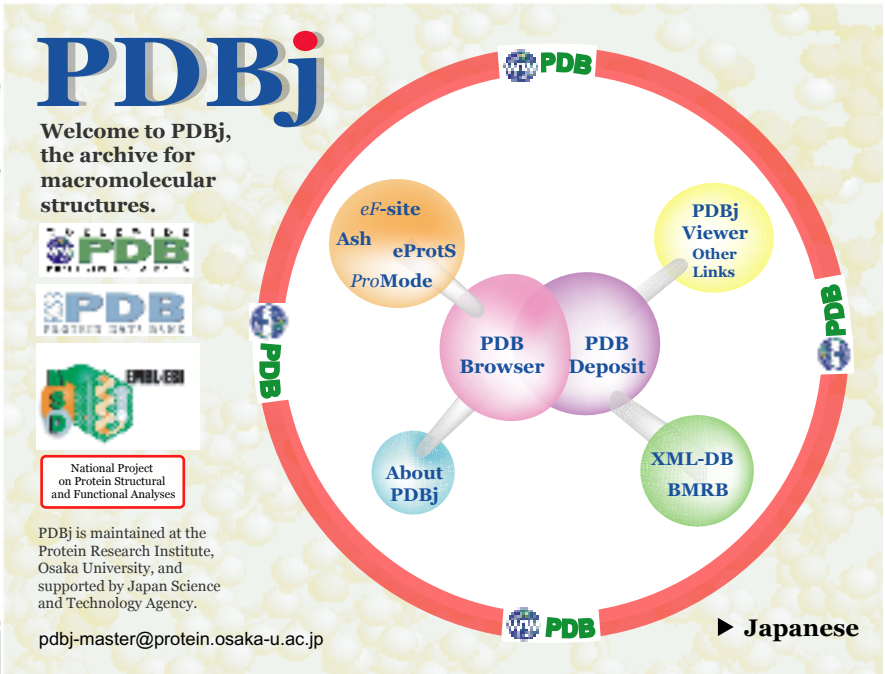
PDBj

Protein Data Bank Japan

News Letter Vol. 3

December 2003

http://www.pdbj.org/



PDBj
Protein Data Bank Japan
Protein Data Bank Japan http://www.pdbj.org/
News Letter Vol. 3

1. RCSB, EBI, and PDBj collaborate to form worldwide Protein Data Bank (wwPDB) to manage PDB archive

The Protein Data Bank (PDB; http://www.pdb.org/) was founded 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 the three member institutions of the Research Collaboratory for Structural Bioinformatics (RSCB).

PDB data are deposited and used by the global community. At present, data are processed by three groups: the PDB managed by the 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. All processed data are sent to the RCSB who oversee the process and distribute the data worldwide.

In recognition of the international nature of the resource, the RCSB, MSD-EBI, and PDBj have formed the worldwide Protein Data Bank (wwPDB; http://www.wwpdb.org/). The mission of the wwPDB is to maintain a single archive of macromolecular structural data that is freely and publicly available to the global community. An international advisory board will be formed to support this organization.

The wwPDB acknowledges the importance of international equality in the ability to deposit and access this data. This equality reflects the international and interdisciplinary nature of scientific research.

The PDB Archive

The "PDB archive" is the collection of flat files in the legacy PDB file in the mmCIF format that follows the PDB exchange dictionary. This dictionary describes the syntax and semantics of PDB data that are processed and exchanged during the process of data annotation. It is designed to provide consistency in the data that are processed by different people in different organizations. The term "PDB archive" does not refer to the Web sites, browsers, software, and database query engines that will continue to be developed at various sites.

The definitive PDB exchange dictionary is kept at http://deposit.pdb.org/mmcif/. All changes or modifications to the PDB exchange dictionary will be agreed upon by the members of the wwPDB (the RCSB, MSD-EBI, and PDBj) and will be reflected in the version number of the dictionary. The legacy PDB format will not be modified unless there is a compelling reason for a change. If this situation occurs, there must be agreement by all members of wwPDB and 90 days notice would be made to the community.

