RCSB PDB NEWSLETTER

NUMBER 21 ◆ SPRING 2004 ◆ *Published quarterly by the Research Collaboratory for Structural Bioinformatics Protein Data Bank*

Weekly RCSB PDB news is available on the Web at www.rcsb.org/pdb/latest_news.html

Links to RCSB PDB newsletters are available at www.rcsb.org/pdb/newsletter.html

Contents

Snapshot: April 1, 2004

24,908 released atomic coordinate entries

Molecule Type		Experimental Technique	
22,528	proteins, peptides, and viruses	21,230	diffraction and other
1,301	nucleic acids	12,324	structure factor files
1,061	protein/nucleic acid complexes	3,678	NMR
18	carbohydrates	1,868	NMR restraint files

Participating RCSB Members

Rutgers: rutgers.rcsb.org

SDSC/UCSD: www.pdb.org

CARB/NIST: nist.rcsb.org

E-mail: info@rcsb.org

FTP: ftp.rcsb.org

Message from the RCSB PDB

The National Science Foundation (NSF) has renewed for five years funding for the PDB under the management of the RCSB. NSF has supported the PDB continuously since 1975, and a multi-agency support partnership first formed in 1989. For the past five years, that partnership has included NSF, the National Institute of General Medical Sciences (NIGMS), the Department of Energy (DOE), and the National Library of Medicine (NLM). The partnership has been expanded now to include the National Cancer Institute (NCI), the National Center for Research Resources (NCRR), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), and the National Institute of Neurological Disorders and Stroke (NINDS).

Wolfgang Bluhm, left, demonstrates the reengineered PDB site to Eric Sayers (NCBI) at the RCSB PDB's exhibition booth at the 48th Annual Meeting of the Biophysical Society, held February 14-18 in Baltimore, MD.

The new support agreement, which began Jan. 1, calls for the continued management of the PDB by three members of the Research Collaboratory for Structural Bioinformatics (RCSB): Rutgers, The State University of New Jersey; the San Diego Supercomputer Center at the University of California, San Diego; and the University of Maryland/National Institute of Standards and Technology's Center for Advanced Research in Biotechnology. This new era for PDB opens following the recent announcement of the wwPDB (www.wwpdb.org), an international agreement to coordinate the deposition and distribution of molecular structure data.

The PDB has continued to grow and evolve since its inception in 1971. Last year, more than 4,600 new molecular structures were added. On an average day, visitors download various structural files more than 120,000 times. During the next five years, the RCSB PDB will meet challenges that include the expanded integration of its information with other biological resources, keeping up with the increasing complexity and volume of deposited structures, meeting the demands for more complex queries, and providing more detailed annotation of the experiments and the structures. The PDB will also continue to serve an ever expanding, diverse and global user community. ✦

MIRROR SITES

Cambridge Crystallographic Data Centre (UK): pdb.ccdc.cam.ac.uk

National University of Singapore: pdb.bic.nus.edu.sg

Osaka University (Japan): pdb.protein.osaka-u.ac.jp

Max Delbrück Center for Molecular Medicine (Germany): www.pdb.mdc-berlin.de

Data Deposition and Processing

PDB Chemical Component Dictionary Format Description Available

The PDB is constantly improving its descriptions of small molecules. Its Chemical Component Dictionary (formerly called the HET Group Dictionary) is available in PDB and mmCIF formats, and is updated weekly. This resource, created by the curation efforts of the PDB teams at Rutgers University and CARB/NIST, is under active development.

A guide to the formats used in the PDB Chemical Component Dictionary is available at **deposit.pdb.org/cc_dict_tut.html**. This guide provides descriptions and examples of the contents of the PDB and mmCIF format Chemical Component Dictionaries, as well as a description of the contents of entries in the Ligand Depot.

Ligand Depot (**ligand-depot.rutgers.edu**) has been created as a data warehouse that integrates databases, services, tools, and methods related to small molecules that are bound to macromolecules. It was created to help users explore the PDB Chemical Component Dictionary and the small molecule contents of the PDB.

Comments and suggestions on the PDB Chemical Component Dictionary, the format description guide, and Ligand Depot are greatly appreciated, and may be sent to **info@rcsb.org**.

Validation of Protein Structures for the PDB

A paper published in the Macromolecular Crystallography volume of *Methods in Enzymology* describes the procedures used for structure validation and processing by the PDB. The article is an overview of the deposition process, including descriptions of the checks and reports generated by ADIT. The validation and standardization of all the data in the PDB are also outlined.

