# PDBj

Protein Data Bank Japan

Protein Data Bank Japan (PDBj)
News Letter Vol. 1
August 2002

http://www.pdbj.org/
and
http://pdbj.protein.osaka-u.ac.jp/

# Protein Data Bank Japan (PDBj) News Letter Vol. 1
http://www.pdbj.org/

August 2002

## 1. Introduction

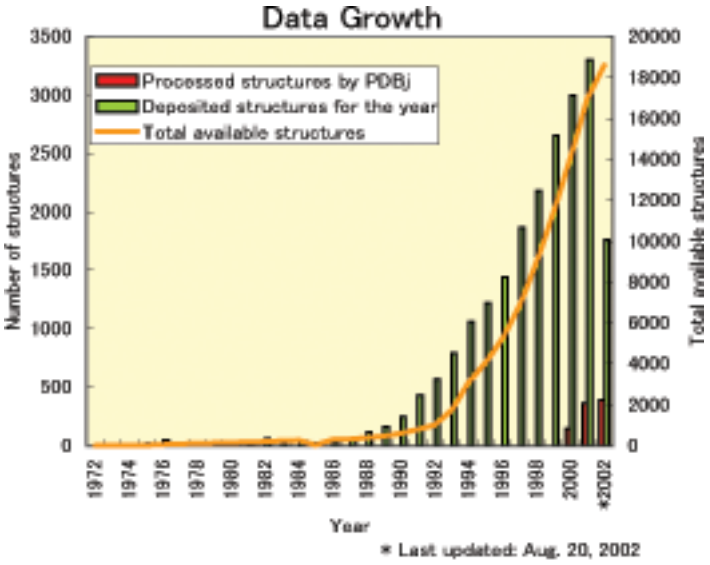
Now the science worldwide is focusing on structual genomics - structure determination in a high throughput system to elucidate all structures in a given proteome and enable to understand its complete biochemical pathway. It is anticipated that these initiatives will produce a great influx of data. PDB (Protein Data Bank) exists to meet the demands of managing and distributing such a tremendous amount of information.

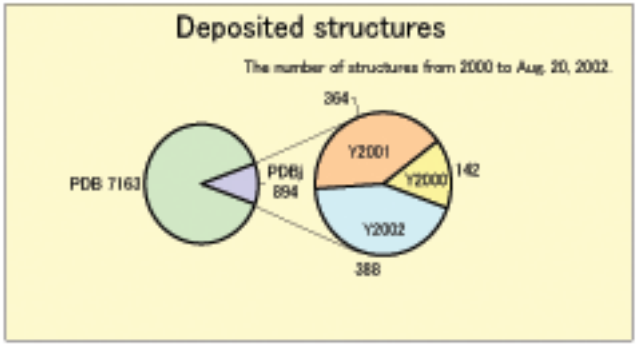
The three-dimentional structures of proteins and other biological macromolecules stored in the PDB are essential for biological sciences. Understanding the macromolecular structures will aid reserchers in understanding how proteins work. This information could give benefits to pharmaceutical and biotechnological industries in understanding various diseases and in developing drugs that enables to target diseases more effectively and with less chance of harmful side effects.

The PDB was established at Brookhaven National Laboratory in 1971. The full responsibility for the operation and enhancement of the PDB was transferred to the RCSB (Research Collaboratory for Structual Bioinformatics: http://www.rcsb.org/pdb/) in July 1999, which consists of Rutgers University, the National Institute of Standards and Technology (NIST) and the San Diego Supercomputer Center (SDSC).

The PDB has been working with the European Bioinformatics Institute (EBI) on data depositions at the EBI site, and on a data exchange dictionary, which is undergoing testing (http://msd.ebi.ac.uk/).

In Japan, the Institute for Protein Research (IPR) of Osaka University has become the representative archive and managed the database for Asia and Oceania with RCSB. Since July 2000, we have started accepting deposition of new structures from researchers in the region. We have also started projects to enhance the structual protein databases as Protein Data Bank of Japan (PDBj), with financial assistance from the Institute for Bioinfomatics Research and Development of Japan Science and Technology Corporation (BIRD-JST : http : // bird. jst .go.jp). As the number of depositions has increased rapidly in recent years, now over 18,000 structures are registered and released. About 50% of the recent structures are from US, 30% from Europe and 10% from Asia. Japan accounts for about 7-8%.



## 2. Data deposition at PDBj

Group for PDB Database Curation (Left to Right): Reiko Igarashi, Yumiko Kengaku, Masami Kusunoki, Takashi Kosada and Shinobu Saeki

The PDBj has cooperated with RCSB, the PDB operation group in USA, and started automatic entry using ADIT, data entry servers, since July 2000. The all entries are processed by the PDBj exclusive staffs. The structure data is put in the database faster and more accurate by interacting with the depositors to data authentication etc.

## 3. Research activities of PDBj

Although the maintenance of the database, including processing newly deposited structures, is the principle role of PDBj, we are also engaged in various research activities.

