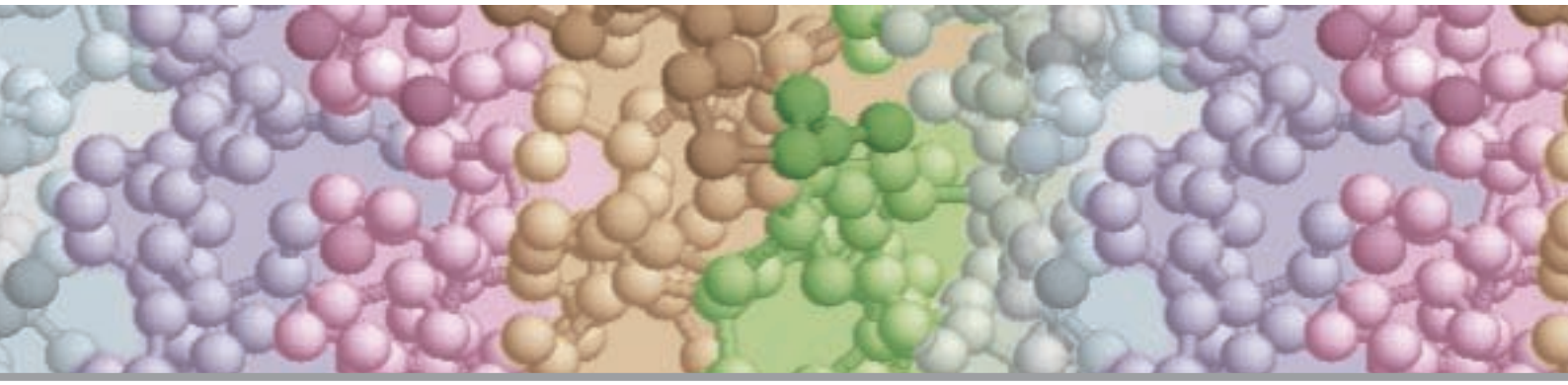
PDBj
Protein Data Bank Japan

http://www.pdbj.org/

News Letter Vol. 8, No. 3

March 2007



PDBj is maintained at the Protein Research Institute, Osaka University, and supported by Japan Science and Technology Agency.

News

The Third wwPDBAC meeting held at Tokyo on October 27, 2006.

On October 27, 2006, the Third wwPDB Advisory Committee (wwPDBAC) meeting was held in Tokyo, Japan, and was chaired by Dr. Stephen K. Burley (Structural GenomiX, Inc.). The other participants were: (1) PDB Site Representatives: Prof. Brian Matthews (RCSB, University of Oregon), Prof. Neil Isaacs (MSD, University of Glasgow), Prof. Robert Kaptein (MSD, Utrecht University), Dr. Kei Yura (PDBj, JAEA), Prof. Soichi Wakatsuki (PDBj, Photon Factory), Prof. Masatsune Kainosho (BMRB, Tokyo Metropolitan University), Prof. Hideo Akutsu (BMRB, Osaka University) (2) Community Stakeholder Representatives: Prof. Edward N. Baker (IUCr, University of Auckland), Dr. R. Andrew Byrd (ICMRBS, National Cancer Institute), Prof. Jose-Maria Carazo (Macromolecular EM, CNB), (3) PDB Site Leaders: Prof. Helen M. Berman (RCSB, Rutgers University), Dr. Kim Henrick (MSD, EBI), Prof. Haruki Nakamura (PDBj, Osaka University), Prof. John L. Markley (University of Wisconsin) and (4) Funding Agency Representatives: Mr. Toshiyuki Koike (JST).

At the beginning, Prof. Haruki Nakamura, Head of PDBj, welcomed the wwPDBAC members, and the wwPDBAC mission statement was confirmed: To help ensure that the Protein Data Bank is maintained for the public good as a secure, single, global archive for experimental structural biology data that is freely accessible in perpetuity.

Then, Prof. Helen M. Berman made an overview of recent wwPDB progress: Continued growth of archive, website updates, publications and presentations, time stamped archive, wwPDB team building with actual exchange visits, phone conferences and VTC, annotation documents, remediation, and BMRB (BioMagResBank) membership to the wwPDB.

Regarding the last issue, Prof. John L. Markley introduced the BMRB, which joined the wwPDB in 2006 and acts as the worldwide archive for biomolecular NMR data. NMR data related to structures is cross-referenced to the PDB, and PDBj mirrors BMRB and supports external BMRB depositions. BMRB and PDB are working closely to capture and annotate NMR data associated with deposited coordinate sets by having a “one stop” deposition system for NMR structures.