J. Westbrook, Z. Feng, K. Burkhardt, H.M. Berman: Validation of protein structures for protein data bank. *Methods Enzymol.* (2003) **374**: 370-85.

PDB Deposition Statistics

In the first quarter of 2004, approximately 1,273 structures were deposited in the PDB archive.

Of the structures received, 83% were deposited with a "hold until publication" release status; 3% were deposited with a specific release date; and 14% were deposited with a "release immediately" status.

84% of these entries were the result of X-ray crystallographic experiments; 13% were determined by NMR methods.

Release of New Deposition Tool at BMRB

BioMagResBank (BMRB) and RCSB PDB have jointly released a new ADIT-NMR deposition tool, which is available from the BMRB website (**www.bmrb.wisc.edu**). The current ADIT-NMR release provides the biological NMR community with a deposition interface for submitting a wide variety of quantitative experimental results from NMR spectroscopic studies of biological macromolecules to BMRB. The interface will be familiar to those who have submitted atomic coordinates and NMR constraints to the PDB.

ADIT-NMR has been developed as a collaborative project between the RCSB members PDB and BMRB. Through the PDB/BMRB collaboration, the ADIT-NMR tool will soon become a one-stop site for the deposition of all NMR derived data including atomic coordinates and constraints. The tool combines the BMRB dictionary for NMR experimental results with the software platform created by the RCSB PDB for the ADIT deposition system to create an easy to use interface for submitting NMR spectral parameters (chemical shifts, J-coupling constants, residual dipolar couplings, etc.), relaxation parameters, hydrogen exchange and pKa data to the BMRB. All members of the biological NMR community are encouraged to make use of ADIT-NMR in depositing their data at the BMRB and the RCSB PDB.

BMRB is the publicly accessible depository for NMR results from peptides, proteins, and nucleic acids recognized by the International Society of Magnetic Resonance and by the IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy.

Data Query, Reporting, and Access

Website Statistics

The PDB is available from several Web and FTP sites located around the world. Users are also invited to preview new features at the RCSB PDB beta test site, accessible at **beta.rcsb.org/pdb**.

The access statistics are given below for the primary RCSB PDB website at **www.pdb.org**.

Access Statistics for **www.pdb.org**

	DAILY AVERAGE			MONTHLY TOTALS		
MONTH	HITS	FILES	SITES	KBYTES	FILES	HITS
Jan 04	230,365	173,524	113,681	293,504,938	5,379,258	7,141,343
Feb 04	266,211	196,747	131,202	222,479,516	5,508,931	7,453,926
Mar 04	257,754	189,513	120,288	254,218,454	5,685,405	7,732,636

QuickSearch Available on RCSB PDB Site

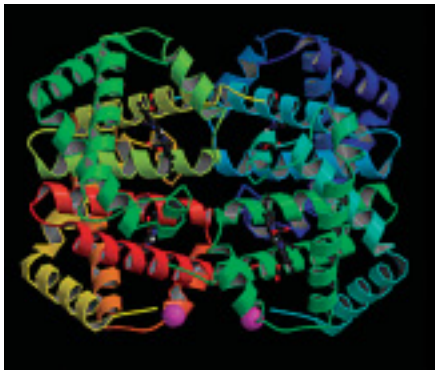
A new keyword search feature that searches across the PDB archive and/or the RCSB PDB website static pages is available. It supports the search syntax of the Lucene-based keyword search currently in production. An "exact word match" and "full text" search is performed on an index of the mmCIF files and an index of the static RCSB PDB Web pages. The structures returned by the search can be browsed, refined, and explored

using the Query Result Browser and Structure Explorer. The static page results are listed as links and displayed with the keyword highlighted in the context in which it appears. Please write to info@rcsb.org with comments or suggestions.

Accessing and Understanding Biological Units

When crystallographic structures are deposited in the PDB, the primary coordinate file generally contains one asymmetric unit—a concept that has applicability only to crystallography, but is important to understanding the process in obtaining the functional biological molecule. An introduction to biological units in the PDB archive is now accessible at www.rcsb.org/pdb/biounit_tutorial.html. This useful guide:

- Defines "asymmetric unit" and "biological molecule."
- Indicates where information about the biological unit can be found in PDB and mmCIF coordinate files.
- Describes how the biological unit files in the PDB have been derived.



