PDBj

Protein Data Bank Japan

News Letter Vol. 4

July 2004

http://www.pdbj.org/

1. News

International Workshop

The International Workshop: Integrated Databases and DataGrid for Structural Biology and Molecular Biology, organized by JST (Japan Science and Technology Agency) and the Institute for Protein Research, Osaka University, was held on March 1st and 2nd, 2004 at Osaka University.

Snapshot of the Workshop

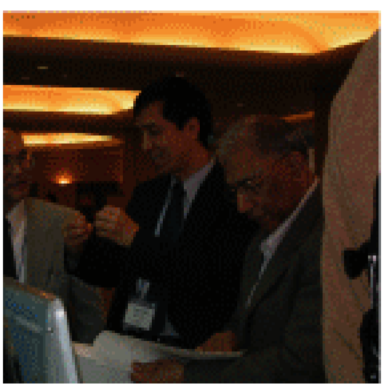
The future directions of integrated databases and datagrid technology were discussed by invited speakers from Japan and abroad: Haruki Nakamura (Osaka Univ.), Hideo Matsuda (Osaka Univ.), Philip Bourne (UCSD), Satoru Miyazaki (DDBJ), Hideaki Sugawara (DDBJ), Peter Rice (EBI), Kei Yura (JAERI), Tin Wee Tan (National Univ. Singapore), Tadashi Imanishi (JBIRC) John Westbrook (RCSB), Nobutoshi Ito (Univ. Tokyo Med. Dent.), Philip McNeil (EBI), Kengo Kinoshita (Osaka Univ.), Masaharu Isoyama (PRF), Andreas Heger (Helsinki Univ.), Daron Standley (JST-BIRD) .

For understanding and effectively utilizing genome information, it is necessary to combine a wide range of life science information. This, in turn, requires integrating high quality and well-managed databases. With developing networks, it has been possible recently to decentralize the environment using DataGrid. New direction for API and Workflows for those databases were another of the topics of this Workshop.

PRICPS 2004

The 1st Pacific-Rim International Conference on Protein Science and the 4th annual meeting of the Protein Science Society of Japan (PRICPS) was held April 14th- 18th, 2004, at Pacifico Yokohama, Japan. Our activities were introduced at the exhibition booth, where both demonstrations and presentations were given.

