HHS Public Access

Author manuscript

Structure. Author manuscript; available in PMC 2018 February 21.

Published in final edited form as:

Structure. 2017 September 05; 25(9): 1317–1318. doi:10.1016/j.str.2017.08.001.

PDB-Dev: a Prototype System for Depositing Integrative/Hybrid Structural Models

Stephen K. Burley[1,2,3,*], **Genji Kurisu**[4], **John L. Markley**[5], **Haruki Nakamura**[4], **Sameer Velankar**[6], **Helen M. Berman**[1], **Andrej Sali**[7], **Torsten Schwede**[8], and **Jill Trewhella**[9,10]

[1]RCSB Protein Data Bank, Institute for Quantitative Biomedicine, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA [2]Skaggs School of Pharmacy and Pharmaceutical Sciences and San Diego Supercomputer Center, University of California, San Diego, La Jolla, CA 92093, USA [3]Rutgers Cancer Institute of New Jersey, Rutgers, The State University of New Jersey, New Brunswick, NJ 08903, USA [4]Protein Data Bank Japan (PDBj), Institute for Protein Research, Osaka University, Osaka 565-0871, Japan [5]BMRB, BioMagResBank, Biochemistry Department, University of Wisconsin-Madison, Madison, WI 53706, USA [6]Protein Data Bank in Europe (PDBe), European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire CB10 1SD, UK [7]Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry, California Institute for Quantitative Biosciences, University of California, San Francisco, UCSF MC 2552, Byers Hall at Mission Bay, 1700 4th Street, Suite 503B, San Francisco, CA 94158, USA [8]SIB Swiss Institute of Bioinformatics and Biozentrum, University of Basel, Klingelbergstrasse 50-70, 4056 Basel, Switzerland [9]School of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia [10]Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA

With this Letter to the Editor, the World-wide PDB (wwPDB) Partnership (wwpdb.org) and the wwPDB Integrative/Hybrid (I/H) Methods Task Force would like to announce public release of a prototype system for depositing I/H structural models, PDB-Development (or "PDB-Dev") (Vallat et al., 2016c). The URL for PDB-Dev is https://pdb-dev.wwpdb.org.

Essential mechanisms in biology frequently involve large macromolecular assemblies (or machines). In favorable cases, their structures can be determined by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy or electron microscopy (3DEM) alone, culminating in deposition of atomic structural models (with x, y, z atomic coordinates) into the global PDB archive (PDB; pdb.org). Many such biological machines are, however, poorly suited to single experimental method approaches. Researchers are increasingly forced to combine various experimental data and information from measurements and computational analyses to generate "hybrid" or "integrative" structural models of macromolecular assemblies (Ward et al., 2013). In addition to X-ray crystallography, NMR, and single- particle 3DEM data, structural information in the form of spatial restraints can be obtained from a multitude of measurements, including small-angle scattering, atomic force

*Correspondence: stephen.burley@rcsb.org.

microscopy, chemical cross-linking, co-purification, Förster resonance energy transfer, electron paramagnetic resonance, mass spectrometry, hydrogen/deuterium exchange, cryoelectron tomography with sub-tomogram averaging, correlative fluorescent light microscopy, and various proteomics and bioinformatics analyses (Ward et al., 2013). I/H approaches have yielded informative structural models of very large macromolecular assemblies, such as the nuclear pore complex and its sub-complexes, the type III secretion system needle, the proteosomal lid sub-complex, the ESCRT-I complex, and an RNA ribosome- binding element from the turnip crinkle virus genome. Despite great need, there are, at present, no standard mechanisms available to represent, deposit, validate, biocurate, archive, disseminate, and visualize I/H models, their supporting experimental data and metadata, and the protocols used to compute the structural models so that they are freely available to researchers and educators around the world. Moreover, some examples of I/H structures have already been submitted to the PDB resource, but currently remain unprocessed owing to the lack of appropriate infrastructure.

To address this challenge, the wwPDB organization sponsored an I/H Methods Task Force workshop in October 2014 at the EMBL-European Bioinformatics Institute (EMBL-EBI). Participants included 38 researchers from Europe, Asia, and North America with expertise in experimental protein structure characterization, computational modeling, visualization, and data archiving (https://www.wwpdb.org/task/hybrid). These experimental and computational scientists, together with wwPDB representatives, contributed to the workshop. Three breakout groups discussed challenges involved in managing I/H structural models and their supporting experimental data. Five consensus recommendations emerging from the meeting were summarized in a White Paper published in *Structure* (Sali et al., 2015), the most important of these being the urgent need for creation of data standards and establishment of a federated system of data resources to standardize representation, validation, archiving, and dissemination of I/H structural models and supporting data.

Two subgroups have been created within the wwPDB I/H Methods Task Force to begin implementing these recommendations (https://pdb-dev.wwpdb.org/contributors/). First is the Model Validation Subgroup, led by Andrej Sali (UCSF) and Torsten Schwede (SIB), which addresses development of methods for I/H model representation, validation, and visualization; and second is the Federation Subgroup, led by Jill Trewhella (Sydney/Utah) and Helen Berman (Rutgers/RCSB PDB), which focuses on building a network of resource partners that can exchange I/H related experimental and structural information in a concerted, systematic manner.

