

2024 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。)

自然言語処理の埋め込み法を用いたゲノムスケールでの蛋白質間相互作用予測法の高精度化

(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。

本研究では、ゲノムスケールでの蛋白質間相互作用 (以下, PPI) 予測システムの高精度化を目指し、そのための要素技術として蛋白質の立体構造予測の深層学習モデルである AlphaFold2 (以下, AF2) の活用も進めている。AF2 は、分子機能解明や創薬研究等において重要なツールである。さまざまな蛋白質が関わる分子間相互作用の予測において、これまで構造情報が未知の場合、配列情報のみを用いる方法が基本であったが、AF2 によって予測された構造の積極的な活用が期待される。しかしながら、未知の蛋白質に対する AF2 の予測精度、蛋白質全体だけでなく PPI の部位及びその周辺構造の予測精度について十分な検証がされていない。そこで本研究では、AF2 の信頼性を検証するために、蛋白質全体及び PPI 部位及びその周辺構造の AF2 の予測精度について検証を行った。具体的には AF2 開発以降に PDB に登録された蛋白質を検証用データとして用いて、予測構造上の PPI 部位及びその周辺構造と、天然構造上のそれらの構造に対して Distance Map (以下, DM) 及び RMSD を用いて比較を行った。

その比較の結果、DM を用いた解析では、検証用データに含まれる蛋白質の天然構造とその予測構造を比較した構造ペア数全体の約 89.2%において、PPI 部位及びその周辺構造ともに高い精度 ($\leq 2.0 \text{ \AA}$) 予測されていることが確認できた。一方、天然変性領域上にある PPI 部位及びその周辺構造では、構造予測が困難な事例が複数確認できた。RMSD を用いた解析においても同様に、蛋白質の天然構造とその予測構造を比較した構造ペア数全体の約 98.4%において、PPI 部位及びその周辺構造ともに高い精度 ($\leq 3.0 \text{ \AA}$) 予測されていることが確認できた。すなわち、一部の天然変性領域を含む蛋白質においては、構造予測が困難な事例もあったが、AF2 は全体構造と PPI 部位及びその周辺構造を高い精度で予測できる可能性が確認できた。今後、AF2 やその AlphaFold3 の最新版を利活用することで、PPI 及び PPI 部位予測等の精度向上が図られることが期待される。