MS10-P04 | VALIDATION IN THE PDB

Armstrong, David (PDBe, Hinxton, GBR); wwPDB consortium (wwPDB consortium, AUT)

The Protein Data Bank (PDB) is one of the oldest biological archives. Having started with 7 entries in 1971, it has grown to over 150,000 entries. The way the quality of archival entries was understood and evaluated has evolved over the years. In 2008-2009, the wwPDB, the worldwide organisation which manages the PDB, set up community expert task forces to obtain recommendations on validation metric for coordinate models in the PDB archive, supporting experimental data and the correspondence between them. These recommendations were implemented in the wwPDB validation pipeline and presented in the wwPDB validation reports.

The wwPDB validation reports summarise the overall quality of each PDB entry in a simple slider image. The sliders allow non expert users to quickly judge the quality of each PDB entry at a glance and select the best quality model for their work, without having to understand all the validation metrics that were used to create the sliders. The wwPDB validation service also produces a comprehensive report which contains details of every outlier in the PDB entry in PDF and XML format.

The wwPDB validation service is made freely available via the wwPDB website (validate.wwpdb.org) and through an API and users are encouraged to use these services prior to deposition to the PDB. Following curation of a PDB entry an official wwPDB validation report is produced which depositors are encouraged to provide to the referees of their publication and is required by an increasing number of journals.