### 3-1. Development of PDBj-ML (XML for PDB)

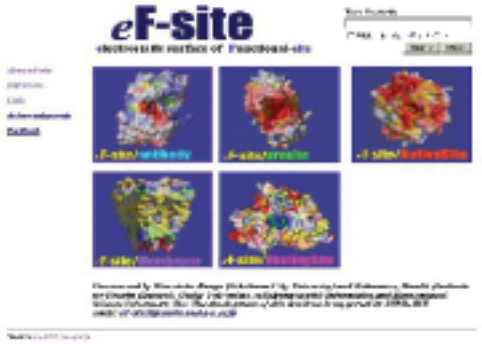
The protein structures in PDB are kept in so-called PDB format. Although this conventional format has undergone some revision, its limitation is widely felt. On the other hand, following a huge success of HTML, now XML is regarded as one of the most attractive way to exchange data between computers. We have, therefore, developed an XML format for PDB, PDBj-ML, and all the files are now available in the format, which is hoped to aid data exchange with other biological databases.


### 3-2. Development of structure/function database

Only a limited amount of information of molecular functions is found in the current PDB. On the other hand, papers reporting these structures usually have such information in addition to the structual data. Along with manual additions by annotators, we are also looking into the possibility of automated procedures. As the first step, we are developing computer software that extracts relevant information automatically from "remarks" in the PDB files.

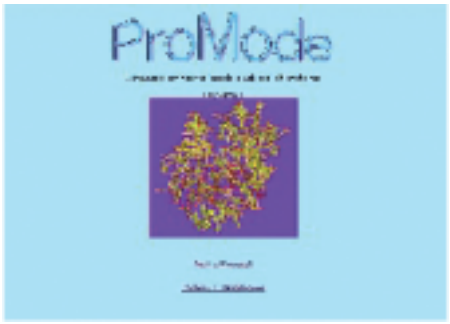
### 3-3. Database of the protein molecular surface (eF-site)

Much of interaction of proteins takes place on its surface. Thus surface properities, such as shape and electrostatic potential, are of great interest. eF-site is a database of such information, developed here at IPR (http://www.pdbj.org/eF-site/). Now as a part of the PDBj project, various improvements are being made to the database. You can browse and download the surface similar to that of your protein. This may aid to identify the function of proteins whose function is yet to be understood.



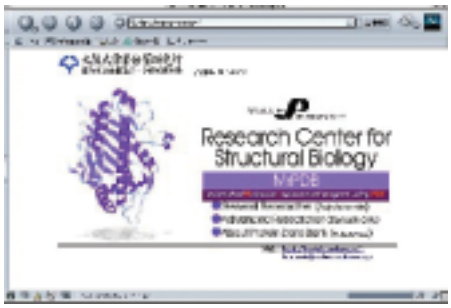
### 3-4. Database of the protein dynamics (ProMode)

It is very important to consider dynamics of proteins when one tries to understand its molecular function. Various methods exist to simulate the motion of proteins but such calculations require a vast amount of computational resources. A database of the dynamics would benefit many scientists who do not have access to such resources. We adopt the normal mode analysis for the calculations and the result are being accumulated for public access.




### 3-5. Automatic molecular replacement system (MrPDB)

When a new crystal structure of a protein is to be determined and the 3D structure of its homologous protein has already been known, the molecular replacement method is expected to accelerate the procedure very much. As more protein structures are determined and registered in PDB, this molecular replacement method becomes applicable to more protein crystals. PDBj offers an automatic analysis system, called MrPDB, applying the molecular replacement method. This system was originally developed at IPR by Masami Kusunoki.




### 3-6. Database of the experimental data of NMR (BioMagResBank)

A group from the University of Wisconsin, USA, is developing a database for experimental data of NMR. PDBj has started its mirror site (http://bmrb.protein.osaka-u.ac.jp) in Japan and plans to accept the data depositions in the near feature.


## 4. PDBj Staff

### Head

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

### Group for PDB Database Curation

Kusunoki, Masami (Associate Prof., IPR, Osaka Univ.)

Kosada, Takashi (Technical Assistant, IPR, Osaka Univ.)

Igarashi, Reiko (Research Assistant, JST)

Kengaku, Yumiko (Research Assistant,JST)

Saeki, Shinobu (Research Assistant, JST)

## Group for Development of PDBj-ML

Ito, Nobutoshi (Researcher, JST)

Kobayashi, Kaori (Technical Assistant, JST)

Sakamoto, Hisashi (Technical Assistant, JST)

Shimizu, Yukiko (Research Assistant, JST)

Takahashi, Aki (Research Assistant, JST)

## Group for Development of Secondary Databases and Analysis Systems

Kinoshita, Kengo (Instructor, Grad., Sch. of Integrated Sci., Yokohama City Univ.)

Wako, Hiroshi (Prof., School of Social Sciences, Waseda Univ.)

Okawa, Takenao (Associate Prof., Grad. Sch. of Informatics Sci., Osaka Univ.)

Toh, Hiroyuki (Research Director, Dept. of Bioinfomatics, Biomolecular Engineering Research Institute)

## Group for NMR Database (BMRB)

Akutsu, Hideo (Prof., IPR, Osaka Univ.)

Nakatani, Eiichi (Technical Assistant, JST)

## Secretary

Kamada, Chisa (JST)

## Office address

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@protein.osaka-u.ac.jp

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@protein.osaka-u.ac.jp