In PDB entry **2hhb**, *the biological unit and the asymmetric unit are the same.*

PDB ID: **2hhb**

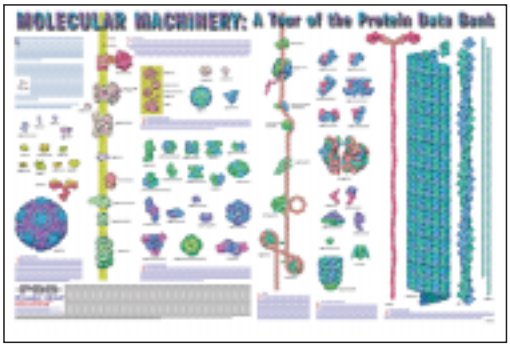
Fermi, G., Perutz, M. F., Shaanan, B., Fourme, R.: The crystal structure of human deoxyhaemoglobin at 1.74 Å resolution. J Mol Biol *(1984) 175: 159.*

Coordinate files for the biological units for applicable structures are accessible from the View Structure and Download/Display File sections of the Structure Explorer pages on the primary PDB website and its mirrors. The biological unit coordinate files can also be downloaded from the PDB FTP site at ftp://ftp.rcsb.org/pub/pdb/data/biounit/coordinates/. Images for the asymmetric and biological units are available from each entry's View Structure section of the Structure Explorer page. ✦

RCSB PDB Outreach

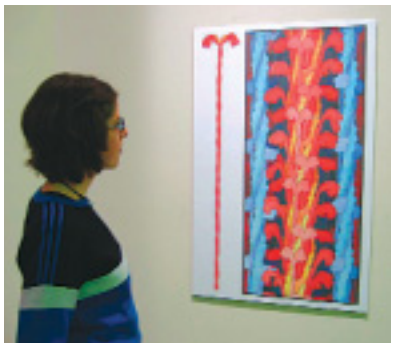
Molecular Machinery Poster Available as PDF

The poster, "Molecular Machinery: A Tour of the Protein Data Bank" by David S. Goodsell, is available as downloadable PDF files, one as a two-sided, 8 1/2" x 11" quick reference guide (5MB) at www.rcsb.org/pdb/molecules/poster_quickref.pdf, and the other as a 24" x 36" poster (31MB) at www.rcsb.org/pdb/molecules/poster_full.pdf.



Art of Science *Exhibit Visits the University of Wisconsin-Madison*

The RCSB PDB's *Art of Science* exhibit, which includes posters from the *Molecule of the Month* series, was displayed in the Biochemistry Atrium at the University of Wisconsin-Madison in February and March. The exhibit was sponsored by the Center for Eukaryotic Structural Genomics (CESG) and the BioMagResBank (BMRB).



The artistry of actin at the University of Wisconsin.

The PDB: A Case Study in Management of Community Data

A paper in the inaugural issue of *Current Proteomics* describes the development of the PDB and the expansion of its community of data depositors and users. The lessons learned from the development of the PDB into a Web-based archive containing approximately 25,000 released structures and more than 160,000 hits per day may be applicable to the ongoing development of new data and knowledge resources in proteomics.

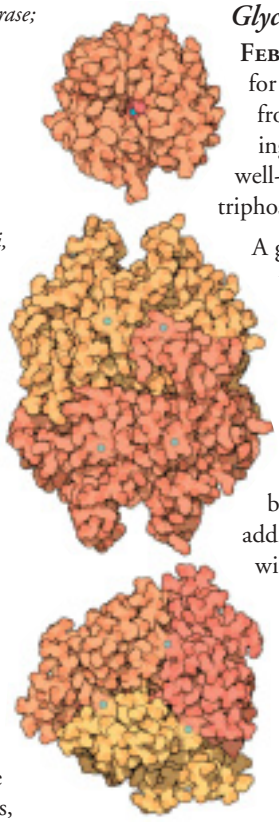
H.M. Berman, P.E. Bourne, J. Westbrook: The Protein Data Bank: A case study in management of community data. *Current Proteomics* (2004) **1**: 49-57.

