PDB NEWSLETTER

NUMBER 18 ◆ SUMMER 2003 ◆ PUBLISHED QUARTERLY BY THE PROTEIN DATA BANK

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

Links to this and previous PDB newsletters are available at **www.rcsb.org/pdb/newsletter.html**

Contents

Snapshot: June 15, 2003

21,248 released atomic coordinate entries

Molecule Type

19,151	proteins, peptides, and viruses
1,199	nucleic acids
880	protein/nucleic acid complexes
18	carbohydrates

Experimental Technique

18,010	diffraction and other
9,629	structure factor files
3,238	NMR
1,567	NMR restraint files

Participating RCSB Members

SDSC/UCSD: www.pdb.org

Rutgers: rutgers.rcsb.org

CARB/NIST: nist.rcsb.org

E-mail: info@rcsb.org

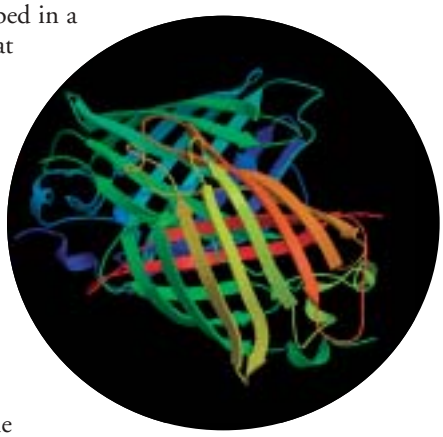
FTP: ftp.rcsb.org

Message from the PDB

This summer, the PDB will be featuring many of the developments described in this newsletter at meetings.

At the 11th International Conference on Intelligent Systems for Molecular Biology (ISMB), the PDB will be demonstrating many new features for query and reporting at the PDB exhibit booth. The meeting will be held June 29–July 3 in Brisbane, Australia.

These new query and reporting features will also be described in a poster being presented at the 17th Symposium of the Protein Society (July 27–29, Boston, MA).



Green fluorescent protein, the PDB's Molecule of the Month for June, 2003

PDB ID: ***1gfl***

F. Yang, L.G. Moss, G.N. Phillips, Jr. (1996): The molecular structure of green fluorescent protein. Nat. Biotechnol. *14, p. 1246.*

ADIT software will be distributed at the American Crystallographic Association's Annual Meeting (July 26–31, Northern Kentucky Convention Center). This standalone version of ADIT is for use on your own desktop computer.

The PDB Poster prize (see page 5) will also be making its debut at the ACA meeting. The prize will also be awarded at the Asian Crystallographic Association (AsCA) and the European Crystallographic Association (ECM) meetings this summer.

We look forward to seeing you soon.

The PDB ✦

The Protein Data Bank (PDB) is the single worldwide repository for the processing and distribution of 3-D biological macromolecular structure data. The PDB is operated by Rutgers, The State University of New Jersey; the San Diego Supercomputer Center (SDSC) at the University of California, San Diego (UCSD); and the Center for Advanced Research in Biotechnology (CARB) of the National Institute of Standards and Technology (NIST)—three members of the Research Collaboratory for Structural Bioinformatics (RCSB), a non-profit consortium dedicated to improving our understanding of biological systems.

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**

Universidade Federal de Minas Gerais (Brazil): **www.pdb.ufmg.br**

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

Data Deposition and Processing

ADIT Software Available for Download

A standalone version of ADIT for use on your own desktop computer is available for download from **pdb.rutgers.edu/software**.

This version has the same features as the web version of ADIT.

ADIT is an integrated software system for editing and checking PDB structure data entries. The system includes tools to help users prepare and check structure depositions. ADIT is currently available in source form and in binary form for Linux platforms.

The ADIT structure deposition interface

If you prefer to use the web version of ADIT to deposit your structures, we urge you to run format prechecks and validation prechecks prior to deposition.

PDB Deposition Statistics

As of June 1, 2003, more than 2,000 structures have been deposited to the PDB. 78% have been deposited with the release status HPUB; 13% release immediately, and 8% on HOLD.

84% are from X-ray crystallographic studies and 12% are from NMR studies.

55% of these depositions release the sequence in advance of the structure's release.

68% of these depositions were deposited with experimental data.

Submission of Structure Factor Data to the PDB

In a recent message to the PDB list server Gerard Kleywegt and Alwyn Jones have requested crystallographers worldwide to deposit structure factor data for existing and future PDB entries. The PDB strongly supports this request.