A new wwPDB policy regarding Theoretical Models was introduced by Prof. Haruki Nakamura. Based on a discussion among the modelers, structural genomicists, and electron microscopists at a workshop held on November 19-20, 2005, the following recommendation was made and published (Structure (2006) 14, 1211-1217): *PDB depositions will be restricted to atomic coordinates that are substantially determined by experimental measurements on specimens containing biological macromolecules.* The wwPDB has implemented this recommendation on following schedule: On August 15, 2006, the policy was announced with a 60 day period for review. Between August 15th-October 15th, 2006, a transition was executed: "All existing un-processed theoretical model entries as well as entries deposited during this time were not validated or processed. Entries were released as-is without author review or corrections". Finally, on October 15, 2006, theoretical model depositions were no longer accepted.

Other issues were also discussed. In particular, new remediation changes and rollout plans were introduced by the wwPDB members and discussed in depth. One big change is the requirement that remediated data conforms to IUPAC nomenclature standards, especially for hydrogen atoms. This advance is critical for the PDB to fully serve the data archiving needs of the NMR community. There was considerable discussion regarding the impact of remediated PDB releases on various super users and software resources. In addition, it was agreed among all the wwPDBAC members that the four character PDB ID Code will continue to be used but that the first character will be expanded from nine possibilities (1 to 9) to thirty five (1 to 9 and A to Z), which will effectively resolve the problem (1,632,960 unique combinations) without a need for major software adjustments.

At the end of the wwPDBAC meeting, it was announced that the Forth wwPDBAC meeting will be held in September, 2007 at Princeton, NJ in the US. Details of the wwPDBAC meeting and recommendations from the Advisory Committee are provided on the wwPDB web site at http://www.wwpdb.org/news.html.

AsCa' 06: Joint Conference of the Asian Crystallographic Association and the Crystallographic Society of Japan

Joint Conference of the Asian Crystallographic Association and the Crystallographic Society of Japan was held in Tsukuba, Japan, from November 20th to 23rd, 2006. The subject of this conference was crystal growth studies, including all aspects of materials science related to crystallography, such as nano and biological materials, and encompasses X-ray and neutron scattering/diffraction, electron microscopy, XAFS and NMR.
We introduced our activities at the booth, and had a seminar by Prof. Nakamura on November 20th. We received many requests for our databases and services form participants.

Snapshot of the seminar and the booth.

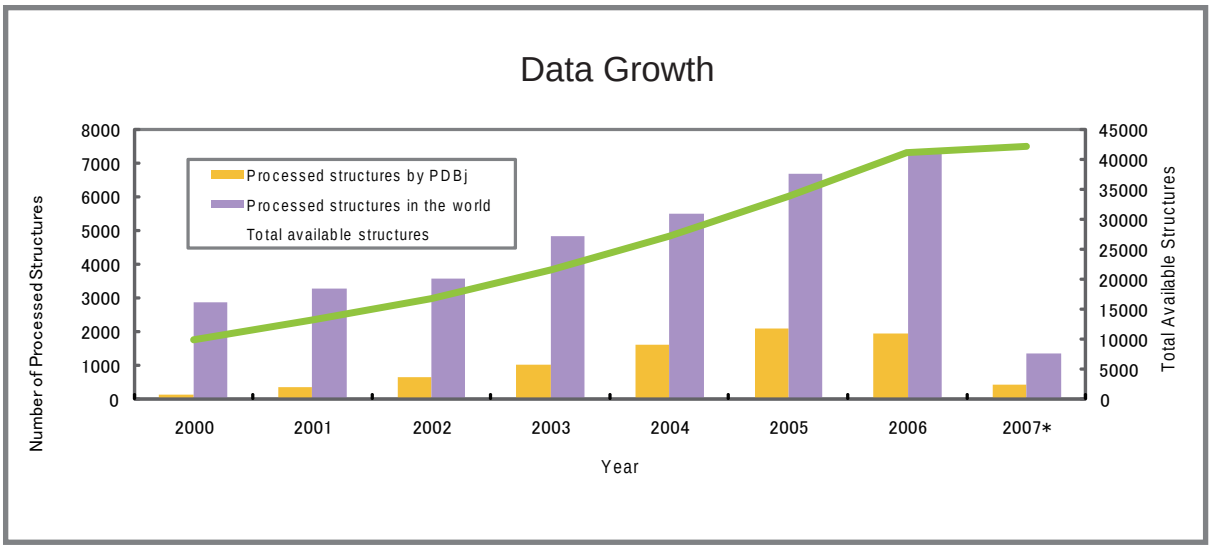
The 3rd DDBJing & PDBjing Workshop at Tokyo Tech.

