

## 研究開発課題別事後評価結果

### ▶ 課題情報

研究開発課題名	蛋白質構造データベースのデータ検証高度化と統合化
研究代表名	栗栖 源嗣（大阪大学 蛋白質研究所 教授）

### ▶ 事後評価結果

#### 1. 総合評価

当初の実実施計画がすべて達成され、期待を上回る成果が得られた。

構造生物学が重要な創薬基盤技術になっているという高い社会的ニーズの中で、構造生物学の国際的な情報基盤構築を目指す研究開発課題であった。厳しい品質管理の下、立体構造情報と NMR 実験情報のアーカイブ構築を実施し、wwPDB 全体の約 24% のデータ登録と公開を遅滞なく実施した点や、急増するクライオ電顕の構造データの品質検証手法を迅速に確立しアーカイブを構築した点を高く評価する。利活用状況から、PDB は構造生物学には欠かせない研究基盤として定着しており、その中で PDBj が wwPDB のアジア拠点として国際的に大きく貢献している点を評価する。

今後、中国・インドに PDB 拠点が立ち上がることから、PDB の日本拠点としての PDBj の意義について、コミュニティの理解をより丁寧にて得ていく必要がある。利用者へのインタビューや利用状況のモニタリングなどを通じ、PDBj が独自に提供するデータ検索・閲覧インターフェースの利便性向上、ASH viewer などの PDBj 独自ツールのアピールなどを通じた一層のプレゼンスの向上を期待する。

#### 2. 研究コミュニティを含むデータ提供者やDB利用者との連携・協業

wwPDB として Nucleic Acids Research 誌に最新情報を掲載するなど、研究コミュニティに対して適切な情報発信がされた。wwPDB の中心メンバーとしての活動や学会での展示ブース出展・講習会の開催等、関連コミュニティと積極的に協議・連携・協力が行われ、適切に研究開発が進められた。システムバイオロジーや代謝工学への展開、COVID-19 の特設ページの開設等、利用者や社会的ニーズに対しても迅速な対応が実施された点を評価する。

#### 3. 利用者にとって有用なデータ基盤、持続的な DB 運用体制構築に向けた取り組み

OneDep の継続開発や BMRBdep の導入で、データ登録の利便性向上が図られた。The Cambridge Structural Database (CSD) との統合化により X 線結晶構造のリガンド複合体解析の利便性が向上した。クライオ電顕データの品質管理や標準化に迅速に対応し、急増するクライオ電顕データが収集された。糖鎖表記が糖鎖コミュニティで使われている標準的な記述法に統一されたことで、糖タンパク質の検索と構造解析の利便性が著しく向上した。上記の通り、ユーザーニーズや構造生物学の進捗に対応し、迅速かつ適切にデータベース開発を進めた点を評価する。

4. 生命科学研  
究やイノベ  
ーションへ  
の波及効果  
(DBを利用  
して得られ  
た研究成果)

月間訪問者数は約 7.2 万件（2021 年度平均（推定））で、研究開発期間初年度の約 1.8 倍増となった。構造生物学には欠かせない情報基盤として定着している。データベースの利活用により多くの研究成果やイノベーションが創出されており、科学技術や産業の発展、社会的課題の解決等に大きく貢献した点を評価する。

本研究開発期間中に公開された画期的な構造予測ツール AlphaFold 2 および RoseTTAFold ならびに AlphaFold 2 を用いた予測構造データベース AlphaFold DB は、PDBj を含む wwPDB による長年の膨大な構造情報があってこそ実現したものであり、公共データベースの重要性や意義を知らしめるものであった。

5. その他特記  
すべき事項

中国・インドの PDB 拠点の立ち上げを支援するなど国際的な貢献が認められた。wwPDB と連携してアノテーター育成を継続的に実施し、安定的運用に寄与した点を評価する。

6. 実施計画の  
達成状況

PDBj を研究開発対象とし、1) PDB アーカイブの構築・データ検証・公開ならびに検証レポートの拡充とセマンティック化、2) 他のデータベースとの統合化と高度化、3) データベースの利用促進・人材育成を目的とした研究開発が実施された。

1)については、wwPDB メンバーとしてアジア・中東地域を分担し、厳しい品質管理の下、立体構造情報と NMR 実験情報をキュレーションし、wwPDB 全体の約 24% のデータ登録と公開が遅滞なく実施された。wwPDB と協力してクライオ電顕と NMR の構造データの品質検証手法を確立し、検証レポートの提供が開始された。X 線結晶解析データの検証レポートの RDF 化を進め、PDBj が独自に高度化した検証ファイルが wwPDB から公開された。糖鎖構造記述法を糖鎖コミュニティ標準の表記に統一し、PDB の糖鎖データを修正するとともに、全ての PDB ファイルに糖鎖構造リポジトリ GlyTouCan へのリンクを自動的に組み込む準備が進められた。継続的な wwPDB との連携により、データ登録システム OneDep の高度化、BMRBdep の開発、NMR 構造データベース登録業務の自動化と簡略化が実現された。

2)については、低分子化合物データベース CSD を統合することで CSD ID で検索できるようにするとともに、CSD の構造化学情報を使った X 線結晶構造解析の電子密度マップの計算係数が公開された。また 3D 構造同士の構造アラインメントサービス ASH viewer の開発が進められた。PDB アーカイブの mmCIF ファイルから蛋白質分子内コンタクト情報を自動抽出して PDBj Mine2 の関係データベースに追加し、立体構造から得られる情報とその構造メタデータの統合検索が実現した。NMR 構造座標と実験データの整合性を検証するツールが開発され、NMR 構造データの検証レポートが充実化した。また OneDep システムで NMR-START または NEF 形式の単一ファイルで NMR 実験データが登録できるようになった。

3)については、蛋白質立体構造データベース専門部会を通じ、企業を含むユーザーの要望を踏まえた開発が実施された。生物工学会等で研究成果を発信し、代謝工

学やシステム生物学の研究者との意見交換が行われた。利用者・研究者向けのセミナーや講習会を毎年開催し、ユーザー拡大と利用促進が図られた。欧州 PDBe と米国 RCSB PDB と合同で OneDep Summit を毎年開催し、アノテーターのスキルアップと世界同一基準でのデータ登録が実現された。wwPDB の Advisory Committee の意向を踏まえ、PDB China 立ち上げを支援した。

上記の通り、当初の実施計画を大幅に上回る成果が得られた。