PDB tries to make the submission of structure factor data as easy as possible. Data can be submitted in any format with an accompanying description of content. Structure factor data for new entries can be uploaded along with coordinate data at the time of deposition. For existing entries depositors can simply mail structure factor data to **deposit@rcsb.rutgers.edu**.

As emphasized in the following message, the structure factor data are an important component of the PDB archive. We ask for the cooperation of all crystallographers to help strengthen the scientific content of the PDB archive by depositing structure factor data for all of their entries both new and old.

Dear colleague!

At present, structure factor data are available for only about *half* of all crystal structures in the PDB. Unless we all make a serious effort now, we must fear that these data will be lost to science for all eternity. Therefore, we would like to encourage all macromolecular crystallographers to check if any of their PDB entries perhaps do not have the associated structure factors deposited. To help you do this, a simple form to query the RCSB database is available at this URL:

fsrv1.bmc.uu.se/eds/eds_sos.html

Simply type (a unique part of) your name and hit the "Check" button to get a list of any and all such entries. If there are any, please try and track down the structure factors (on old disks, tapes, or by asking former students and post-docs, etc.) before they are lost forever. If you find any, please send them to the RCSB (**deposit@rcsb.rutgers.edu**). (By the way: the most likely future user of deposited structure factor data are you yourself!!)

As you may know, we have been working on creating an archive of electron density maps for all crystal structures in the PDB for which structure factor data have been deposited—the Uppsala Electron Density Server (EDS; URL: **fsrv1.bmc.uu.se/eds**). At present, in about one percent of cases, we are unable to calculate a map at all, and for another ~15% of cases we are unable to reproduce the published R-value to within five percentage-points. The webpage with the search form mentioned above also contains a request for you to help us improve our ~85% success rate with EDS map calculations.

These two initiatives combined will help to preserve and improve the wealth of macromolecular crystallographic data in the public databanks and to make them available and easily accessible to the entire scientific community (cell and molecular biologists, medicinal chemists, crystallographers, etc. etc.) now and in the future.

Thank you for your time and help in advance!!

--Gerard Kleywegt & Alwyn Jones

P.S.: please direct technical correspondence about EDS to **eds@xray.bmc.uu.se**

P.P.S.: in this request it has been tacitly assumed that coordinates of all published structures have been deposited already. In cases where this is not so, you are of course also strongly encouraged to dig up the models and deposit them together with the corresponding structure factors.

P.P.P.S.: please help this initiative by bringing it to the attention of colleagues who may not read the electronic crystallographic bulletin boards.

Clarification of the PDB Policy for "Hold for Publication" (HPUB) Entries

To insure that coordinate release for structures with "Hold for Publication" status is consistent with the current PDB Hold Policy, the PDB will place a one-year limit on the length of this hold period. If the citation for a structure is not published within the one-year period, depositors will be given the option to either release or withdraw the deposition.

The one-year limit on the hold period will be applied to new depositions as well as current depositions with "Hold for Publication" status. Depositors with structures currently held for more than one year are currently being contacted and asked to decide if they would like to release or withdraw these entries. ✦

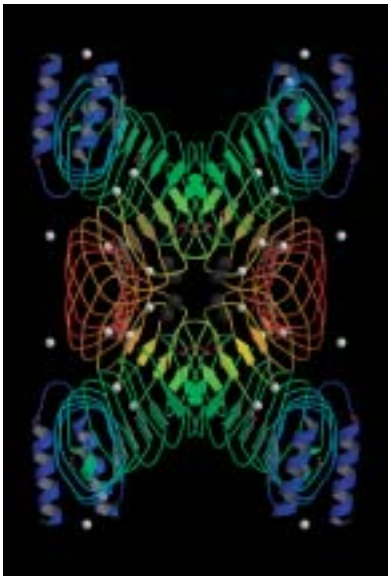
Data Query, Reporting, and Access

PDB Releases XML Data Files for Beta Test

All of the released PDB entries are now available in XML format from the PDB beta FTP site at **ftp://beta.rcsb.org/pub/pdb/uniformity/data/XML/**. Comments are welcomed on these data.