To create data standards for archiving I/H models, we have developed an I/H methods data dictionary that defines the data specifications for archiving I/H structural models, associated spatial restraints, and structural model computation protocols (Berman et al., 2016; Vallat et al., 2016a, 2016b). This dictionary is a modular extension of the PDBx/mmCIF dictionary (Fitzgerald et al., 2005), currently used by the wwPDB for archiving atomic structures of macromolecules. The new dictionary extension addresses two fundamental requirements of I/H methods: (1) descriptions of spatial restraints derived from a variety of experimental and analytical techniques, and (2) definitions of multi-scale, multi-state, and time-ordered ensembles of macromolecular assemblies. Generic representation of multi-scale models

together with descriptions of experimentally determined spatial restraints is an important feature of the new I/H methods dictionary. The first version of the dictionary has been developed using specific models computed with the open-source Integrative Modeling Platform (IMP) software (Russel et al., 2012). The I/H methods dictionary and associated documentation can be found at https://github.com/ihmwg/IHM-dictionary/.

We have also developed a prototype web-based deposition system (https://pdb-dev.wwpdb.org) for I/H models that complies with the current specifications defined in the I/H methods dictionary. Three test cases covering a variety of features of the IMP software have been deposited using this system, including the 7-piece Nup84 sub-complex of the nuclear pore complex from yeast (Shi et al., 2014), the yeast exosome complex, and the yeast mediator complex. These structures can be downloaded and visualized using the ChimeraX visualization suite (Ferrin et al., 2017), which provides built-in support for the new I/H methods data representation.

To maximize the applicability of the I/H methods data dictionary and the prototype deposition system, we now ask the structural biology community for feedback and to provide additional test cases as input to PDB-Dev. To assist in this endeavor or for other general inquiries, please contact the PDB-Dev development team at pdb-dev@mail.wwpdb.org.

Acknowledgments

The wwPDB Consortium and the I/H Methods Task Force thank all community participants who have contributed time and expertise to this project, and wish to acknowledge specifically the efforts of J. Meiler, G. Hummer, F. DiMaio, E. Tajkhorshid, T. Ferrin, B. Vallat, M. Gabanyi, B. Webb, J. Westbrook, T. Goddard, K. Baskaran, J. Stone, R. McGreevy, R. Moretti, O. Smart, A. Gutmanas, J. Berrisford, J. Haas, J. Koefinger, H. Suzuki, G. Bekker, R. Aebersold, N. Kirby, P. Adams, G. Montelione, W. Chiu, G. Phillips, B. Carragher, J. Rappsilber, C. Seidel, G. Vuister, D. Schriemer, P. Griffin, F. Alber, M. Marti-Renom, and D. Svergun. Creation of PBB-Dev was funded by NSF EAGER grant DBI-1519158.

References

Berman HM, Westbrook J, Vallat B, Webb B, Sali A. A Data Dictionary for Archiving Integrative/Hybrid Models. 66th Annual Meeting of the American Crystallographic Association. 2016 85-SA.

Ferrin, T., Huang, C., Peeterson, E., Goddard, T., Couch, G., Meng, E., Morris, S. UCSF ChimeraX. 2017. https://www.rbvi.ucsf.edu/chimerax/

Fitzgerald, PMD., Westbrook, JD., Bourne, PE., McMahon, B., Watenpaugh, KD., Berman, HM. 4.5 Macromolecular dictionary (mmCIF). In International Tables for Crystallography G. In: Hall, SR., McMahon, B., editors. Definition and Exchange of Crystallographic Data. Springer; 2005. p. 295-443.

Russel D, Lasker K, Webb B, Velazquez-Muriel J, Tjioe E, Schneidman-Duhovny D, Peterson B, Sali A. Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. PLoS Biol. 2012; 10:e1001244. [PubMed: 22272186]

Sali A, Berman HM, Schwede T, Trewhella J, Kleywegt G, Burley SK, Markley J, Nakamura H, Adams P, Bonvin AM, et al. Outcome of the First wwPDB hybrid/integrative methods task force workshop. Structure. 2015; 23:1156–1167. [PubMed: 26095030]

Shi Y, Fernandez-Martinez J, Tjioe E, Pellarin R, Kim SJ, Williams R, Schneidman-Duhovny D, Sali A, Rout MP, Chait BT. Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. Mol. Cell Proteomics. 2014; 13:2927–2943. [PubMed: 25161197]

Vallat, B., Webb, B., Westbrook, J., Sali, A., Berman, H. Integrative/Hybrid Methods PDBx/mmCIF dictionary extension. 2016a. https://github.com/ihmwg/IHM-dictionary

Vallat, B., Webb, B., Westbrook, J., Sali, A., Berman, H. Integrative/hybrid methods PDBx/mmCIF dictionary extension documentation. 2016b. https://github.com/ihmwg/IHM-dictionary/tree/master/dictionary_documentation

Vallat, B., Webb, B., Westbrook, J., Sali, A., Berman, HM. The PDB-Dev prototype deposition and archiving system. 2016c. https://pdb-dev.wwpdb.org/

Ward AB, Sali A, Wilson IA. Biochemistry. Integrative structural biology. Science. 2013; 339:913–915. [PubMed: 23430643]