDD: Structure-Based Drug Design）は、標的となる受容体タンパク質の立体構造や相互作用情報に基づいて医薬品設計を理論的に行うもので、有望な候補化合物（リード）を探索、選定し、最適化する工程で用いられる手法である。SBDDを成功裡に行うためには、用いるタンパク質の立体構造の精度が非常に高いことが必要である。本研究ではGタンパク質共役型受容体（GPCR: G Protein-Coupled Receptors）を取り上げ、SBDDに耐えうる程度に正確な立体構造を予測することを目標とする。

GPCRは、細胞膜上に存在し、神経伝達物質やホルモンを認識する生体センサーとしての役割を持つ。GPCRの種類は多岐にわたるが、構造の点で細胞膜貫通ヘリックスを7本持つという共通点を有し、タンパク質中で最大のスーパーファミリーを形成している。現在臨床的に用いられている薬の約半分がGPCRに作用するものと言われており、創薬の大きなターゲットである。

そこで、SBDDが期待されるわけだが、GPCRは膜タンパク質のため結晶構造を得るのが非常に難しく、これまでに実験的に結晶構造が得られているのは、ウシロドプシンだけである。ロドプシン（Rhodopsin）は、目の網膜に存在する物質で、これまで実験的に結晶構造が得られているのは光を吸収する前の「不活性型」だけである。そのため、これを鋳型にして予測した他のGPCRの立体構造は、やはり不活性型になってしまい、活性型と結合するといわれているアゴニスト（作動薬）の探索に用いることはできない。本研究では、これらの問題を解決する手法を提案する。

2. 提案する手法

本研究では、扱う問題を活性型GPCRの立体構造予測問題として定式化し、次に説明するタンパク質の誘導適合（induced-fit）を考慮したモデルを導入する。

タンパク質の高次構造は、水素結合、静電的相互作用、疎水性相互作用など極めて微小な安定化力で維持されており、外界からの熱エネルギー

ギーで常に揺らいでいる。溶液中でのタンパク質は、揺らぎのためにわずかに異なる多くの立体構造の間を移り変わっており、その中で結合を許す構造にのみリガンドは結合し、より安定したエネルギー値をとるように複合体構造が変化する。このように、まるでリガンドがタンパク質の構造を結合に適したものに变えたようにみえる現象を誘導適合という。誘導適合を考慮したタンパク質・リガンドの結合シミュレーションシステムは、その難しさのためにほとんど存在しないのが現状である。我々は、誘導適合を考慮したSBDD実現のために、時間スケールの異なる揺らぎの組合せでGPCRの揺らぎをモデル化することにした。なお、モデル化にあたり、GPCRは脂質二重膜に囲まれているため他の球状タンパク質ほど自由には動けないが、結晶中の環境ほど拘束されてはいないという点と、GPCRを構成するヘリックスの溶液中での軸周りの回転運動がこれまでの研究で示唆されているという点を考慮した。そして、誘導適合による構造変化を結合シミュレーション単体の中で解決するのではなく、他の手法と組み合わせた進化計算法を用いることで解決するアプローチをとることとした。具体的には、以下のフレームワークを提案する。

- (1) ウシロドプシンを鋳型として、ホモロジーモデリングで標的GPCRの初期立体構造を構築する。
- (2) 数十～百ps程度までのGPCR分子の揺らぎを分子動力学 (MD: Molecular Dynamics) でシミュレーション計算し、平均的な構造を求める。
- (3) より時間スケールの大きい揺らぎとして、GPCRを構成する各ヘリックスの軸を中心とした回転運動を仮定する。既に結合することが自明であるアゴニストを訓練例として、結合シミュレーションと実数値遺伝

的アルゴリズムを組み合わせた探索を行うことで、アゴニストの結合に適したヘリックスの回転位置を求める。

この中で特に注意を要する点は、(3)の探索で用いる評価関数である。通常のタンパク質とリガンドの結合シミュレーションでは、様々な評価関数が提案されているが、全ての場合において結合自由エネルギーを精度良く反映するようなスコアは現在までにわかっていない。そこで、本研究では客観的なスコアに加え、専門家の陽にできない知識を取り込むために、探索のサイクルごとに結合の様子をビューワーソフトウェアで提示しユーザーに定性的な判断を下してもらうようにした。その際、「どの程度良いか」という判断は難しいので、「これは良くない」というもののみ選び出し、負の評価を与えるようにした。このように対話的に人間の評価基準を取り入れるので、これは対話型進化計算法のひとつといえる。