The XML data files have been created by software translation of the mmCIF data files (**ftp://beta.rcsb.org/pub/pdb/uniformity/data/mmCIF/**) created as part of the PDB Data Uniformity Project. The mmCIF data files use the data items defined in the PDB Exchange Dictionary (**deposit.pdb.org/mmcif**). The XML data files conform to an XSD style XML Schema (**deposit.pdb.org/mmcif/dictionaries/ascii/pdbx-v0.904.xsd**) derived from the PDB Exchange Dictionary. As a result, the element and attribute names in the XML data files directly correspond to the item names defined the PDB Exchange Dictionary.

The delivery of PDB data in XML format is the product of a collaboration between the Protein Data Bank Japan (PDBj), the Macromolecular Structure Database (MSD) group at European Bioinformatics Institute (EBI), and the Research Collaboratory for Structural Bioinformatics (RCSB).



The inferred biologically active unit of the YopM cytotoxin

PDB ID: ***1g9u***

A.G. Evdokimov, D.E. Anderson, K.M. Routzahn, D.S. Waugh (2001): Unusual molecular architecture of the Yersinia pestis *cytotoxin YopM: a leucine-rich repeat protein with the shortest repeating unit.* J. Mol. Biol. ***312****, p. 807.*

New Features in Beta Testing

PDB users are encouraged to preview the biological unit files, curated (beta) mmCIF files, redundancy reduction cluster data, and new keyword search that are now in a beta testing phase. Comments on these new features are highly appreciated and may be sent to **notify@rcsb.org**:

- **Biological Unit—Images and Coordinate Files**

The biological unit images and corresponding coordinate files for applicable structures are accessible from the Structure Explorer pages on the PDB Beta Web Site at **beta.rcsb.org/pdb**.

The *View Structure* section of the Structure Explorer offers still ribbon images of the assumed biological unit(s) for structures, where relevant, in addition to static images of the asymmetric unit. Links to the coordinate files that are used to generate the biological unit images are also accessible here, as well as from the Download/Display File section of the Structure Explorer.

- **Curated (Beta) mmCIF Files**

The Download/Display File section of the Structure Explorer pages on the Beta Web Site provides links to view or download the curated mmCIF files. These files include remediated data from the Data Uniformity Project (**www.rcsb.org/pdb/uniformity**). The files follow the latest version of the mmCIF dictionary supplemented by an exchange dictionary developed by the RCSB and the MSD-EBI. This exchange dictionary can be obtained from **deposit.pdb.org/mmcif.**

The curated mmCIF files for a set of query results can be downloaded by selecting the Download Structures or Sequences option from the pull down menu at the top of the Query Result Browser page.

Curated mmCIF files for all PDB structures are available in gzip (.gz) format at **ftp://beta.rcsb.org/pub/pdb/uniformity/data/mmCIF.gz/**. UNIX-compressed versions of these files (.Z) are available at **ftp://beta.rcsb.org/pub/pdb/uniformity/data/mmCIF/**.

- **Redundancy Reduction Cluster Data**

The results of the weekly clustering of protein chains in the PDB are available for beta testing at **ftp://ftp.rcsb.org/pub/pdb/derived_data/NR/**. These clusters are used in the "remove sequence homologs" feature on the PDB web sites. Files that list the clusters and their rankings at 50%, 70% and 90% sequence identity are available. Smaller rank numbers indicate higher (better) ranking. Chains with rank number 1 are ranked as the best representative of their cluster.

The contents of these files and the details of the clustering and ranking are further described at **ftp://ftp.rcsb.org/pub/pdb/derived_data/NR/README** and **www.rcsb.org/pdb/redundancy.html**.

- **New Keyword Search**

A much improved keyword search is now available on the beta web site's home page, SearchLite, and the "Text Search" box on SearchFields. This new search engine (powered by Lucene) queries an index derived from the curated mmCIF files, and should return more accurate search results.

PDB Web Site 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 PDB beta test site, accessible at **beta.rcsb.org/pdb**.