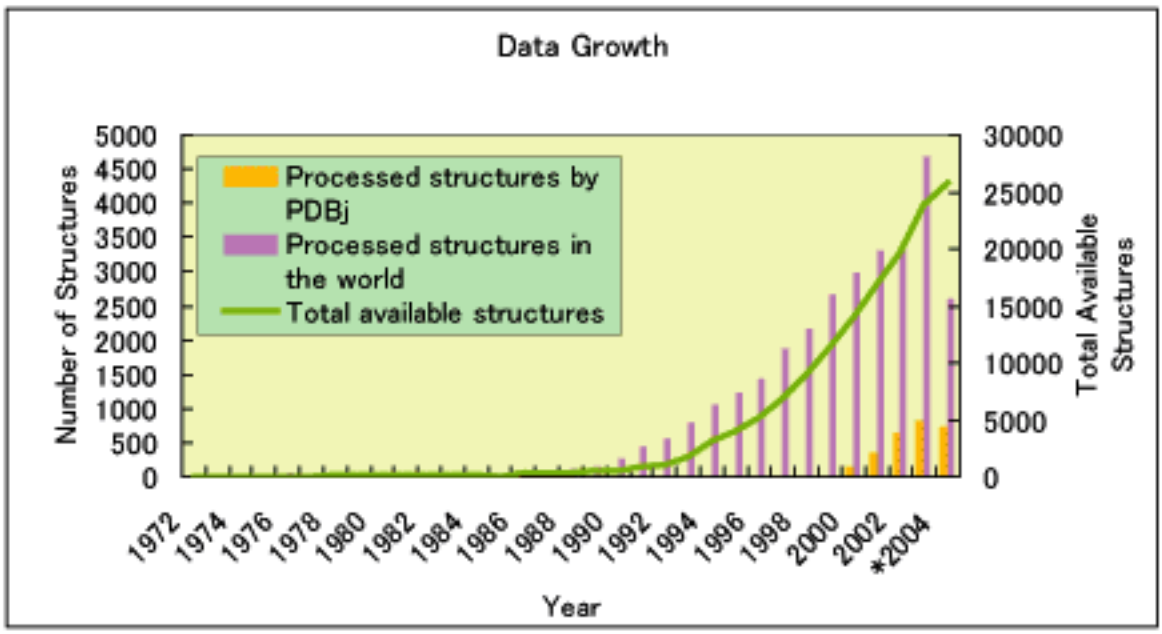
Exhibition booth at PRICPS 2004



Prof. Wako demonstrates ProMode at the PDBj booth.

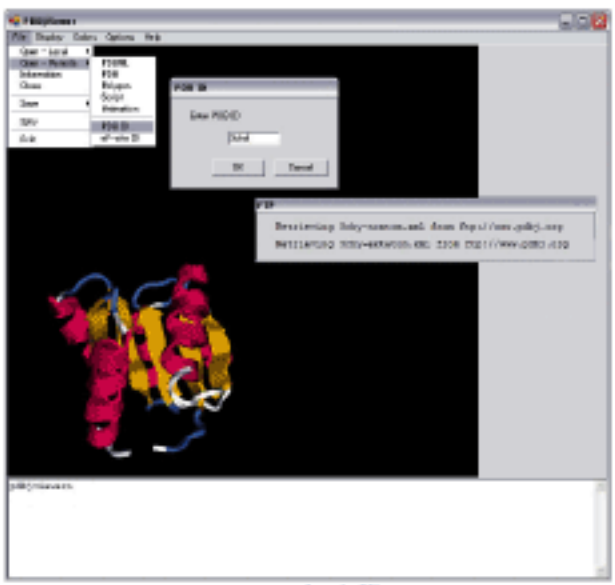
2. PDB Deposition Statistics

* Last updated: June 30, 2004



3. PDBj Viewer *(jV)*

PDBj Viewer was first developed for visualizing protein molecular structures and surfaces interactively on the eF-site database, which is a database of electrostatic surfaces for functional sites. The viewer has since been extended to be used both as a stand-alone program and as an applet. Because "Chime" and "RasMol" are now popular applets and stand-alone programs, respectively, our viewer was designed to have compatibility with "Chime" and "RasMol" to some extent. The display function of each molecular surface, commonly managed as a polygon, is a unique feature.



example of jV2

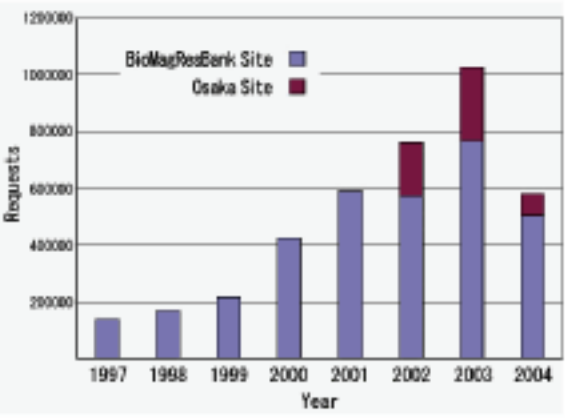
Now, PDBj Viewer has an applet version "*jV1*" and a stand-alone version "*jV2*," that has a limitation and operates only under Windows. The latest viewer, *jV2*, is implemented not only to display protein surfaces ("RasMol" does not have this function), but also to have various other functions, including the simultaneous display of multiple molecules, animation, directly parsing pdbML files, and data downloading through the PDBj web site, which is necessary for use at PDBj and eF-site.

At the end of this autumn, a new version, *jV3*, is going to be released including MAC OS-X and Linux, and can be used both as a stand-alone and an applet.

