Growing the PDB Core Archive Using ORCiD Login for Depositing 3D Biostructures with OneDep

Gregg V Crichlow[1], Zukang Feng[1], Yuhe Liang[1], Ezra Peisach[1], Irina Persikova[1], Jasmine Y Young[1], wwPDB Team wwPDB Team[2], Stephen K Burley[3]

[1]RCSB Protein Data Bank, Rutgers, The State University of New Jersey, [2]RCSB Protein Data Bank, Rutgers, The State University of New Jersey, PDBe, EMBL-European Bioinformatics Institute,, PDBj, Institute for Protein Research, Osaka University, EMDB, EMBL-European Bioinformatics Institute, BMRB, UConn Health, University of Connecticut, RCSB Protein Data Bank, San Diego Supercomputer Center, University of California San Diego, [3]RCSB Protein Data Bank, Rutgers, The State University of New Jersey

gregg@rcsb.rutgers.edu

The Protein Data Bank (PDB) was established in 1971 as the first open-access digital data resource in biology with just seven X-ray crystallographic structures of proteins. Today, the single global PDB archive houses more than 200,000 experimentally-determined three-dimensional (3D) structures of biological macromolecules that are made freely available to many millions of users worldwide with no limitations on usage. 3D biostructure information facilitates basic and applied research and education across the sciences, impacting fundamental biology, biomedicine, biotechnology, and energy sciences. The Worldwide Protein Data Bank partnership (wwPDB, wwpdb.org) currently includes five Full Members (RCSB PDB, PDBe, PDBj, BMRB, and EMDB) and one Associate Member (PDBc), which together manage the PDB, EMDB, and BMRB Core Archives. wwPDB Members are committed to ensuring that structural biology data are Findable, Accessible, Interoperable, and Reusable (FAIR).

Depositing macromolecular structures together with experimental data to the PDB is mandatory for publication of new 3D biostructures in many peer-reviewed scientific journals. The global wwPDB OneDep software system is jointly maintained by wwPDB Members.

OneDep currently supports complete data deposition, rigorous validation, and expert biocuration of incoming PDB structures generated using macromolecular crystallography (MX), 3D electron microscopy (3DEM), nuclear magnetic resonance (NMR) spectroscopy, and micro-electron diffraction (microED). The system provides user interfaces for PDB Data Depositors to log in, upload files, review validation reports and correct their data as needed, and provide mandatory metadata information. PDB identifiers and associated PDB DOIs are assigned at the time of submission. wwPDB Members are committed to ensuring the highest possible data quality for PDB structures. OneDep software undergoes constant improvement to enhance metadata completeness and quality, support advances in science and technology, and incorporate community feedback. Recent OneDep improvements encompassed additional data validation for MX, 3DEM, and NMR structures, and ORCiD login to access all your depositions. Herein, we present OneDep in detail, introduce recently added features, and explain how PDB Data Depositors can use the new streamlined ORCiD login system to initiate/access their PDB depositions. RCSB PDB is funded by the National Science Foundation (DBI-1832184), the US Department of Energy (DE-SC0019749), and the National Cancer Institute, National Institute of Allergy and Infectious Diseases, and National Institute of General Medical Sciences of the National Institutes of Health under grant R01GM133198.