The access statistics are given below for the main PDB Web site at **www.pdb.org**. ✦

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

	DAILY AVERAGE			MONTHLY TOTALS		
MONTH	HITS	FILES	SITES	KBYTES	FILES	HITS
May 03	179,189	136,960	92,164	154,325,959	4,108,804	5,375,686
Apr 03	194,929	148,353	106,378	205,700,581	4,450,604	5,847,881
Mar 03	174,789	132,651	109,763	149,005,949	4,112,207	5,418,463

PDB Outreach

The PDB Talks with Frances C. Bernstein

Frances C. Bernstein

Frances C. Bernstein (Bernstein + Sons) was a member of the PDB team at Brookhaven National Laboratory from 1974 until 1998, where she personally processed or validated most of the entries 15 through 9,000 in the PDB. During this time, she was involved with the creation of the current PDB format and testing of new processing programs for this format as they were created. She has also contributed extensively to the evolving data definitions used in databases and applications programs for macromolecular studies. Ms. Bernstein's expertise has resulted in the creation of tools such as pdb2cif and the enhancement of others including AutoDep and RasMol, and she has freely provided many helpful suggestions to the RCSB.

The PDB recently interviewed Ms. Bernstein regarding her contributions to the resource.

PDB: How did you become involved with the PDB?

Ms. Bernstein: I had two young sons and was working one day per week maintaining and modifying the crystallographic software in the Chemistry Department at Brookhaven National Laboratory. Thomas Koetzle was the sole person involved with the PDB after the untimely death of Walter Hamilton and he asked me if I could work another day per week for the PDB. As the PDB grew and my sons got older, I was able to work more hours per week and then full-time for many years.

PDB: What are some of the interesting milestones of the PDB's growth that you had participated in?

Ms. Bernstein: The original data format was based on Bob Diamond's output format and in 1976 I was involved with the design of what is now known as PDB format, as well as some of the processing and validation software. I set up the original procedures to keep track of what was done in processing and the revision history of each entry. For quite some years, we were only three people at the PDB (two of these part-time) and so we all had our hands in everything that was going on. In addition to processing data, I handled data distribution and the associated record keeping, set up the mailing list, drafted Newsletters, etc. As the volume of data grew and more staff were added, we all became more specialized and my focus was on data processing and the improvement of processing and checking software. The huge increase in growth of the PDB started in the early 1990's and an excellent chart is at **www.rcsb.org/pdb/holdings.html#growth**.

We tried to keep things as seamless as possible for depositors and users so most changes were evolutionary as computing changed. The most dramatic changes came with widespread use of e-mail and then the Internet with web pages and entries immediately available on-line.

PDB: Generally, how did the ways in which data were deposited and validated change during your time with the PDB?

Ms. Bernstein: Originally data were deposited on magnetic tape but a few data sets came on punched cards and one or two were typed from printed coordinates (and it took years to get all the typos out of those). Correspondence was typed by a secretary, processing jobs were submitted on punched cards, and the pace was very much slower. E-mail speeded up interaction with depositors and users enormously and the biggest change came when it was possible to make data freely available to everyone on the Internet. I think it is fair to say that each data deposition had things to check and resolve and we wrote code to check all entries whenever we identified something that could be systematized in software.

PDB: What would you cite as your most rewarding contribution to the PDB?

Ms. Bernstein: I have derived tremendous satisfaction from knowing that my work, and that of all the others who have worked for the PDB over the years, is facilitating some of the enormous advances in medicine, and science in general. On a more specific level, I believe that I helped the depositor community understand that we were striving for the same goal as they were: to provide the best possible data to the user community.

PDB Focus: DNA Day

April 25, 2003 marked the 50th anniversary of the publication of the description of the structure of the double helix. Teachers and students are encouraged to celebrate these historic achievements on this "DNA Day". Many web sites have compiled a wealth of information about this event—a few are listed below.

The National Human Genome Research Institute (**www.genome.gov**) has a variety of teaching resources for National DNA Day at **www.genome.gov/10506367**.

The Nature Publishing Group has compiled the original articles, historical perspectives, and examinations of DNA in medicine, society, and as a biological molecule in "Double Helix: 50 years of DNA" at **www.nature.com/nature/dna50**.

The Cold Spring Harbor Laboratory has a "Celebration of 50 Years of DNA" at **www.dna50.org**, which provides resources and a schedule of events around the world.

The 50th Anniversary Conference (held on April 25th) and other resources from the University of Cambridge are available at **www.admin.cam.ac.uk/univ/science/dna**.

King's College sponsored "A Day of Celebrations" on April 22 with DNA information at **www.kcl.ac.uk/depsta/ppro/dna**.

PBS aired a NOVA feature on the "Secret of Photo 51" about Rosalind Franklin's role in the discovery of the structure of DNA.

Other 50th anniversary events, including articles and meetings, are included at **www.dna50.org.uk**.