Molecules of the Quarter: Carbonic Anhydrase, Glycolytic Enzymes, Calcium Pump

The *Molecule of the Month* series, by David S. Goodsell, explores the functions and significance of selected biological macromolecules for a general audience (www.rcsb.org/pdb/molecules/molecule_list.html). Three features during this past quarter are included below:

Carbonic Anhydrase: Breathing in, Breathing Out

January, 2004—Breathing is a fundamental function in life. In our lungs, oxygen diffuses into the blood, binds to hemoglobin, and is transported to all the cells of our body. Carbon dioxide is a byproduct of sugar and fat breakdown and must be removed from the body. However, less than 10% of the carbon dioxide that diffuses out of cells dissolves in the blood plasma, about 20% binds to hemoglobin, while 70% is converted to carbonic acid to be carried to the lungs. Carbonic anhydrase, an enzyme in red blood cells, aids in the conversion of carbon diox-



Alpha (top), beta, and gamma forms of carbonic anhydrase; PDB entries **1ca2**, **1ddz**, *and* **1thj**, *respectively.*

PDB ID: **1ca2**

Eriksson, A. E., Jones, T. A., Liljas, A.: Refined structure of human carbonic anhydrase II at 2.0 Å resolution. Proteins *(1988)* ***4****: 274.*

PDB ID: **1ddz**

Mitsuhashi, S., Mizushima, T., Yamashita, E., Miyachi, S., Tsukihara, T.: X-Ray Structure of Beta-Carbonic Anhydrase from the Red Alga, Porphyridium purpureum*, Reveals a Novel Catalytic Site for CO_2 Hydration* J.Biol.Chem. *(2000)* ***275****: 5521.*

PDB ID: **1thj**

Kisker, C., Schindelin, H., Alber, B. E., Ferry, J. G., Rees, D. C. : A left-hand beta-helix revealed by the crystal structure of a carbonic anhydrase from the archaeon Methanosarcina thermophila. EMBO J *(1996)* ***15****: 2323.*

ide to carbonic acid and bicarbonate ions. When red blood cells reach the lungs, the enzyme helps to convert the bicarbonate ions back to carbon dioxide, which we breathe out.

Since its identification in 1933, carbonic anhydrase has been found abundant in all mammalian tissues, plants, algae, and bacteria. This ancient enzyme has three distinct classes— alpha, beta, and gamma. While all three require a zinc ion at the active site, members of one class share very little sequence or structural similarity with the other two classes, suggesting that each class evolved independently. Carbonic anhydrase from mammals belongs to the alpha class, the plant enzymes belong to the beta class, while the enzyme from methane-producing bacteria that grow in hot springs forms the gamma class. PDB entries **1ca2**, **1ddz**, and **1thj** are examples of the alpha, beta, and gamma carbonic anhydrase enzymes, respectively. The zinc ions in the active sites are blue. The alpha enzyme is a monomer, while the gamma enzyme is trimeric. Although the beta enzyme is a dimer, there are four zinc ions bound to the structure indicating four possible enzyme active sites.

Since this enzyme produces and uses protons and bicarbonate ions, carbonic anhydrase plays a key role in the regulation of pH and fluid balance in different parts of our body. When there is a build up of the fluid that maintains the shape of our eyes, the fluid often presses on the optic nerve in the eye and may damage it. This condition is called glaucoma. In recent years, inhibitors of carbonic anhydrase have been used to treat glaucoma.

For more information on carbonic anhydrase, see www.rcsb.org/pdb/molecules/pdb49_1.html

Glycolytic Enzymes

February, 2004—Glucose is a convenient high-energy fuel for cells because it is stable, soluble, and easy to transport from storage to where it's needed. Glycolysis (sugar breaking) is a ten-step cellular process to burn glucose in small, well-controlled steps to capture the energy as ATP (adenosine triphosphate).

A glucose molecule is primed with two phosphates (using up two ATP molecules), broken in two, reshaped, and dehydrated, forming four ATP molecules in the process, or a net gain of two ATPs.

One glycolytic enzyme removes several hydrogen atoms from the sugar, transferring them to the small carrier molecule NAD (nicotinamide adenine dinucleotide). Many cells, including most of our own, eventually combine the hydrogens with oxygen to form water, building additional ATP in the process. In a reaction used to make wine and beer, yeast cells use alcohol dehydrogenase to add the hydrogen atoms back to the broken sugar molecule, forming alcohol. In extreme exercise, muscles add the hydrogen atoms back in a different way to form lactic acid.