The 3rd DDBJing & PDBjing Workshop, in corporation with DDBJ, was held December 20th and 21th, 2006, at Tokyo Institute of Technology. We introduced our activities and lectured our services forcusing mainly on PDBML and *e*F-site. The participants had hands-on training using PCs.

Snapshot of the Workshop.

Statistics

The statistics data is available at the wwPDB page (http://www.wwpdb.org/stats.html).



* Last updated : March, 2007

Services

Improvement of xPSSS

Our primary data browser, xPSSS (XML-based Protein Structure Search Service) has been improved. Now users can more easily focus on their desired targets using Keyword search. The search results are also sorted in a variety of ways, as the user requires. Moreover, additional experimental data, such as the crystallization conditions, which have been collected by the PDBj annotators from literature sources, are now fully displayed. Finally, in order to provide an historical record, pages have been prepared for downloading PDB files that are currently obsolete.

XQuery Advice system.

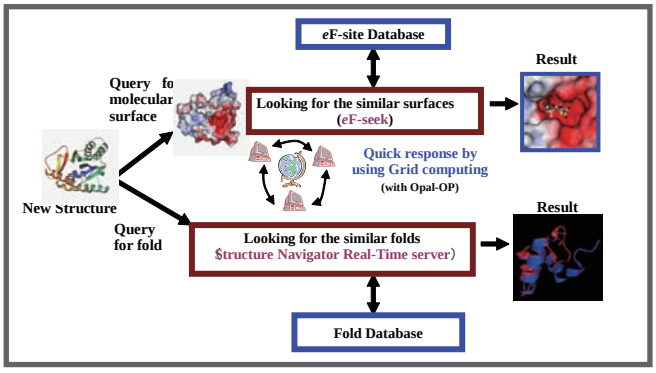
As before, by utilizing the advantage of an XML-DB, XPath searches can be performed. Now, XQuery searches are also available. In addition, we are now developing an XQuery Advice system, which will assist the user in submitting an XQuery, without knowing the details of the PDBML schema or XQuery syntax. This system will be available in April, 2007.

Grid Web service system for analog data searches

Over the last several years, PDBj has been developing query systems for protein shapes (folds) and molecular surfaces, using Structure Navigator-RT and eF-seek, respectively. Generally speaking, the information contained in a protein fold or surface is analog rather than digital. In contrast, text data, such as keywords or amino acid sequences are essentially digital. Because analog queries are more complex than digital queries, they require specialized application programs that use significant computing resources. Since both the available data and computer resources are expected to grow continuously in the near future, it is necessary to apply Grid technology to the analog search problem and to make use of computing resources other than those inside PDBj. To this end we have implemented our analog query tools as web services using OPAL and Opal-OP Grid technology.

OPAL is a software tool that has been developed by Sriram Krishnan, Wilfred Li, and Prof. Peter Arzberger at UCSD. OPAL can wrap command-line application programs in order to make them available as Web services. By using OPAL, a Web service can be submitted, the status of the job monitored, and the result obtained. It is also possible to cancel the job through OPAL. In order to make GT4 (Globus Toolkit 4) one of the Operation Providers Kohei Ichikawa, in the Graduate School of Information Science and Technology at Osaka University, in collaboration with the above UCSD group, has developed the Opal Operation Provider (Opal-OP). By submitting the job through Globus with the local job scheduler, jobs are distributed to several computing servers using the Globus general security system. Thus, reliable and stateful replies are expected.

PDBj has implemented an Opal-OP version of Structure Navigator-RT that enables jobs to automatically be distributed to two PC clusters, one at the Cybermedia Center at Osaka University and the other at the Institute of Medical Science at the University of Tokyo. This system makes it possible to process even large structure queries relatively quickly (i.e., in under a minute). This Grid architecture should provide a stable, secure, and efficient computing infrastructure for analog queries for the foreseeable future.