3. 用いている要素技術

提案手法中で用いている主要要素技術には、分子動力学シミュレーション、実数値遺伝的アルゴリズム、ドッキングシミュレーションがある。それらを簡単に説明する。

■ 分子動力学シミュレーション

本研究では“分子とは部分電荷を持った原子がバネでつながったもの”とみなして計算を行う古典MDの立場を用いる。個々の原子の運動を、以下のニュートンの運動方程式を解くことで求める。なお、本研究で用いたMDソフトウェアは、TINKERである。

$$m_i \frac{d^2}{dt^2} \mathbf{r}_i = -\nabla_i U(\mathbf{r}_1 \cdots \mathbf{r}_n), \quad i = 1, \dots, n$$

$m_i, \mathbf{r}_i$ は粒子*i*の質量と位置

U は系全体のポテンシャルエネルギー

■ 実数値遺伝的アルゴリズム

遺伝的アルゴリズム (GA: Genetic Algorithms) は生物の適応進化に関するダーウィンの自然選択説を計算機上で模擬する最適化手法である。GAについては離散値空間を探索領域とする組合せ最適化問題への応用がよく知られているが、連続値空間を探索領域とする問題には、実数値GAという手法が提案されている。本研究では、実数値GAのうち単峰性正規分布交叉 (UNDX: Unimodal Normal Distribution Crossover) を用いている。

■ ドッキングシミュレーション

標的分子と低分子化合物 (リガンド) の結合について、以下の式で与えられる結合自由エネルギー変化 (ΔG_{bind}) のできるだけ小さい構造を探索する。

$$\Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand})$$

探索の評価関数としてのスコアには様々なものが提案されている。本研究では、統計的スコア関数法 (既知の結晶構造を統計的に処理して、各官能基のペアの相対的な配置の確率分布を求めてスコア化する方法) を用いているGOLDというソフトウェアを用いる。

4. ロイコトリエン受容体を用いた実行例

実際の適用例として、様々なGPCRのなかからロイコトリエン (LT: Leukotriene) 受容体を取り上げて計算を行った。ロイコトリエンは、生体防御に重要な役割を果たす生理活性脂質である一方、気管支喘息の発症・増悪因子として知られている。現在までに4種のLT受容体のクローニングがなされているが、そのなかでロイコトリエンB4受容体 (LT4R1_HUMAN) のアミノ酸配列情報を用いて実験を行った。

LT4R1_HUMANの配列についてウシロドプシン (1L9H) を鋳型にしてホモロジーモデリングを行った結果を初期構造として、TINKERで

全原子緩和のためのアニーリングを行った後、MDシミュレーションを行った。パラメータ等は以下の通りである。力場はamber99を用い、温度: 310K、圧力1気圧、Time Step: 1.0fsの条件下で文献から得られた制約条件を加えて、65.4ps間計算を行い、平均原子位置を求めた。

次に、活性型LT受容体ではヘリックス3, 5, 6がリガンドとの結合に関与しているらしいという文献情報があるので、この3本のみを自由に回転させた立体構造を実数値GAでの個体と定義した。実数値GAとドッキングシミュレーションを組み合わせた探索については、以下のように行った。

- 1) 初期集団の生成: 実数値GAでの集団プールにおける個体数は30とし、初期集団をランダムに生成する。
- 2) 世代ごとの生存選択: 集団プールからランダムに交叉のための親を選択し、UNDXで子を6個生成する。子個体 (6個) + 親個体 (2個) の計8個に対して立体構造の評価を下すために、すでにアゴニストとして知られている 12-keto-LTB4 (PubChem No. sid8233) を使ってGOLDでドッキングシミュレーションを行う。1つの個体当たりGOLDスコアの高い順に5つのリガンドのコンホメーションを記録しておき、ユーザーが画面でその結合状態をチェックする。ユーザーが結合状態を「良くない」と評価したものには負の評価を与える。そうして、各個体の最高スコアが決定したら、高い順に2個体を選択し (エリート戦略)、集団中の両親と置き換える。
- 3) 反復: ある停止条件が満たされるまで、ステップ2)を繰り返す。

