OneDep, a system for the deposition, biocuration, and validation of experimentally determined structures of biological macromolecules submitted to the Protein Data Bank (PDB) archive, has been developed as a global collaboration by the Worldwide Protein Data Bank partners (wwPDB, http://wwpdb.org)(Berman et al., 2003; Young et al., 2017). This system was designed to ensure that the wwPDB could meet the challenges of rapidly evolving technologies in structural biology and the evolving archiving requirements of the scientific community over the coming decades.

The OneDep system is underpinned by a flexible data architecture based on the PDBx/mmCIF dictionary (http://mmcif.wwpdb.org)(Westbrook et al., 2005). OneDep software tools leverage PDBx/mmCIF dictionary metadata to: validate data integrity; produce structure and experimental data files in multiple file formats; define data transformations required for cross-repository data exchange; and define mappings between the schema of a variety of database implementations. Extensibility is another key feature of PDBx/mmCIF dictionary technology enabling the wwPDB to respond to a rapidly evolving scientific landscape. To support scientific advancement and ensure the highest data quality and completeness, a working group of community experts in structural biology software works with the wwPDB to define new data content and implement these content extensions in software tools across the structure determination pipeline.

This talk will present how PDBx/mmCIF has been incorporated into the wwPDB OneDep data and software architecture, and how this architecture is enabling the development of more informative versioning of data in the PDB repository.

wwPDB Members: RCSB PDB (supported by NSF DBI-1338415, NIH, and DOE); PDBe (EMBL-EBI, Wellcome Trust, BBSRC, EU, and MRC); PDBj (JST-NBDC); and BMRB (NIGMS).


Keywords: PDBx/mmCIF, wwPDB, OneDep Deposition Biocuration and Validation