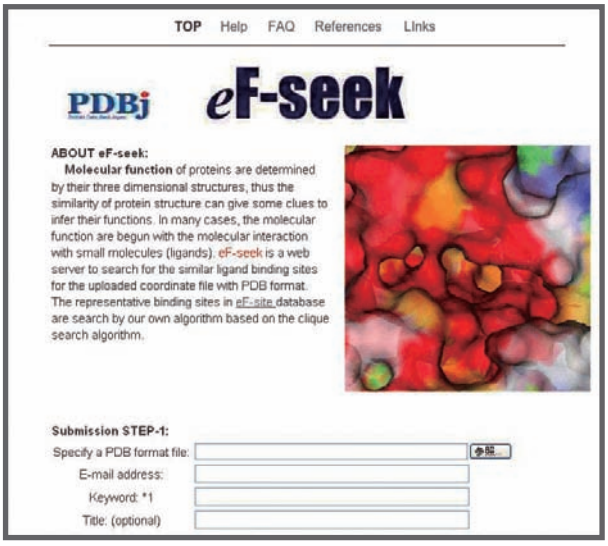
The basic concept of Opal-OP.

Start of *e*F-seek service

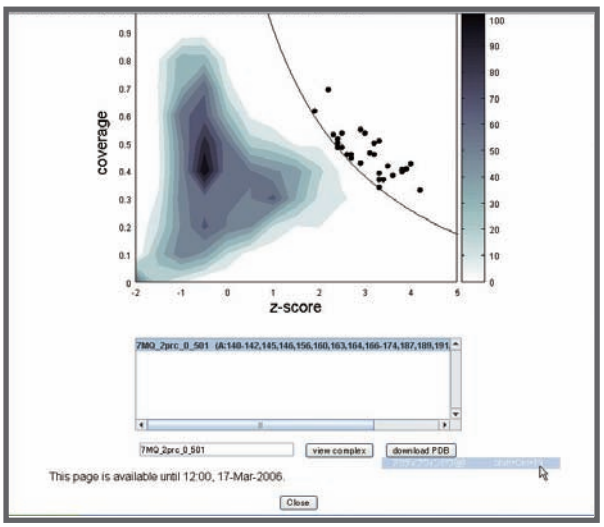
PDBj has developed a database of protein molecular surfaces, *e*F-site. Now, we have also started a new search service, *e*F-seek, that searches protein molecular surfaces in *e*F-site, and returns those that have similar geometries and electrostatic potentials to the query protein. In this service, users upload their own structure in the PDB flat format from the *e*F-seek frontpage, and our application program searches for similar surfaces. Finally, the results are returned on the Web result page.

In this web page, the candidates (shown as black filled circles) can be selected by a mouse, and the predicted complex structure can be displayed in *j*V by selecting the "view complex" option. The predicted structures are also downloaded as PDB files.

Since the search requires large computing resources, it may not be realistic for only a single computer system. Therefore, collaborating with the Institute for Medical Science at the University of Tokyo, a new GRID computing system has been developed for distributing the jobs. Consequently, the computation speed is about 30 times quicker than if the job were submitted to a single machine, and the search result for a typical protein will be given within a day.



The eF-seek frontpage.



The result page.

Staff

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

Group for PDB Database Curation
Kusunoki, Masami, Dr. (IPR, Osaka Univ.)
Kosada, Takashi (IPR, Osaka Univ.)
Matsuura, Takanori, Dr. (IPR, Osaka Univ.)
Igarashi, Reiko (JST-BIRD)
Kengaku, yumiko (JST-BIRD)
Ikegawa, Yasuyo (JST-BIRD)
Matsuura, Kanna (JST-BIRD)
Inoue, Mayumi (IPR, Osaka Univ.)

Group for Development of new tools and services
Standley, Daron M., Dr. (JST-BIRD and MEI center, Osaka Univ.)
Kinjo, Akira, Dr. (IPR, Osaka Univ.)
Iwasaki, Kenji, Dr. (IPR, Osaka Univ.)
Suzuki, Hirofumi, Dr. (IPR, Osaka Univ.)
Yamashita, Reiko (JST-BIRD)
Shimizu, Yukiko (JST-BIRD)
Kudou, Takahiro (IPR. Osaka Univ.)

Group for NMR Database
Akutsu, Hideo, Prof. (IPR, Osaka Univ.)
Takayama, Yuuki, Dr. (IPR, Osaka Univ.)
Nakatani, Eiichi（JST-BIRD)
Harano, Yoko (IPR, Osaka Univ.)

Group for Medical Institute of Bioregulation, Kyushu University
Toh, Hiroyuki, Prof. (MIB, Kyushu Univ.)

Group for School of Biomedical Science, Tokyo Medical and Dental University
Ito, Nobutoshi, Prof. (Tokyo Med. Dent. Univ.)

Derived Databases
Kinoshita, Kengo, Dr. (IMSUT, Tokyo Univ.)
Wako, Hiroshi, Prof. (Waseda Univ.)

Secretary
Kamada, Chisa (JST-BIRD)

Contacting

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
URL: http://www. pdbj.org/