



Title	Enhancing Precision in Antibody-Antigen Complex Structure Prediction Through Parametric Optimization of RosettaAb Docking Scoring Function [an abstract of dissertation and a summary of dissertation review]
Author(s)	Sangeetha Udani Ratnayake, RATNAYAKE MUDIYANSELAGE
Citation	北海道大学. 博士(情報科学) 甲第16067号
Issue Date	2024-06-28
Doc URL	<a href="http://hdl.handle.net/2115/92806">http://hdl.handle.net/2115/92806</a>
Rights(URL)	<a href="https://creativecommons.org/licenses/by/4.0/">https://creativecommons.org/licenses/by/4.0/</a>
Type	theses (doctoral - abstract and summary of review)
Additional Information	There are other files related to this item in HUSCAP. Check the above URL.
File Information	RATNAYAKE_MUDIYANSELAGE_Sangeeth_review.pdf (審査の要旨)



[Instructions for use](#)

## 学位論文審査の要旨

博士の専攻分野の名称 博士 (情報科学) 氏名 RATNAYAKE  
MUDIYANSELAGE Sangeetha Udani Ratnayake

審査担当者 主査 准教授 長田 直樹  
副査 教授 遠藤 俊徳  
副査 客員教授 今西 規

## 学位論文題名

### Enhancing Precision in Antibody-Antigen Complex Structure Prediction Through Parametric Optimization of RosettaAb Docking Scoring Function

(RosettaAb Docking スコアリング関数のパラメータ最適化を通じた抗体-抗原複合体構造予測の精度向上)

抗体医薬は、特定の病原体やがん細胞に対して高い特異性を持つタンパク質であり、現代医療において重要な役割を果たしている。これらは人体の免疫システムが自然に生成する抗体を模倣または改良したもので、特定のターゲットに結合し、病気の進行を阻止または遅らせることが可能である。抗体のデザインは、疾患特異的な治療効果を高めるために重要であり、ターゲットの精度と治療成果を大きく左右するが、実現しうる抗体の種類は天文学的な組み合わせになるために、効率よく抗体をデザインするためには、実験的な手法に加えて、計算機科学的なアプローチが重要となっている。

タンパク質の立体構造予測にはさまざまな方法があるが、スコアリング関数と呼ばれる、ドッキングした2つの分子の分子間における結合親和性を概算するために使用される関数を利用したものが広く用いられている。その中でも RosettaAntibodyDesign(RAbD) と呼ばれるソフトウェアは、抗原・抗体構造の予測において最も幅広く用いられているものの1つである。このソフトウェアは抗原・抗体構造の予測に特化したものであるものの、予測性能があまりよくないことが知られている。原因の一つとして、スコアリング関数が、一般的なタンパク質同士の結合とは違う物理化学的特徴をもつ抗原・抗体構造に対して最適化されていない可能性が指摘される。例えば、一般的なタンパク質の結合では疎水性のアミノ酸が結合に重要な役割を果たすが、抗原・抗体構造では親水性のアミノ酸が重要な役割を果たすことが知られている。そこで本研究では、RAbD をケーススタディとして、これまで詳細に検討されてこなかった抗原・抗体構造の予測精度とスコアリング関数との関連について詳細な解析を行い、スコアリング関数に用いられるパラメータの最適化を行うことによって、よりよい抗原・抗体構造の予測法を提案する。

本論文は大きく2つのパートに分かれているが、どちらの研究でも、RAbD や他の多くのソフトウェアで用いられている decoy distribution というものに注目して研究を行っている。Decoy とは、立体構造の予測最適化の前段階として、ランダムな初期位置をもつ分子から簡便な手法で求められた安定的な構造である。実験的に真の構造がわかっている場合、100 個程度の異なった decoy について、横軸に実験で確かめられた構造とのズレ (RSMD)、縦軸に推定された自由エネルギーをプロットしたものが、decoy distribution と呼ばれる。自由エネルギーを計算するスコアリング関数が最適なものであれば、decoy distribution は漏斗型 (funnel-like shape) を取ることが経験的にわかっている。

最初の研究では、立体構造が既知のヒトおよび哺乳類の抗原・抗体をデータベースから取得し、RAbD から得られる decoy distribution の定性的評価を行った。その結果、多くの場合、RAbD は非常に悪い形の decoy distribution を示した。このことは、このデータセットに対して RAbD がうまく機能していないことを示している。しかし、RAbD のスコアリング関数の中から、アミノ酸の疎水性の程度に基づき結合性を評価する項を削除することによって、わずかながら性能の向上が見られることを示した。また、抗体がもつアミノ酸配列のうち、どのような特徴をもつアミノ酸が抗体に存在すると理想的な decoy distribution が得られないかについて、ロジスティック回帰モデルを用いた解析を行ったところ、芳香族アミノ酸などが抗体の結合部位にある場合に RAbD がうまく働かないことが多い、などの特徴を発見した。

次の研究では、decoy distribution を指標としてスコアリング関数のパラメータを最適化する試みを行った。そのために、decoy distribution の形を客観的に評価できる量的なスコアを 3 種類提案し、それぞれを用いて、100 個の抗原・抗体構造ごとにパラメータの最適化を行った。最適化後のパラメータを用いて再び抗原・抗体の立体構造予測を行ったところ、多くのケースにおいて decoy distribution は改善され、RMSD も減少した。また、最適化の指標としてはピアソンの相関係数が最も性能が良かった。これらのことは最適化後のスコアリング関数を用いることによって、より正確な抗体・抗原構造が予測できたことを示している。さらに、抗体・抗原構造ごとに予測されたパラメータセットを主成分分析によってクラスタリングし、多くの抗原・抗体構造に適用可能なパラメータセットを探索した。このパラメータセットは、100 個中 91 個の抗原・抗体構造に対して RMSD の減少をもたらした。また、検証用のデータセットとして新たに取得した 10 個の抗原・抗体構造についてこのパラメータセットを試したところ、標準的なパラメータセットと比較して、10 個すべてにおいて RMSD の減少が確認された。

以上を要するに、著者は、抗原・抗体構造の計算機科学的予測という社会的にも重要な問題において、独自の視点により新たな予測向上の手法を示したものであり、情報科学、特に、生命情報学の進歩に貢献するところ大なるものがある。よって著者は、北海道大学博士 (情報科学) の学位を授与される資格あるものと認める。