

Validating the Structure at the wwPDB

Jasmine Y. Young

RCSB PDB, Rutgers, The State University of New Jersey, Piscataway, New Jersey, United States

The Protein Data Bank (PDB) is the single global repository for three-dimensional structures of biological macromolecules and their complexes. Over the past decade, the size and complexity of macromolecules and their complexes with small molecules deposited to the PDB have increased significantly. The PDB archive now holds more than 125,000 experimentally determined structures of biological macromolecules, which are all publicly accessible without restriction. These structures provide essential information to a large, diverse user community worldwide. There are more than 590 millions data file downloads from the PDB archive and more than 1 million unique IP addresses access the archive every year.

The worldwide PDB (wwPDB) has developed a global unified system, called OneDep, for deposition, biocuration, and validation of macromolecular structures to the PDB to meet the evolving archiving requirements of the scientific community over the coming decades. The validation module of this system produces validation reports for structures determined by X-ray crystallography, nuclear magnetic resonance spectroscopy, and electron microscopy. These reports were developed with the benefit of recommendations from expert task forces representing the respective communities.

The importance of validation, content of the wwPDB validation report, and accessing the stand-alone, anonymous wwPDB validation server will be presented.

wwPDB members are RCSB PDB (supported by NSF, NIH, and DOE), PDBe (EMBL-EBI, Wellcome Trust, BBSRC, NIGMS, and EU, and MRC), and PDBj (JST-NBDC-JST), and BMRB (NIGMS).