

(様式 1-1)

提出日：2026 年 5 月 18 日

2025 年度 大阪大学蛋白質研究所 拠点事業

研究成果報告書

(1) 事業名（下記より該当事業名を選択し、ほかは削除してください。）

共同研究員

(2) 研究代表者

氏名：村上洋一

所属機関名・部局名・職名：東京情報大学・総合情報学部総合情報学科・教授

(3) 研究課題名（申請時に記載したものと同一課題名を記入してください。）

ゲノムスケールでの蛋白質間相互作用予測及び統合データベースへの実装

(4) 蛋白質研究所受入担当教員

教員名：水口賢司（研究室名：計算生物学研究室）

(5) 研究成果の概要

＊背景および目的、方法と結果について、公開して差し支えない範囲で記載。

生命現象の包括的理解や効率的な創薬にはゲノムスケールでの蛋白質間相互作用（PPI）ネットワークの解明が不可欠であり、統合データウェアハウス（TargetMine）への予測データの実装が望まれている。その実現には「2つの蛋白質が相互作用するか否か」を判定する技術が必要となるが、相互作用を直接担う「部位（界面）」の正確な把握は、TargetMineにおける蓄積データの解像度を向上させる上でも重要である。そこで本研究課題では、ゲノムスケール予測の核心的基盤として、多様性と非対称性ゆえに予測が困難とされるヘテロ蛋白質の相互作用部位予測の高精度化を目的とした。

目的達成のため、最新の大規模言語モデル（PLM）と立体構造予測を融合した4工程のパイプラインを構築した。まずPDB等から取得したヒト由来ヘテロ複合体データから学習用データセット(8,742本)を作成した。特徴抽出では、進化的情報を捉える「ESM-C」と構造トークンを学習した「ProstT5」を併用し、連結特徴量からLightGBMにより残基単位の部位予測モデルを構築した。さらにAlphaFold2の予測構造を利用し、陽性予測残基が空間的（半径12.3Å）に孤立している場合に偽陽性とみなして排除する構造的フィルタリングを導入した。

評価の結果、物理化学的特性に強いESM-Cと二次構造認識に優れたProstT5の相補的な相乗効果により、界面の識別能力が向上した。構造フィルタリングによるノイズ抑制も寄与し、ヘテロ蛋白質の部位予測において先行手法をわずかに上回る高い予測精度（MCC 0.522、F1-Score 0.614、ROC-AUC 0.86）を達成した。今後はこの部位予測技術を、2蛋白質間の相互作用の有無を判定する予測モデルへと応用し、当初の目的であるゲノムスケールでの相互作用予測の実現とTargetMineへの実装を目指す。