ライフサイエンスデータベース統合推進事業
統合化推進プログラム(平成 26 年度採択)
課題名「蛋白質構造データバンクの高度化と統合的運用」(研究代表者　中村　春木)

PDBj と PDBj-BMRB における生体高分子の構造とダイナミクス情報

藤原　敏道
(大阪大学蛋白質研究所)

タンパク質、核酸、糖鎖などの生体高分子の立体構造情報は、構造決定を行った研究者がデータを寄託する形でのデータバンク方式によって Protein Data Bank (PDB)として集積されている。PDB は、1971 年に米国で誕生し 2014 年 5 月中旬には 10 万件を超える構造情報が登録されて、さらに増加している。2003 年には、米国 The Research Collaboratory for Structural Bioinformatics (RSCB)-PDB、欧州 The European Bioinformatics Institute (EBI)-PDBe とともに大阪大学蛋白質研究所が運営する PDBj (PDB Japan、 http://pdbj.org/)が協力して、国際的な worldwide PDB (wwPDB、http://wwpdb.org/)を設立した。その後、核磁気共鳴 (NMR) 法によって解析された NMR 実験データを集積している BMRB (BioMagResBank)が wwPDB に加わり、PDBj のグループも PDBj-BMRB としてデータの登録と公開サービスに協力している。さらに電子顕微鏡画像データのデータベースである The Electron Microscopy Data Bank (EMDB)とも連携を行い、国際連携の下、構造生物学の多様な一次データが wwPDB に集積されている [1)]。PDBj と BMRB-PDBj では wwPDB に共通した品質管理によるデータ登録作業と、世界共通データのダウンロードサイトの運営および独自に開発した種々のサービス・ツールや二次データベースの提供を行っている [2、3)]。また、NBDC による第一期の統合化プログラムにより全登録データの RDF 化などセマンティック化も達成している。

本年度からの NBDC による第二期のプログラムにおいては、これらのデータ登録の高度化（データ検証の改良と検証レポートの公開・配布、新たな PDBx/mmCIF フォーマットの利用の促進）と配布事業の継続に加えて、①リガンド・蛋白質複合体解析による医薬スクリーニングデータの精密化・登録支援、②異なる階層のデータと蛋白質構造情報との統合化のために特に蛋白質複合体に対する多面的解析ツール・サービス（分子表面データベース[4)]、基準振動による蛋白質ダイナミクス・データベース[5)]、PDB と EMDB の双方に対する構造形状の同時比較・検索等）の開発と公開、③統合化された生体系 NMR データの国際的なアノテーション・登録受付を行う。NMR データは生体高分子のダイナミクスに関する情報を与えるためそれらに特化した情報を BMRB に蓄積すると共に、タンパク質動的部位を視覚的に表示するツールによって公開していく。特に NMR 構造の収斂領域を自動的かつ定量的に検出するツールも開発し公開しており[6)]、ダイナミクス情報の解析へ応用していく。また、NMR データの解析と登録の支援ツールの開発と公開を行う。さらに、NMR データと相互作用情報、一塩基置換情報などとリンクしたデータの公開をセマンティック技術により実現していく。

参考文献

1) Berman H. M. et al.、 Structure、 21、 1485-1491、 2013.
2) 中村春木、生物物理、 53、 44-46、 2012.
3) Kinjo A. R. et al.、 Nucl. Acids Res.、 40、 D453-D460、 2012.
4) Murakami Y. et al.、 Protein Science、 22、 1379-1391、 2013.
5) Wako H.、 Endo S.、 Comput. Biology. Chem.、 44、 22-30、 2013.
6) Kobayashi N.、 J. Biomol. NMR、 58、 61-67、 2014.