An updated and expanded web site for the

Image of a right-handed double-stranded B-DNA helix, from the Nucleic Acid Database (NDB ID: BDL001/PDB ID: 1bna)

Nucleic Acid Database (NDB), the repository of structural information about nucleic acids, was released on April 25. The NDB now has a new look and layout, a greatly revised Atlas, a new database that includes
X-ray and NMR structures, and a new search engine at **ndbserver.rutgers.edu**.

The PDB has many education resources related to nucleic acid structure at **www.rcsb.org/pdb/education.html**, including DNA's turn as Molecule of the Month in November 2001—**www.rcsb.org/pdb/molecules/pdb23_1.html**.

PDB Art at Purdue University

Images from the PDB's "Art of Science" exhibit were on display at Purdue University from April 11 through May 17, 2003. Along with molecular images from Purdue's Structural Biology Center, the PDB installments were featured in "Watson's Crick", a commons area in the Department of Biological Sciences where local artists present their work.

The exhibit was held in conjunction with the university's Spring Fest events.

The PDB would like to see the "Art of Science" travel to other places. If you would be interested in sponsoring this exhibit at your institution, please let us know at **info@rcsb.org**.

Qualifying posters that display this "PDB" sticker will be considered for the PDB Poster Prize at the ACA, AsCA and ECM meetings this summer

Protein Data Bank CD-ROM News

- **First Update Release**

The April 2003 release of the PDB CD-ROM sets, issue 104, is an incremental set of 1,317 experimentally determined structures and 23 models. The structures, on one CD-ROM disk, have been shipped.

Structures re-released for any reason between January and April are included in this update. A list of files that have become obsolete since the last update are included so users can update their set of structures.

July and October issues will only contain the structures released during those quarters. New subscribers will receive the January release and all subsequent updates.

Experimental data —NMR constraints and X-ray structure factors —will be handled in the same manner as the structures—a complete set in January, and incremental updates for the three subsequent quarters. New subscribers will receive the January release and all updates.

Questions should be directed to **info@rcsb.org**. Ordering information is available at **www.rcsb.org/pdb/cdrom.html**.

- **Previous PDB CD-ROM Sets Available**

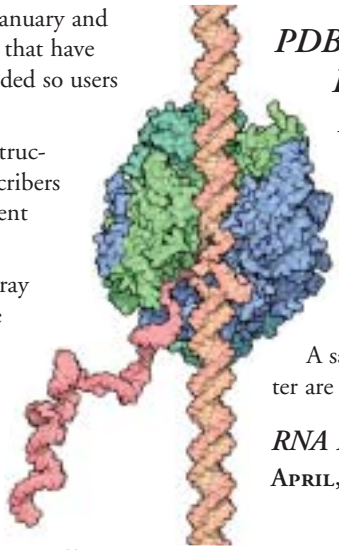
The PDB has extra CD-ROM sets that were issued prior to January 2003. These sets are copies of the FTP archive that were made at the time of the pressing, and include coordinate and experimental data. Software is not included. Multiple sets may be requested.

To request any of these CD-ROM sets, send your address and the number of sets you would like to receive to **info@rcsb.org** or Protein Data Bank, NIST, Mail Stop 8314, Gaithersburg, MD 20899-8314. They will be distributed as long as supplies last on a first come first served basis. The CD-ROM sets will be distributed in reverse order of their date, starting with the most recent.

PDB Poster Prize

The PDB is pleased to announce the initiation of the PDB Poster Prize, which will recognize student poster presentations involving macromolecular crystallography. The prize will be awarded to the best posters by undergraduate or graduate students at each of the meetings of the IUCr Regional Associates—the American Crystallographic Association (ACA), the Asian Crystallographic Association (AsCA), and the European Crystallographic Association (ECM)—as well as at the IUCr Congress itself. Each award will consist of two educational books; this year's prize will be signed copies of *Biochemistry*—Vol I by Donald and Judith G. Voet, and *Introduction to Macromolecular Crystallography* by Alexander McPherson. Winners will be announced on the PDB web site and in the PDB, ACA, and IUCr newsletters.

Details including how to enter can be found at **www.rcsb.org/pdb/poster_prize.html**.