今回は、時間的な制約から60世代の計算を行った。その進化計算によるスコアの推移を図1

に示す（スコアは高いほど良い）。なお、60世代後の回転角の最適解は、TM3: 127°、TM5: 174°、TM6: -26°であった。この最適解を持つLT受容体の立体構造とリガンドの結合状態を図2に示す。ヘリックス3, 5, 6とリガンドが相互作用しているらしいことが見てとれる。図3は3つのヘリックスの回転角の探索空間を示す。各ドットは探索されたポイントを示し、色が濃い方が、GOLDのスコアが高い。これによりそれほど偏りのない探索が行われたことがわかる。

最後に、本手法で得られたLT受容体構造の良さを検証するために、構造を作る計算に用いた以外のアゴニストでドッキングシミュレーションを行った。同様のことをホモロジーモデリン

グだけで作成したLT受容体の立体構造でも行い、得られた両者のGOLDスコアを表1に示す。表1から、本手法で得られた活性型GPCRの立体構造は、単にホモロジーモデリングから得られたものと比較してアゴニストとの良好な結合を示していることがわかる。

【今後の研究の方向、課題】

今後は、進化計算の評価関数スコアの精度をよくするための更なる検討を行うとともに、予測された構造の評価手法を開発する。また、ロイコトリエン以外のGPCRについても検証していきたい。

【成果の発表、論文等】

- 1) 石野洋子：進化計算を用いたGPCR活性型コンホメーションの探索研究，第35回人工知能学会・分子生物情報研究会（SIG-MBI），2007年3月（石川県）。

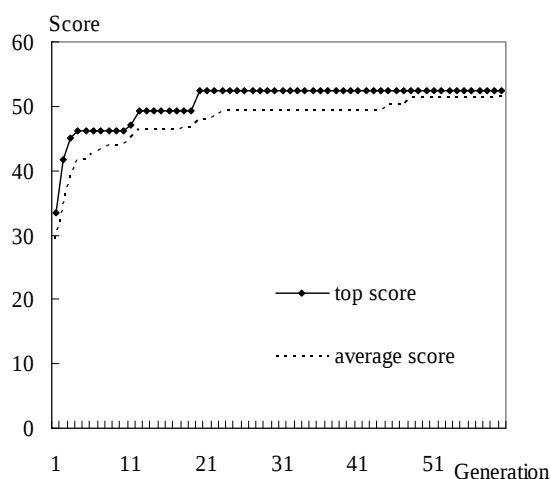


図1. 進化計算によるスコア変化

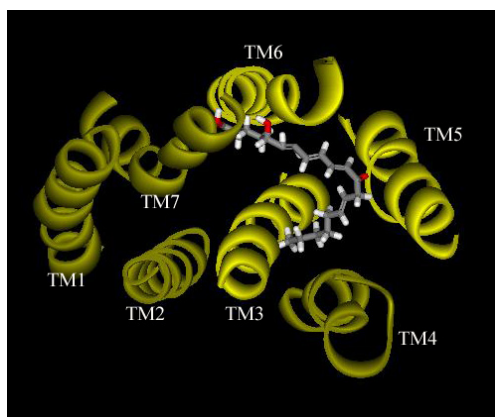


図2. 得られたドッキング構造

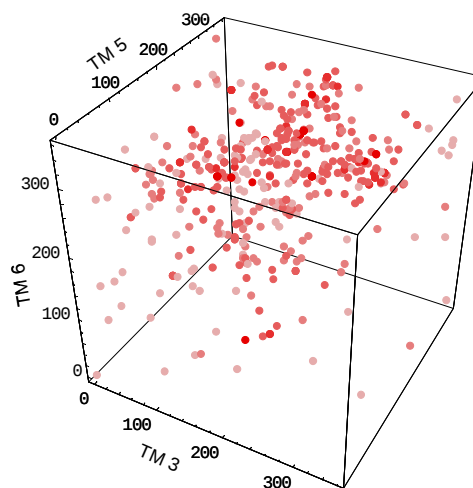


図3. 探索空間

表1. 他のアゴニストのGOLD Score

アゴニスト名	ホモロジーモデリング だけで作った構造	MD+進化計算で作った 構造
sid801685	1.9	45.5
sid801683	-16.1	43.7
sid8234	-12.4	42.5