4. BMRB

BMRB (BioMagResBank) is a database to collect NMR experimental results, such as chemical shift values of individual atoms in proteins and nucleic acids, developed and managed by Prof. John L. Markley at the University of Wisconsin, Madison. NMR researchers deposit the data obtained from their experiments, and retrieve registered data via the BMRB website.

In March, 2002, PDBj opened a Mirror site (http://brmb.protein.osaka-u.ac.jp), collaborating with Prof. John L. Markley. The total number of data entries in the BMRB amounted to 3072 in July 2004, and the number of accessions increases every year. Recently, a new deposition interface was developed at BMRB after the data format was renewed.



Statistics for the number of processing at the BMRB web server.

5. xPSSS

As written in the previous newsletters, PDBML, a new XML format for PDB, has been developed in collaboration between PDBj and RCSB and the entire content of the current database is now available in this format as well as in the conventional "PDB format". To make use of the XML-specific features of this format, we have built a native XML-DB and created a web interface called xPSSS (XML-based protein structure search service, http: //www.pdbj.org /xpsss) to search the new database. In addition to simple keyword searches, various data items, such as citation journal, resolution and deposition date, can be used as search keys through the web interface.

The xPSSS frontpage

The most unique feature of xPSSS is, however, the implementation of XPath search, which makes use of the hierarchical structure of pdbML. It allows users to perform a variety of complicated searches on any items in the database simultaneously. For example, one can search for entries with more than two polypeptide chain heavier than 30kDa and with the function of DNA-binding that actually have the coordinates of the DNA molecules. XPath also enables users to control the format of the output of their searches. One can search for beta-turns and get the amino-acid sequence of the turn as output. Although the examples shown so far need a browser to access the service, SOAP (Simple Object Access Protocol) interface for direct access is also provided to remove this requirement. That is, users can access the database from their programs running on their local machines through the network. This makes large scale searches, such as systematic analysis of the entire database, feasible which would have been impractical using

a browser manually. It also opens up the possibility of making organic links to other databases with a SOAP interface to perform inter-database searches, key to the notion of a "DataGrid". A drawback of this XPath service is that users have to understand its grammar and the structure of pdbML to gain much out of it. We are, thus, planning to create some sort of GUI to reduce users' own efforts in the future. In the meantime, the number of examples of XPath search will be increased and any suggestions and comments are very much appreciated.

A new feature of xPSSS, the Functional Details page, has been prepared so as to display functional information of the individual protein, Biological process, Biochemical function, and Cellular location, with the residual sites if provided. The information sources are original literature and several databases: the original PDB data as the SITE information, GO (GeneOntology, http://www.geneontology.org/), eF-site (http://www.pdbj.org/eF-site/), Swiss-Prot (http://us.expasy.org/), and CATRES (Catalytic Residue Dataset, http://www.ebi.ac.uk/thornton-srv/databases/CATRES/). The importance of function information for protein 3D structures is now recognized as critical for biological science, genome science, and proteomics.

6. PDBj Staff

Director
Nakamura, Haruki (Prof., IPR, Osaka Univ.)

Group for PDB Database Curation
Kusunoki, Masami, Dr.
Kosada, Takashi
Igarashi, Reiko
Kengaku, yumiko
Saeki, Shinobu
Ikegawa, Yasuyo

Group for NMR Database
Akutsu, Hideo, Prof.
Matsuki, Yoh, Dr.
Nakatani, Eiichi

Secretary
Kamada, Chisa

Group for Development of Secondary Database and Analysis Systems
Ohkawa, Takenao, Dr.
Kinoshita, Kengo, Dr.
Toh, Hiroyuki, Prof.
Wako, Hiroshi, Prof.
Kaneta, Yoshikazu

Group for Development of PDBj database
Ito, Nobutoshi, Prof.
Paehler, Arno, Dr.
Standley, Daron M., Dr.
Yamashita, Reiko
Yoshihara, Atsuro
Shimizu, Yukiko
Takahashi, Aki

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: pdbjADMIN@protein.osaka-u.ac.jp

PDBj PDBj PDBj PDBj
PDBj PDBj PDBj PDBj
PDBj PDBj PDBj PDBj