PDB ID: ***1i6h***

A.L. Gnatt, P. Cramer, J. Fu, D.A. Bushnell, R.D. Kornberg (2001): Structural Basis of Transcription: An RNA Polymerase II Elongation Complex at 3.3 Å Resolution. Science ***292****, p. 1876.*

PDB Molecules of the Quarter: RNA Polymerase, Hemoglobin, and Green Fluorescent Protein (GFP)

The Molecule of the Month series explores the functions and significance of selected biological macromolecules for a general audience. These features, written and illustrated by Dr. David S. Goodsell of the Scripps Research Institute, are available at **www.rcsb.org/pdb/molecules/molecule_list.html**.

A sample of the molecules featured during this past quarter are included below:

RNA Polymerase: The RNA Factory

April, 2003—RNA is a versatile molecule. In its most familiar role, RNA acts as an intermediary, carrying genetic information from the DNA to the machinery of protein synthesis. RNA also plays more active roles, performing many of the catalytic and recognition functions normally reserved for proteins. In fact, most of the RNA in cells is found in ribosomes—our protein-synthesizing machines—and the transfer RNA molecules used to add each new

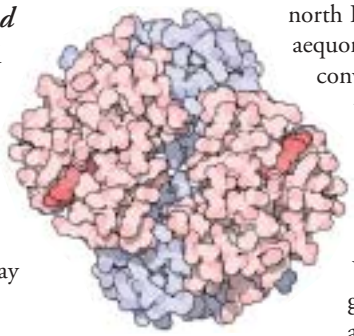
amino acid to growing proteins. In addition, countless small RNA molecules are involved in regulating, processing and disposing of the constant traffic of messenger RNA. The enzyme RNA polymerase carries the weighty responsibility of creating all of these different RNA molecules.

RNA polymerase is a huge factory with many moving parts. The one shown in PDB entry **1i6h** is from yeast cells. It is composed of a dozen different proteins. Together, they form a machine that surrounds DNA strands, unwinds them, and builds an RNA strand based on the information held inside the DNA. Once the enzyme gets started, RNA polymerase marches confidently along the DNA copying RNA strands thousands of nucleotides long.

Further information about RNA polymerase can be found at **www.rcsb.org/pdb/molecules/pdb40_1.html.**

Hemoglobin: Red Blood, Blue Blood

May, 2003—Ever wondered why blood vessels appear blue? Oxygenated blood is bright red: when you are cut, the blood you see is brilliant red oxygenated blood. Deoxygenated blood is deep purple: when you donate blood or give a blood sample at the doctor's office, it is drawn into a storage tube away from oxygen, so you can see this dark purple color. However, deep purple deoxygenated blood appears blue as it flows through our veins, especially in people with fair skin. This is due to the way that different colors of light travel through skin: blue light is reflected in the surface layers of the skin, whereas red light penetrates more deeply. The dark blood in the vein absorbs most of this red light (as well as any blue light that makes it in that far), so what we see is the blue light that is reflected at the skin's surface. Some organisms like snails and crabs, on the other hand, use copper to transport oxygen, so they truly have blue blood.



PDB ID: ***2dhb***

W. Bolton, M.F. Perutz (1970): Three dimensional Fourier synthesis of horse deoxyhaemoglobin at 2.8 Ångstrom units resolution. Nature ***228****, p. 551.*

Hemoglobin is the protein that makes blood red. It is composed of four protein chains, two alpha chains and two beta chains, each with a ring-like heme group containing an iron atom. Oxygen binds reversibly to these iron atoms and is transported through blood. Each of the protein chains is similar in structure to myoglobin (presented in the January 2000 Molecule of the Month), the protein used to store oxygen in muscles and other tissues. However, the four chains of hemoglobin give it some extra advantages.

Further information about hemoglobin can be found at **www.rcsb.org/pdb/molecules/pdb41_1.html**.

Green Fluorescent Protein (GFP): Ready-Made

June, 2003—The green fluorescent protein, shown in PDB entry **1gfl**, is found in a jellyfish that lives in the cold waters of the north Pacific. The jellyfish contains a bioluminescent protein--aequorin--that emits blue light. The green fluorescent protein converts this light to green light, which is what we actually see when the jellyfish lights up. Solutions of purified GFP look yellow under typical room lights, but when taken outdoors in sunlight, they glow with a bright green color. The protein absorbs ultraviolet light from the sunlight, and then emits it as lower-energy green light.

