MS107.O04

PDBx/mmCIF: the foundation for the wwPDB onedep system

John Westbrook[1], John Markley[2], Haruki [3], Sameer Velankar[4], Stephen Burley[5]

[1]*Rutgers, The State University Of NJ, Piscataway, United States,* [2]*BMRB, BioMagResBank, University of Wisconsin-Madison, Madison, United States,* [3]*Protein Data Bank Japan, Institute for Protein Research, Osaka University, Suita, Osaka, Japan,* [4]*Protein Data Bank in Europe, European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, United Kingdom,* [5]*RCSB PDB, Rutgers, The State University of New Jersey, Piscataway, United States*
E-mail: jwest@rcsb.rutgers.edu

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.

Berman, H. M., Henrick, K. & Nakamura, H. (2003). Nat Struct Biol 10, 980.

Westbrook, J., Henrick, K., Ulrich, E. L. & Berman, H. M. (2005). Vol. G. Definition and exchange of crystallographic data, International Tables for Crystallography, edited by S. R. Hall & B. McMahon, pp. 195-198. Dordrecht, The Netherlands: Springer.

Young, J., Westbrook, J. D., Feng, Z., Sala, R., Peisach, E., Oldfield, T. J., Sen, S., Gutmanas, A., Armstrong, D. R., Berrisford, J. M., Chen, L., Chen, M., Di Costanzo, L., Dimitropoulos, D., Gao, G., Ghosh, S., Gore, S., Guranovic, V., Hendrickx, P. M. S., Hudson, B. P., Igarashi, R., Ikegawa, Y., Kobayashi, N., Lawson, C. L., Liang, Y., Mading, S., Mak, L., Mir, M. S., Mukhopadhyay, A., Patwardhan, A., Persikova, I., Rinaldi, L., Sanz-Garcia, E., Sekharan, M. R., Shao, C., Swaminathan, G. J., Tan, L., Ulrich, E. L., Van Ginkel, G., Yamashita, R., Yang, H., Zhuravleva, M. A., Quesada, M., Kleywegt, G. J., Berman, H. M., Markley, J. L., Nakamura, H., Velankar, S. & Burley, S. K. (2017). Structure http://dx.doi.org/10.1016/j.str.2017.01.004.

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