

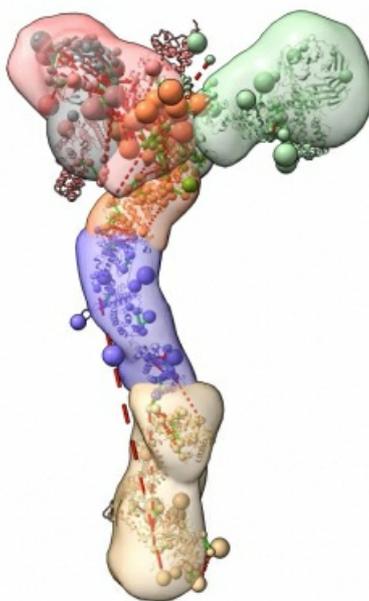
A Data Dictionary For Archiving Integrative/Hybrid Models

Brinda Vallat¹, Benjamin Webb², John Westbrook¹, Andrej Sali², Helen Berman¹

¹RCSB Protein Data Bank, Rutgers, The State University Of New Jersey, Piscataway, United States, ²Department of Bioengineering and Therapeutic Sciences, Department of Pharmaceutical Chemistry and California Institute for Quantitative Biosciences, University of California at San Francisco, San Francisco, United States
E-mail: brinda.vallat@rcsb.org

Structural characterization of complex macromolecular assemblies is increasingly being carried out using integrative and hybrid (I/H) methods. Traditional structure determination methods such as X-ray crystallography and NMR spectroscopy are generally insufficient for tackling such complex assemblies. Recently, methods have been developed that combine spatial restraints derived from a variety of complementary experimental techniques, including cryo-electron microscopy, small angle scattering, chemical crosslinking, mass spectrometry and other proteomics and bioinformatics methods. A feature of integrative modeling is that it allows for multi-scale, multi-state, and time-ordered ensembles, which are very different from the mono-scale atomistic models currently archived in the Protein Data Bank [1]. Our goal is to facilitate the archiving of I/H models so that they can be available to the broader biological research community. We have created a data dictionary that captures the details of I/H models including a variety of experimentally-derived spatial restraints and the modeling of multi-scale, multi-state, time-ordered ensembles. The dictionary is an extension of the PDBx/mmCIF dictionary [2] used by the Protein Data Bank to archive macromolecular structures. The dictionary and supporting documentation are publicly accessible [3]. Creating an archive for I/H models is vital to the evolving structural biology community and the data dictionary is a critical step in this direction. This work is supported by NSF EAGER award number 1519158.

1. Berman, H.M. et al. (2000). *Nucleic Acids Res.* 28(1), 235-42.
2. PDBx/mmCIF Dictionary Resources (2013). Available online, <http://mmcif.wwpdb.org/>.
3. Integrative/Hybrid Methods Dictionary (2016). Available online, <https://github.com/ihmwg/IHM-dictionary>.



Keywords: [Integrative/Hybrid Methods](#), [PDBx/mmCIF](#), [Protein Data Bank](#)