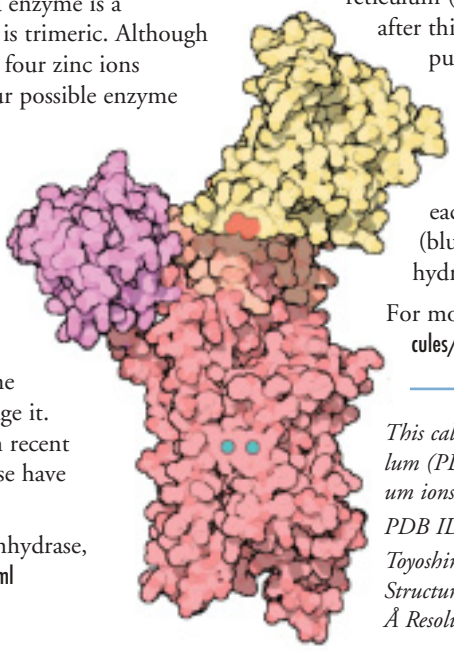
For more on glycolytic enzymes, see www.rcsb.org/pdb/molecules/pdb50_1.html

Calcium Pump

March, 2004—Every time we move, our muscle cells use calcium ions to coordinate a massive molecular effort. These cells release a flood of calcium ions from a special intracellular container, the sarcoplasmic reticulum, which surrounds the bundles of actin and myosin filaments. The calcium ions rapidly spread and bind to tropomyosins on actin filaments. They shift shape slightly and allow myosin to bind and begin climbing up the filament, contracting the muscle.

The calcium pump, found in the membrane of the sarcoplasmic reticulum (right) from PDB entry **1eul**, allows muscles to relax after this frenzied wave of calcium-induced contraction by pumping calcium ions back into the sarcoplasmic reticulum. This allows the muscle to relax. The pump has a big domain poking out of the sarcoplasmic reticulum, and a region that is embedded in the membrane, forming a tunnel to the other side. For each ATP broken, the pump transfers two calcium ions (blue spheres) through the membrane, and two or three hydrogen ions in the opposite direction.

For more on the calcium pump, see www.rcsb.org/pdb/molecules/pdb51_1.html ✦



This calcium pump in the membrane of the sarcoplasmic reticulum (PDB entry **1eul***) allows muscles to relax by pumping calcium ions back into the sarcoplasmic reticulum.*

PDB ID: **1eul**

Toyoshima, C., Nakasako, M., Nomura, H., Ogawa, H.: Crystal Structure of the Calcium Pump of Sarcoplasmic Reticulum at 2.6 Å Resolution. Nature *(2000)* ***405****: 647*

PDB Community Focus: Helen M. Berman

Helen M. Berman, Director of the RCSB Protein Data Bank, completed an AB degree in chemistry at Barnard College in 1964, and in 1967 received her PhD in natural science from the University of Pittsburgh where she studied with George Jeffrey in the Department of Crystallography. In 1969 she went to the Institute for Cancer Research (ICR), Fox Chase Cancer Center in Philadelphia to work with Jenny Glusker. She became an Assistant Member of ICR in 1973 and then rose through the ranks to Senior Member. She was also an Adjunct Professor at the University of Pennsylvania and Director of Research Computing at Fox Chase. In 1989 she moved to Rutgers University where she currently serves as a Board of Governors Professor of Chemistry and Chemical Biology.

Dr. Berman's crystallographic studies have focused on nucleic acids, protein-nucleic acid complexes, and collagen. She has also done systematic analyses of the hydration patterns of biological molecules, including nucleic acids and collagen. Since the earliest days of her career, she has been interested in establishing methods to collect and archive structural data so that systematic studies of the data could be facilitated. She was part of the original team that developed the PDB at Brookhaven National Laboratory in 1971, and in 1991 she founded the Nucleic Acid Database (NDB; http://ndbserver.rutgers.edu/). In 1998, she led the team of Research Collaboratory for Structural Bioinformatics (RCSB) members that won the contract to manage the PDB.

Throughout her career, Dr. Berman has been an active participant in the scientific community. She has served on numerous advisory boards for the National Science Foundation, the National Institutes of Health, and on journal editorial boards. She has served as President of the American Crystallographic Association (ACA) and has also held leadership positions in the Biophysical Society and the International Union of Crystallography (IUCr). She received the 2000 Biophysical Society Award for Distinguished Service and is a Fellow of the Biophysical Society and of the American Association for the Advancement of Science. Under Dr. Berman's leadership, the RCSB began its second five-year period of PDB management in January 2004. During the first five years, the number of released structures in the archive had more than doubled, and the pace of depositions continues to increase at a steady rate. The RCSB PDB staff solicited Dr. Berman's views on RCSB PDB's accomplishments to date, and her vision for the future.

Q The PDB was born at a Cold Spring Harbor Symposium in 1971. What was that meeting like?

