



## Announcing the worldwide Protein Data Bank

In recognition of the growing international and interdisciplinary nature of structural biology, three organizations have formed a collaboration to oversee the newly formed worldwide Protein Data Bank (wwPDB; <http://www.wwpdb.org/>). The Research Collaboratory for Structural Bioinformatics (RCSB), the Macromolecular Structure Database (MSD) at the European Bioinformatics Institute (EBI) and the Protein Data Bank Japan (PDBj) at the Institute for Protein Research in Osaka University will serve as custodians of the wwPDB, with the goal of maintaining a single archive of macromolecular structural data that is freely and publicly available to the global community.

The wwPDB represents a milestone in the evolution of the Protein Data Bank (PDB; <http://www.pdb.org/>)<sup>1,2</sup>, which was established in 1971 at Brookhaven National Laboratory as the sole international repository for three-dimensional structure data of biological macromolecules. Since July 1, 1999, the PDB has been managed by three member institutions of the RCSB: Rutgers, The State University of New Jersey; the San Diego Supercomputer Center at the University of California, San Diego; and the Center for Advanced Research in Biotechnology of the National Institute of Standards and Technology.

The wwPDB recognizes the importance of providing equal access to the database—both in terms of depositing and retrieving data—from different regions of the world. Therefore, the wwPDB members will continue to serve as deposition, data processing, and distribution sites. Deposition procedures will not be altered by the formation of the wwPDB; data can still be deposited using ADIT at the RCSB and PDBj or by using AutoDep at the EBI.

To ensure the consistency of PDB data, all entries will be validated and annotated following a common set of criteria. All processed data will be sent to the RCSB, which distributes the data worldwide. All format docu-

mentation will be kept publicly available and the distribution sites will mirror the PDB archive using identical contents and subdirectory structure. However, each member of the wwPDB will be able to develop its own web site, with a unique view of the primary data, providing a variety of tools and resources for the global community.

An Advisory Board consisting of appointees from the wwPDB, the International Union of Crystallography and the International Council on Magnetic Resonance in Biological Systems will provide guidance through annual meetings with the wwPDB consortium. This board is responsible for reviewing and determining policy as well as providing a forum for resolving issues related to the wwPDB. Specific details about the Advisory Board can be found in the wwPDB charter, available on the wwPDB web site.

The RCSB is the 'archive keeper' of wwPDB. It has sole write access to the PDB archive and control over directory structure and contents, as well as responsibility for distributing new PDB identifiers to all deposition sites. The PDB archive is a collection of flat files in the legacy PDB file format<sup>3</sup> and in the mmCIF<sup>4</sup> format that follows the PDB exchange dictionary (<http://deposit.pdb.org/mmcif/>). This dictionary describes the syntax and semantics of PDB data that are processed and exchanged during the process of data annotation. It was designed to provide consistency in data produced in structure laboratories, processed by the wwPDB members and used in bioinformatics applications. The PDB archive does not include the websites, browsers, software and database query engines developed by researchers worldwide.

The members of the wwPDB will jointly agree to any modifications or extensions to the PDB exchange dictionary. As data technology progresses, other data formats (such as XML) and delivery methods may be included in the official PDB archive if all the wwPDB members concur on the alteration. Any new formats will follow the naming and

description conventions of the PDB exchange dictionary. In addition, the legacy PDB format would not be modified unless there is a compelling reason for a change. Should such a situation occur, all three wwPDB members would have to agree on the changes and give the structural biology community 90 days advance notice.

The creation of the wwPDB formalizes the international character of the PDB and ensures that the archive remains single and uniform. It provides a mechanism to ensure consistent data for software developers and users worldwide. We hope that this will encourage individual creativity in developing tools for presenting structural data, which could benefit the scientific research community in general.

### ACKNOWLEDGMENTS

The RCSB PDB is supported by funds from the National Science Foundation, the Department of Energy, and the National Institutes of Health. The MSD-EBI is supported by funds from the Wellcome Trust, the European Union (TEMBLOR, NMRQUAL, SPINE, AUTOSTRUCT, and IIMS awards), CCP4, the Biotechnology and Biological Sciences Research Council (UK), the Medical Research Council (UK), and the European Molecular Biology Laboratory. PDBj is supported by grant-in-aid from the Institute for Bioinformatics Research and Development, Japan Science and Technology Corporation (BIRD-JST), and the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

**Helen Berman<sup>1</sup>, Kim Henrick<sup>2</sup> & Haruki Nakamura<sup>3</sup>**

<sup>1</sup>RCSB, Piscataway, New Jersey; Rockville, Maryland and La Jolla, California, USA.

<sup>2</sup>MSD-EBI, Hinxton, UK. <sup>3</sup>PDBj, Institute for Protein Research, Osaka University, Osaka, Japan.

e-mail: [berman@rcsb.rutgers.edu](mailto:berman@rcsb.rutgers.edu)

1. Berman, H.M. *et al.* *Nucleic Acids Res.* **28**, 235–242 (2000).
2. Bernstein, F.C. *et al.* *J. Mol. Biol.* **112**, 535–542 (1977).
3. Callaway, J. *et al.* Protein Data Bank Contents Guide: Atomic coordinate entry format description. (Brookhaven National Laboratory, 1996).
4. Bourne, P.E., Berman, H.M., Watenpaugh, K., Westbrook, J.D. & Fitzgerald, P.M.D. *Methods Enzymol.* **277**, 571–590 (1997).