You might be saying: who cares about this obscure little green protein from a jellyfish? It turns out that GFP is amazingly useful in scientific research, because it allows us to look directly into the inner workings of cells. It is easy to find out where GFP is at any given time: you just have to shine ultraviolet light, and any GFP will glow bright green. So here is the trick: you attach the GFP to any object that you are interested in watching. For instance, you can attach it to a virus. Then, as the virus spreads through the host, you can watch the spread by following the green glow. Or, you can attach it to a protein, and watch through the microscope as it moves around inside cells.

Further information about the green fluorescent protein can be found at **www.rcsb.org/pdb/molecules/pdb42_1.html**. ✦

PDB Education Corner

PDB's Education Corner features a different teacher each quarter, offering an account of how he or she uses the PDB to educate students. Educators will find this information useful to inspire their own courses and methods of teaching that incorporate the PDB.

This quarter's column is by Prof. Tim Herman, Director of the Center for BioMolecular Modeling at the Milwaukee School of Engineering:

The Center for BioMolecular Modeling (CBM) at the Milwaukee School of Engineering uses rapid prototyping technologies to produce physical models of proteins and other molecular structures based on atomic coordinates obtained from the Protein Data Bank. These physical models are used both by researchers, who find them useful as "thinking tools" and by students who are just beginning to explore the molecular world.

The CBM directs science outreach and professional development programs targeted to both high school science teachers and undergraduate educators (**www.rpc.msoe.edu/cbm**). At the high school level, an NIH-funded SEPA (Science Education Partnership Award) allows us to offer a two-week summer course entitled "Genes, Schemes and Molecular Machines". In this

A physical model of the p53 tumor suppressor protein from PDB entry ***1tsr****.*

PDB ID: ***1tsr***

Y. Cho, S. Gorina, P.D. Jeffrey, N.P. Pavletich (1994): Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. Science ***265****, p. 346.*

course, teachers are shown how physical and computer-generated models of proteins can be used to make the molecular world real for their students. Using a recently modified version of RasMol (RP-RasMol), teachers are directly involved in the design and construction of the physical models. In the summer of 2002, a team of six high school science teachers produced the first-ever physical models of the ribosome, based on recently deposited atomic coordinate data (**www.rpc.msoe.edu/sepa/HandsOn.pdf**) More recently, we have begun to involve high school students in the design of physical models through our SMART Team program. SMART Teams (Students Modeling A Research Topic) consist of a high school teacher who has participated in our summer course and a group of 3-5 students who work with a local research lab to produce a physical model of the protein under investigation in the lab (**www.rpc.msoe.edu/cbm/sepa/Nature_article.pdf**). Our first SMART Team, known as Team Anthrax, designed and constructed physical models of the three proteins involved in anthrax pathogenesis in the months immediately following the anthrax attacks in the fall of 2002. Currently, seven SMART Teams are working with researchers in Wisconsin. Plans are underway to replicate this program in other areas including San Diego, upstate New York and Salt Lake City. [For information about how to participate in a SMART Team, contact Prof. Herman at **herman@msoe.edu**.]

At the undergraduate level, the CBM works with undergraduate educators to explore ways in which physical models can enhance the use of molecular visualization tools. With support from an NSF-CCLI award, the Center plans to launch a Summer Modeling Institute at which undergraduate faculty will have access to our physical modeling technologies to design and produce models that will be used in courses on their local campuses. A Model Lending Library has been created to allow any undergraduate educator to borrow models resulting from this project for use in their classroom.

Molecular models for science education—including a DNA Construction Kit—can be obtained from 3D Molecular Designs (3DMD), a recent spin off of the CBM (**www.3dmoleculardesigns.com**). ✦

Related Links: Software

The PDB's Software List page offers links to free programs and utilities, such as RCSB-developed software, mmCIF tools, and resources supporting crystallography, NMR, structure analysis and verification, modeling and simulation, and molecular graphics. A few of these links include:

ADIT—Workstation Version (alpha release)
pdb.rutgers.edu/mmcif/ADIT/index.html
A package for editing and checking structure data entries

OpenMMS Toolkit—openmms.sdsc.edu
A suite of Java source code that includes an mmCIF parser, RDBMS loader, XML translator, and Corba server

BioEditor—bioeditor.sdsc.edu
A tool for creating and viewing dynamic, formatted structure annotations; for Windows

PDB Job Listings

PDB career opportunities are posted at **www.rcsb.org/pdb/jobs.html**. The current available openings are:

Systems/Web/Database Programmer

The Protein Data Bank has a position open for a systems programmer and database administrator who will be responsible for maintenance of software, website, and supporting databases associated with a digital library of the Protein Data Bank (PDB) archive. The duties will include maintenance of web application software and database tools to manage the historical record of the PDB. The ideal candidate will have a strong background in UNIX, website, and database administration. Proficiency in programming in a UNIX/LINUX environment with languages such as JAVA, C/C++, or PERL is required. Experience in the following is highly desirable: SQL, JDBC, and application development using ORACLE or DB2. This position is at the PDB site at the Center for Advanced Research in Biotechnology in Rockville, MD. Please send resume to Dr. Gary L. Gilliland at **gary.gilliland@nist.gov**.

Data Archive/Digital Librarian

The Protein Data Bank has a position open for a data archivist who will organize and manage the historical record of the Protein Data Bank. The ideal candidate will have a strong background in library science and the physical sciences. Experience in the management of large document and media collections is required. The position is at the PDB site at the Center for Advanced Research in Biotechnology in Rockville, MD. Please send resume to Dr. Gary L. Gilliland at **gary.gilliland@nist.gov**.

System and Applications Programmer

The Protein Data Bank at Rutgers University has a position open for an applications programmer to support and develop software for data processing operations at the Protein Data Bank. Programming areas include: macromolecular structure analysis and validation, molecular graphics, web application development, distributed object and relational database applications, and general scientific programming. Experience developing and maintaining object oriented software on UNIX platforms is required. Experience in the following is highly desirable: C/C++, JAVA, and Corba. Please send resume to Dr. Helen Berman at **pdbjobs@rcsb.rutgers.edu**.

Biochemical Information Specialist

The Protein Data Bank at Rutgers University has a position open for a Biochemical Information Specialist to curate and standardize macromolecular structures for the Protein Data Bank. A background in biological chemistry, as well as some experience with UNIX-based computer systems, is required. Experience in crystallography and/or NMR spectroscopy is a strong advantage. The successful candidate should be well-motivated, able to pay close attention to detail, and meet deadlines. This position offers the opportunity to participate in an exciting project with significant impact on the scientific community. Please send resume to Dr. Helen Berman at **pdbjobs@rcsb.rutgers.edu**.

PDB Project Team Leaders

The overall operation of the PDB is managed by the PDB Project Team Leaders. Technical and scientific support are provided by the 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 Coordinator
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

PDB Members

Rutgers

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

Anthony Adelakun
anthony@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

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

Tania Rose Posa
tania@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

David Archbell
dave@sdsc.edu

Dr. Peter Arzberger
parzberg@sdsc.edu

Bryan Banister
bryan@sdsc.edu

Tammy Battistuz
tammyb@sdsc.edu

Dr. Wolfgang F. Bluhm
wbluhm@sdsc.edu

Dr. Nita Deshpande
nita@sdsc.edu

Dr. Ward Fleri
ward@sdsc.edu

Dr. Douglas S. Greer
dsg@sdsc.edu

Jeff Ott
jott@sdsc.edu

David Padilla
dpadilla@sdsc.edu

Thomas Solomon
tsolomon@sdsc.edu

Wayne Townsend-Merino
wayne@sdsc.edu

Peggy Wagner
wagner@sdsc.edu

CARB/NIST

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

Dr. T.N. Bhat
bhat@nist.gov

Phoebe Fagan
phoebe.fagan@nist.gov

Jeremy Kohansimeh
jkohansi@nist.gov

Dr. Veerasamy Ravichandran
vravi@nist.gov

Michael Tung
michael.tung@nist.gov

Padma Priya Paragi Vedanthi
padmap@nist.gov

Statement of Support

The PDB is supported by funds from the National Science Foundation, the Office of Biological and Environmental Research at the Department of Energy, and two units of the National Institutes of Health: The National Institute of General Medical Sciences and the National Library of Medicine, in addition to resources and staff made available by the host institutions.

PROTEIN DATA BANK
www.pdb.org

U.S. Department of Commerce
National Institute of Standards and Technology
100 Bureau Drive, Stop 8314
Gaithersburg, MD 20899-8314

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

Return Service Requested

FIRST-CLASS MAIL
POSTAGE & FEES PAID
NIST
PERMIT No. G195