In the future, other data formats (such as XML) and delivery methods may be included in the official PDB archive after full discussion and agreement among members. New formats must follow the naming and description conventions of the PDB exchange dictionary.

The wwPDB members and their roles

As founding members, all three organizations serve as deposition, data processing and distribution sites. All format documentation will be kept publicly available and the distribution sites will mirror the PDB archive using identical contents and subdirectory structure.

Each wwPDB member will develop its own Web sites and views of the same primary data. Any products will be marked with that member's logo.

An Advisory Board will meet with wwPDB members yearly. Each member of wwPDB will appoint two members to the Advisory Board. wwPDB will ask the International Union of Crystallography (IUCr) and the International Council on Magnetic Resonance in Biological Systems (ICMRBS) each to appoint one representative to the Advisory Board.

The three wwPDB members agree that the RCSB will be the 'archive keeper' and will have sole write access to the PDB archive and control over the directory structure and contents. The archive keeper distributes new PDB identifiers to all deposition sites. The use of identifiers by members of the wwPDB resembling PDB 4-letter codes for other purposes is prohibited. The four-letter PDB identifier currently consists of a number (0-9) followed by 3 letters or numbers.

The archival PDB files will be distributed with the reserved conventional names, in the form pdb<entry_id>.ent where <entry_id> is a PDB 4-letter code, e.g. pdb1abc.ent, for PDB format entries; r<entry_id>sf.ent, e.g. r1abcsf.ent for X-ray experimental data; <entry_id>.mr, e.g. 1abc.mr for NMR experimental/constraints, <entry_id>.cif, e.g. 1abc.cif for mmCIF format entries; and <entry_id>.xml e.g. 1abc.xml for canonical XML format entries.
The formal wwPDB charter is available from http://www.wwpdb.org.

Impact / Significance for PDB User Community

wwPDB formalizes the international character of the PDB and ensures that the archive will remain single and uniform. Formation of the wwPDB will be transparent to depositors. PDB depositions will continue to be accepted at all three sites and will not appear any different than it does now. Data will be equally well validated and annotated at all wwPDB sites.

By providing a formal mechanism for standardizing the presentation of PDB data, software developers and users of the data will be assured of consistent data. At the same time, it is hoped that this agreement will allow for individual creativity in how the data are presented and made available to the community.

Acknowledgement

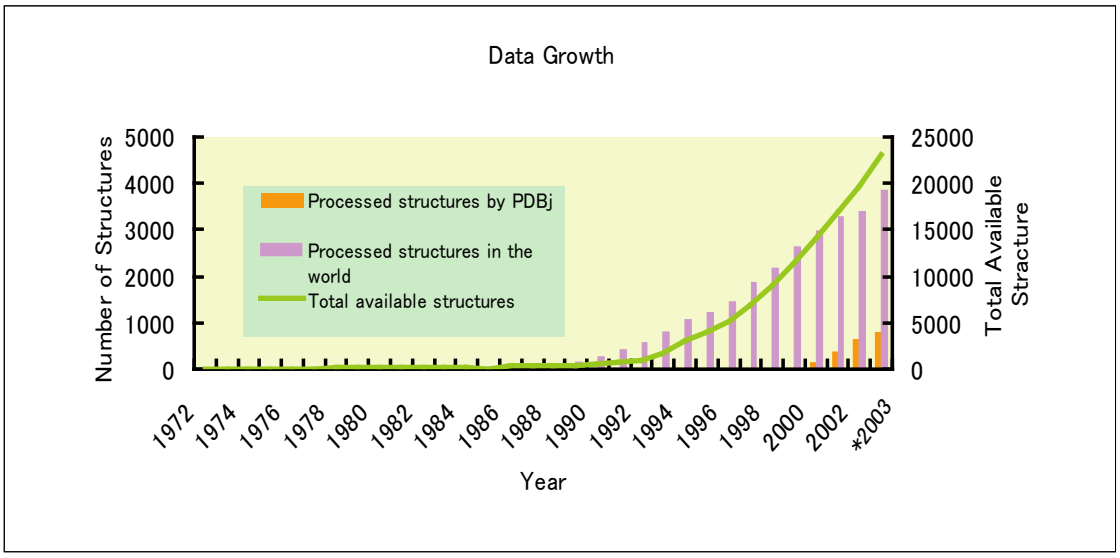
The RCSB PDB is supported by funds from the National Science Foundation, the Department of Energy, and two units of the National Institute of Health: the National Institute of General Medical Sciences and the National Library of Medicine.

