

## 論文の内容の要旨

応用生命工学 専攻  
平成 17 年度博士課程 入学  
氏 名 森田 瑞樹  
指導教員名 清水 謙多郎

### 論文題目

## タンパク質-リガンド結合部位の予測手法の開発

### 1. タンパク質-リガンド結合部位の予測手法の開発

あらゆる生物の生体内においてタンパク質は他の何らかの生体分子と相互作用をしていると考えられ、それによってタンパク質は機能を発現している。タンパク質は水を除けば生体中に最も多く存在する生体高分子であるので、タンパク質の働きの詳細を知ることは生命現象を理解するうえで非常に重要である。そのためタンパク質と他の生体分子との相互作用は研究対象として大変注目されており、個々の相互作用のレベルからそれらの相互作用が織りなすネットワークのレベルまで幅広く研究されている。この中の一つの研究対象にタンパク質-リガンド結合部位の同定および予測がある。ここでリガンドは受容体タンパク質に結合する分子および酵素に結合する基質を指す。リガンドがタンパク質のどこに結合するかを知ることはタンパク質の機能解析の第一歩であり、また医薬品の探索や設計においても必要不可欠の情報であるため、それを同定・予測するための手法を開発することは非常に重要である。

近年、構造ゲノムプロジェクトの進展によって多数のタンパク質の立体構造が利用できるようになってきており、また、タンパク質の構造には機能に関する多くの情報を含むと考えられることから、本研究ではタンパク質の立体構造を利用し、そのどの部分にリガンドが結合するか、空間上の位置を示す手法を開発することとした。

立体構造のみを利用したリガンド結合部位の予測手法の主なものは大きく 2 通りに分け

られる。1つはタンパク質の形状に基づく手法、もう1つはエネルギーの計算に基づく手法である。形状に基づく手法はタンパク質の表面にある溝や凹みを検出し、その大きさや深さに基づいてリガンド結合部位を予測する手法である。一方のエネルギー計算に基づく手法はタンパク質周囲のエネルギー分布を計算し、タンパク質との引力が強い一帯を探すことによってリガンド結合部位を予測する手法である。本研究では生体分子の結合は生体分子間に働く引力的なエネルギーによって引き起こされるとの仮定に依拠し、エネルギーの計算に基づいてリガンド結合部位を予測する手法を開発することにした。

タンパク質の立体構造を利用したリガンド結合部位予測手法の最終的な目標は、リガンドが結合していないタンパク質の立体構造（リガンド非結合状態の構造）を用いてリガンドが結合する部位を予測できるようにすることである。しかし、これまでに発表された手法の多くではリガンドが結合した立体構造（リガンド結合状態の構造）を用いて手法のテストがなされている。計算機によるタンパク質-リガンド結合部位の予測は1982年のKuntzらの研究からはじまったが、25年以上経った今もなおリガンド非結合状態の構造のどこにリガンドが結合をするかを十分な精度で予測をすることができない。よって本研究ではリガンド結合状態の構造を用いた場合のみならず、リガンド非結合状態の構造を用いた場合においても高い精度でリガンド結合部位の予測ができる手法の開発を目指した。

## 2. 予測手法の概要

本研究で開発した手法では次のようにしてリガンド結合部位を予測する。タンパク質の表面近くにプローブを置き、タンパク質とプローブとの間に働く van der Waals 相互作用の強さを計算しながらプローブを挿引してゆく。このようにして強く相互作用をするようなまとまった場所を探し、そのような場所の中で全体として最も強く相互作用をする場所をリガンド結合部位の予測部位の第一候補とする（図1）。

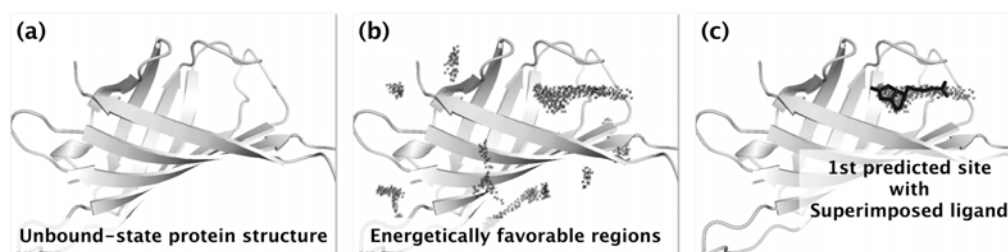


図1 本研究で開発した手法によるリガンド結合部位予測の概念図。(a) タンパク質の立体構造, (b) タンパク質との間に強い相互作用が働く場所を探す, (c) その中で最も強く相互作用をする場所をリガンド結合部位の第一候補とする。