A It was an enormously exciting meeting, especially for a young crystallographer. Virtually all the pioneers of the field were there presenting the results of their research. A particularly vivid image I have is of a large group of people sitting on the grass in a circle around Max Perutz talking about hemoglobin.

Q What has shaped the PDB the most the since its beginnings as an archive containing seven structures?

A There has been a progression of influences on the PDB. First the focus was on getting structures into the PDB. In the 1970s Tom Koetzle single handedly wrote personal letters to every protein crystallographer asking them to participate. In the 1980s, the community of protein crystallographers began to organize under the leadership of Fred Richards to try to encourage people to deposit structures. The IUCr set up more formal committees to achieve the same thing. In 1989, guidelines were set forth requiring deposition and release of macromolecular structures. At the same time the technology improved making structure determination much faster. By the early 1990s the number of depositions began to rise quickly and the problem of how to keep up with the data emerged. A backlog of structures began to build. The PDB was a victim of its own success. The use of modern data management methods as well as the development of an efficient team of annotators has helped to solve this problem. The challenges in the 2000s will be high throughput structure analysis, large macromolecular assemblies, and the demand for archiving more information about each experiment and its results.

Q You have been actively involved with the PDB since its beginning. What continues to draw you to this project?

A The idea that by looking at groups of structures it will be possible to derive new knowledge has always been a compelling concept. I have done this in my own research and have always wanted to make it possible for others. If all the data are organized properly, it should be possible to mine it efficiently. If many people can do this on large data sets, it should be possible to learn about basic concepts, such as protein folding, and use the knowledge to create new drugs.

Q Some people think that being involved with scientific infrastructure, such as the PDB, is not doing *real* science and is therefore less important. Do you agree?

A Not at all. It is one of the most important things that I can do. To do it right requires knowledge of the data and the technology needed to collect and disseminate it. Once the infrastructure is in place, new science will emerge. To facilitate that process and to see what emerges, and to imagine what new science will be facilitated is what motivates me.

Q The RCSB is comprised of organizations in different parts of the country. How does the collaboration work?

A That is complicated. Each site has its own set of projects that contribute to the PDB. However, each project must also interact with all the others. To make this work we have developed various computer-based forums. Daily communication by email and phone is critical, as are personal visits and personnel exchanges among the groups. Recently, we have begun to use video conferencing. Once a year we have a retreat that allows everyone to be together for a couple of days to talk about the various projects, to plan for the coming year, and for PDB staffers to get to know one another.

Q Who makes up the PDB's user community?

A It used be only crystallographers, and later NMR spectroscopists, but now it has expanded to include biologists, computational biologists, educators, and students.

Q How does the PDB interact with its users?

A We have various electronic mail services that allow users to ask questions and bring problems to our attention. We attend many different meetings and participate in a variety of ways. At some meetings we have an exhibit booth. We are also organizing workshops for the purpose of educating different parts of the community about what we do and how to use the various tools. Outreach is a key element of the PDB because it gives us the feedback we need to improve what we do.

Q Recently, the RCSB formalized the ongoing collaboration with the Macromolecular Structure Database-European Bioinformatics Institute and PDBj (PDB Japan) to form the wwPDB. How does this affect the PDB?

A The wwPDB was organized to make sure that the PDB remains as a single archive. When users from around the world access a flat file with ID **1XXX** they can be assured that they will always get the same file. This organization will also help us develop new collaborations that will enhance the use of the PDB files. Science is international and wwPDB acknowledges that.

Q Currently, 28 of your structure determinations are in the PDB. What has been your experience as a PDB depositor?

A Watching my students deposit files in the PDB allows me to see how we can improve the process.

Q You are active in other areas of research, including protein-nucleic acid interactions, structure determinations, and databases. Has this activity influenced your work with the PDB?

A As a depositor and a user, I have both perspectives. I have used structural data to do systematic analyses of macromolecules. The need to have easy access to the data was a motivating force for me in helping to improve the PDB. As a user of the PDB, I can see how we can make it easier for others to use.

Q What do you think the PDB will be like in the next 30 years?

A In 1971, it was almost beyond our imagination that structure determination of proteins could be completed in a few days with the results instantly accessible on a desktop computer. But here we are, and we now know for sure that in the future there will be even more structures, new methods for structure determination, much larger structures and we will have more information about each structure. All aspects of the process will be fully automated. The really exciting thing to think about is what people will do with all the data. This will depend on the ingenuity of new generations of biologists, some of whom are not yet born, who will certainly find ways to use all this information and give us the ultimate knowledge about how molecules function. ✦

RCSB PDB Job Listing

Structural Bioinformatics Project Leader

The PDB group at the University of California, San Diego is seeking a Project Manager to guide the PDB in its next five-year phase of development. The Project Manager will work collaboratively with the PDB software architects, programmers, and scientists, at UCSD and the RCSB PDB partner sites, to expand the PDB's functionality and reliability as a premier biological data and information resource.