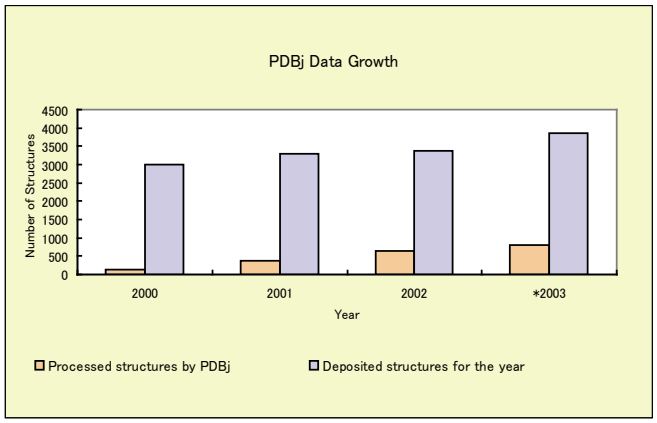
The MSD-EBI is supported by funds from the Wellcome Trust, the European Union (TEMBLOR, NMRQUAL, SPINE, AUTOSTRUCT, and IIMS awards), CCP4, the Biotechnology and Biological Sciences Research Council (UK), the Medical Research Council (UK), and the European Molecular Biology Laboratory.

PDBj is supported by grant-in-aid from the Institute for Bioinformatics Research and Development, Japan Science and Technology Agency (BIRD-JST), and Ministry for Education, Culture, Sports, Science and Technology(MEXT).