本研究で開発した手法を従来手法と比較するために、従来研究の中から形状に基づく手法とエネルギーの計算に基づく手法を1つずつ選び、それらの手法がテストされたのと同じタンパク質で本手法をテストし、その結果を比較した。ここで従来手法として選んだのは2005年に Laurie と Jackson によって発表された Pocket-Finder (形状に基づく手法) と Q-SiteFinder (エネルギー計算に基づく手法) である。

### 3. 予測結果

本研究で開発した手法は従来手法よりも高い予測精度をあげることができた (表 I)。特にリガンド非結合状態での予測精度が従来手法よりも 20%以上向上しており、結合状態での予測精度に近い精度が得られた。また、本手法はリガンドが結合した場所をどれくらい厳密に予測できたかを表す指標である Precision の平均値 (Average precision) が従来手法と比べて高かった。これは従来手法よりも狭い範囲でリガンドの結合する位置を予測できたことを示している。また、本手法と Q-SiteFinder によって得られた Average precision は Pocket-Finder によって得られたものよりも格段に高い。このことはエネルギーの計算に基づく手法は、形状に基づく手法と比べてより狭い範囲でリガンドの結合する位置を予測できたことを示している。

表 I 各手法によって得られた予測結果の概要

		Ligand in the 1st predicted site (Precision $\geq$ 0.25)	Average precision
This method	Bound	0.800	0.839
	Unbound	0.743	0.771
Q-SiteFinder	Bound	0.743	0.739
	Unbound	0.514	0.619
Pocket-Finder	Bound	0.714	0.375
	Unbound	0.514	0.354

### 4. データセットを作製するウェブサーバーの開発

タンパク質-リガンド結合部位の予測手法の開発においては、予測はリガンド非結合状態のタンパク質の立体構造を用いて行い、予測結果の評価はリガンド結合状態の立体構造を

用いて行う必要がある。そのようにするためには、同じタンパク質のリガンド結合状態と非結合状態のペアの立体構造が必須である。またこのようなペアのデータセットは、タンパク質へのリガンドのドッキング・シミュレーション手法の開発や、リガンド結合によって引き起こされるタンパク質の立体構造変化の研究などのためにも欠かすことができない。このように、タンパク質-リガンド結合に関する研究にとってリガンド結合/非結合状態のペアの立体構造データセットは大きな価値があるが、このようなデータセットは現在非常に限定された内容のものしか存在しない。そのため本研究では、リガンド結合状態と非結合状態の立体構造のペア・データセットを作製するウェブサーバーを新たに開発した（図2）。

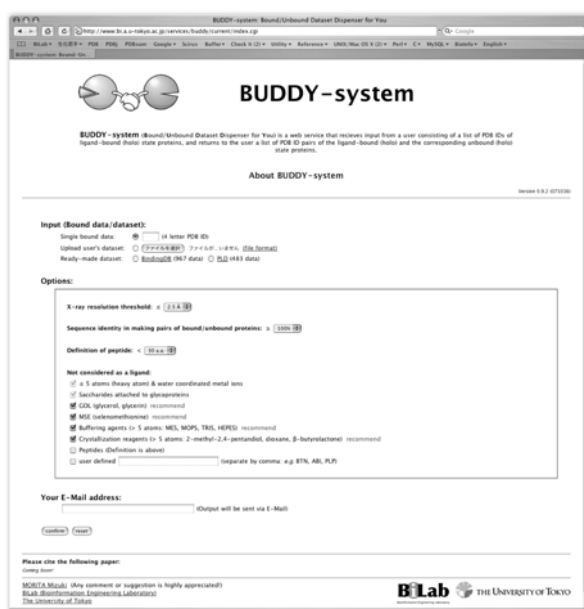


図2 本研究によって開発・公開したウェブサーバー：BUDDY-system

<http://www.bi.a.u-tokyo.ac.jp/services/buddy/>

このウェブサーバーはユーザーがリガンド結合状態のタンパク質の PDB ID を入力すると、同じタンパク質のリガンド非結合状態の PDB ID をそれとペアにしてユーザーに返す。1つの PDB ID だけでなく、複数の PDB ID を同時に入力とすることができる。このウェブサーバーはすでに一般に公開している。