Job functions include: identify and develop requirements for new PDB delivery, query functionality, and usability; develop and implement innovative approaches that will satisfy users' current needs and anticipate their future needs based on progress in the science of structural biology and structural bioinformatics; and work with the scientific community to fulfill the above.

Qualifications include:

- Ph.D. in biological sciences or related field or equivalent combination of education and experience in the field of bioinformatics and computational biology, including expert knowledge and research experience in sequence and protein structure analysis, protein 3-D structure prediction and fold recognition, and protein modeling.
- Extensive background and expertise in project management and research coordination.
- Demonstrated experience working with a team of high-level professional scientists and computer programmers.

Applicants should apply on-line at
joblink.ucsd.edu/bulletin/job.html?cat=new&job_id=31324.

Related Links: File Formats

www.rcsb.org/pdb/info.html#File_Formats_and_Standards

The RCSB PDB offers links to software tools and descriptive resources for file formats used to describe PDB data. A few of these links include:

mmCIF Resources deposit.pdb.org/mmcif/
Background information, Data Dictionaries (including the PDB Exchange Dictionary), Proposed Data Items for Structural Genomics Depositions, mmCIF–PDB correspondences, OMG CORBA API, and software tools

Chemical Component Dictionary and Description
The PDB Chemical Component Dictionary (formerly the HET Group Dictionary) is available in PDB and mmCIF formats. A guide to these dictionaries is available at deposit.pdb.org/cc_dict_tut.html

PDB File Format Contents Guide Version 2.2
www.rcsb.org/pdb/docs/format/pdbguide2.2/guide2.2_frame.html
A description of the PDB file format.

PDB Education Corner by Tommie S. Hata

Tommie Hata is a biology teacher at The Pingry School in Martinsville, NJ.

The Protein Data Bank has become a significant and popular website to supplement our curriculum within molecular biology. Despite attempts by text books to invigorate material with animations and CD-ROM presentations, the material in a book is often static and disconnected from the real world of science. The PDB has given life to molecules, the people behind them, and the process of discovering them.

The Pingry School is a special place. It is a country day school that offers many opportunities to the student body. The students are involved in a number of diverse activities from glass blowing to biomolecular modeling in addition to the standard academic rigor. Our students have a variety of backgrounds and interests, yet have a unifying theme in a motivated attitude towards learning and achieving. The ninth grade biology program at our school has made a distinct effort to reduce the breadth of the curriculum and improve the depth of knowledge. This concept has led away from copious note taking on many broad topics within biology and deeper exploration of the process of science. The PDB has truly made this possible.

The first unit within our curriculum is the structure and function of macromolecules. When students are learning the specifics of protein structure, they predict a structure of a retinol binding protein (RBP) and use the PDB (**1aqb**) to compare their predicted structure to an actual RBP. The students look specifically for the hydrophobic nature in amino acids within the "carrying" portion of the molecule, as compared to hydrophilic portion of the molecule, using computer programs such as MDL Chime and RasMol. The reality of seeing the amino acids in a three dimensional form, comparing different records of RBP, and manipulating the molecule using chime led to many to exclaim "Ahaa!" or "I was right!" and facilitate discussions on the other possibilities of structure. Using Boolean commands in RasMol, such as "sidechain and alpha" or "backbone and helix," further increases the students' awareness of protein structure. On separate occasions, students have generated RasMol scripts to highlight individual amino acid sidechains in hemoglobin that interact with the heme group and to highlight cAMP bound within catabolite activator protein (CAP). The educator is simply a tool in the learning process; the students are given the resources to facilitate learning.

The *Molecule of the Month* feature by Dr. David Goodsell has been integral to deepening knowledge of other biological molecules and processes. Almost every unit is supplemented by having the students read a feature and then answer questions developed by the teachers. Topics such as cellular respiration, photosynthesis, gene regulation, and membrane transport relate to *Molecule of the Month* features. Some students take the next step to view and analyze the PDB files referenced in the features and compared different records.