2. PDB Depositions Statistics

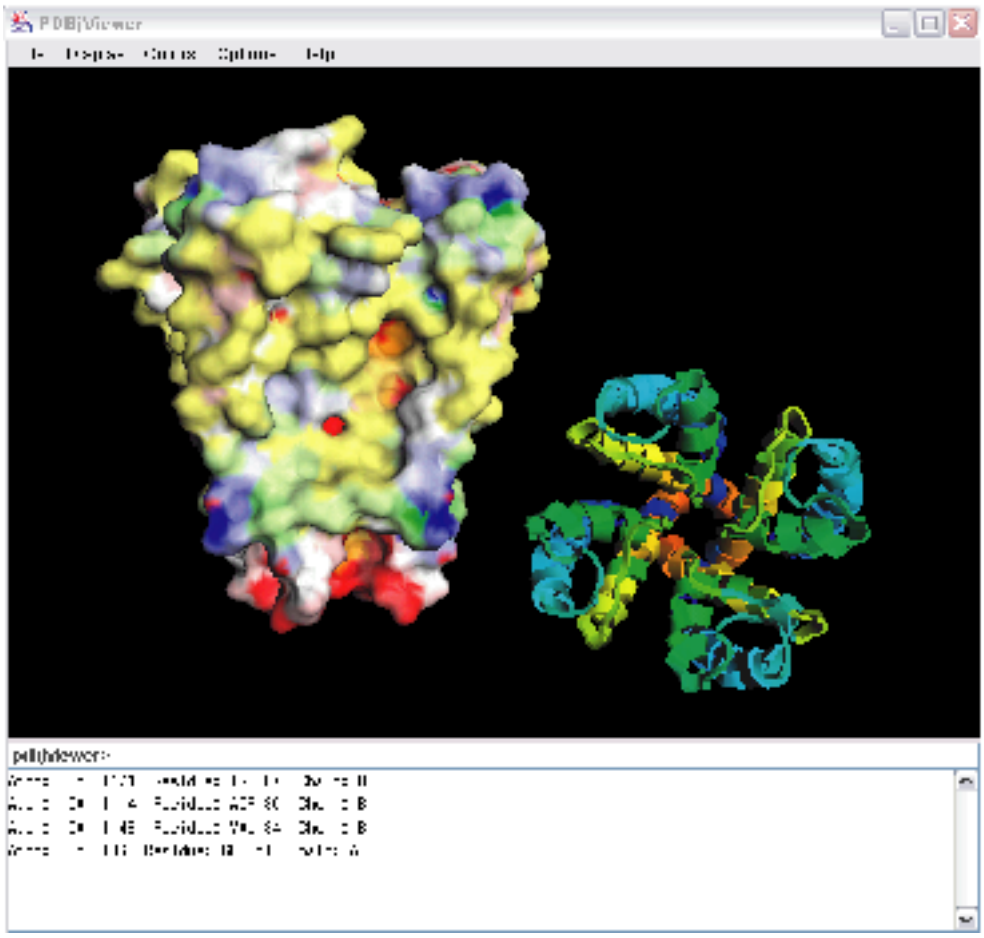
＊Last updated: Nov. 5, 2003





3. A New Molecular Graphics Viewer: PDBjViewer

We developed a new molecular graphics viewer, 'PDBjViewer', to visualize protein molecular structures and surfaces, suitable for using the XML description of PDB data.
The features of PDBjViewer are as follows:



PDBjViewer Screen, displaying the molecular structure and surface of a potassium ion channel, first revealed by R. Mackinnon who won the Novel Prize in 2003.

1) Several molecular models: wire-frame, stick, ball-and-stick, CPK, ribbon, and cartoon, are displayed with the flexible color assignment.

2) Protein molecular structures and surfaces are manipulated in real-time by mouse controls, for rotation, translation, and zooming.

3) It is controlled by pull-down menu commands or the command-line, following the popular 'RasMol' graphics program.

4) It has been developed both as a stand-alone and an applet program.

5) It runs using the Java Runtime Environment. JRE 1.4 (or later version) and Java3D API (OpenGL 1.2.1 or later version) are required.

6) It reads both conventional PDB format data and XML format data. The latter has been tentatively proposed by the collaboration work of the RCSB and PDBj groups for files without atom coordinates[1], for the external atom coordinates files[2] and their schemas[3]. Information other than atomic coordinates described by XML can also be visualized, such as secondary structures.

7) A polygon file written in XML can be read and displayed, mainly for visualizing the protein molecular surfaces, but any other colored polygons that are generated by the user can be displayed as well. More than one polygon file can be opened at the same time and superposed on a molecule image.

8) Two image windows are simultaneously opened. For example, a molecular structure is shown in the left window, and the molecular surface is in the right window.

PDBjViewer was developed by Kengo Kinoshita (Graduate School of Integrated Science, Yokohama City University) and Haruki Nakamura (Institute for Protein Research, Osaka University), as one of the activities of Protein Data Bank Japan (PDBj), supported by grant-in-aid from Institute for Bioinformatics Research and Development, Japan Science and Technology Agency (BIRD-JST). It is now applied in the protein molecular surface database (eF-site[4]) and the encyclopedia of proteins structures (eProtS[5]).

PDBjViewer, Version 1.0 and its user manual are freely distributed from the Web page: http://www.pdbj.org/PDBjViewer/. The source code will also be made available, when the version is established.

1. ftp://beta.rcsb.org/pub/pdb/uniformity/data/XML/all-noatom/
2. ftp://beta.rcsb.org/pub/pdb/uniformity/data/XML/all-extatom/
3. http://deposit.pdb.org/pdbML/pdbx-v0.905.xsd
   http://deposit.pdb.org/pdbML/pdbx-v0.904-alt.xsd)
4. http://www.pdbj.org/eF-site/
5. http://www.pdbj.org/eProtS/

PDBj
Research Center for Structural and Functional Proteomics,
Institute for Protein Research (IPR), Osaka University
3-2 Yamadaoka, Suita, Osaka 565-0871, Japan
TEL (PDBj office): +81-(0)6-6879-4311
TEL (PDBj deposition office): +81-(0)6-6879-8638
FAX: +81-(0)6-6879-8636
E-mail: pdbj-master@protein.osaka-u.ac.jp