In September, The Pingry School established a Students Modeling A Research Topic (SMART) team under the guidance of Dr. Tim Herman at the Milwaukee School of Engineering (see PDB Education Corner from Summer 2003 issue). With help from Dr. Richard Ebright (HHMI/Waksman Institute/ Rutgers), Dr. Helen Berman, and other scientists at the PDB, my seven students have been able to design a physical model of a class I transcription-activation complex (**www.mybiology.com/smartteam.htm**). College-level concepts such as protein structural motifs have become part of our discussions, which would not have been possible without the PDB. Through their interaction with Dr. Berman and the PDB, the students also gain a better understanding and appreciation for both the science and scientists involved in crystallography.

In addition to becoming a tool used in our classroom, the PDB has made scientists and their work tangible to our students. At Pingry, the PDB has morphed from a tool exclusive to the science community to a shortcut on the students' computer desktops. ✦

Annotator Kyle Burkhardt discusses protein structure with Tommie Hata and his class at the RCSB PDB site at Rutgers.

RCSB PDB Leadership Team

The overall operation of the PDB is managed by the RCSB PDB Leadership Team.
Technical and scientific support are provided by the RCSB PDB Members.

Dr. Helen M. Berman, *Director*
Rutgers University
berman@rcsb.rutgers.edu

Dr. Philip E. Bourne
Co-Director
SDSC/UCSD
bourne@sdsc.edu

Judith Flippen-Anderson
Production and Outreach Leader
Rutgers University
flippen@rcsb.rutgers.edu

Dr. Gary Gilliland
Co-Director
CARB/NIST
gary.gilliland@nist.gov

Dr. John Westbrook
Co-Director
Rutgers University
jwest@rcsb.rutgers.edu

RCSB PDB Members

The RCSB PDB is managed by three partner sites of the Research Collaboratory for Structural Bioinformatics:

Rutgers

Rutgers, The State University of New Jersey
Department of Chemistry and Chemical Biology
610 Taylor Road
Piscataway, NJ 08854-8087

Prentice Bisbal
prentice@rcsb.rutgers.edu

Kyle Burkhardt
kburkhar@rcsb.rutgers.edu

Li Chen
lchen@rcsb.rutgers.edu

Sharon Cousin
sharon@rcsb.rutgers.edu

Dr. Shuchismita Dutta
sdutta@rcsb.rutgers.edu

Dr. Zukang Feng
zfeng@rcsb.rutgers.edu

Lew-Christiane Fernandez
fernandz@rcsb.rutgers.edu

Dr. Rachel Kramer Green
kramer@rcsb.rutgers.edu

Vladimir Guranovic
vladimir@rcsb.rutgers.edu

Dr. Shri Jain
sjain@rcsb.rutgers.edu

Dr. Rose Oughtred
rose@rcsb.rutgers.edu

Dr. Irina Persikova
persikova@rcsb.rutgers.edu

Suzanne Richman
richman@rcsb.rutgers.edu

Melcoir Rosas
melcoir@rcsb.rutgers.edu

Dr. Bohdan Schneider
bohdan@rcsb.rutgers.edu

Dr. Huanwang Yang
hyang@rcsb.rutgers.edu

Dr. Jasmin Yang
jasmin@rcsb.rutgers.edu

Christine Zardecki
zardecki@rcsb.rutgers.edu

SDSC/UCSD

San Diego Supercomputer Center
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0537

Dr. Ken Addess
addess@sdsc.edu

Dr. Wolfgang F. Bluhm
wbluhm@sdsc.edu

Dr. Nita Deshpande
nita@sdsc.edu

Ann Kagehiro
kagehiro@sdsc.edu

Jeff Merino-Ott
jott@sdsc.edu

Wayne Townsend-Merino
wayne@sdsc.edu

CARB/NIST

Center for Advanced Research in Biotechnology
National Institute of Standards and Technology
9600 Gudelsky Drive
Rockville, MD 20850

Al Carlson
carlson@umbi.umd.edu

Dr. Veerasamy Ravichandran
vravi@nist.gov

Kathryn Rosecrans
rosecran@umbi.umd.edu

Elizabeth Walker
walkere@umbi.umd.edu

Statement of Support

The RCSB PDB is supported by funds from the National Science Foundation, the National Institute of General Medical Sciences, the Department of Energy, the National Library of Medicine, the National Cancer Institute, the National Center for Research Resources, the National Institute of Biomedical Imaging and Bioengineering, and the National Institute of Neurological Disorders and Stroke.

RCSB PROTEIN DATA BANK
www.pdb.org

San Diego Supercomputer Center
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0537

Return Service Requested