Fall 2005 • Number 27

Published quarterly by the Research Collaboratory for Structural Bioinformatics Protein Data Bank

RCSB PDB PROTEIN DATA BANK NEWSLETTER

Weekly RCSB PDB news is available online at www.pdb.org

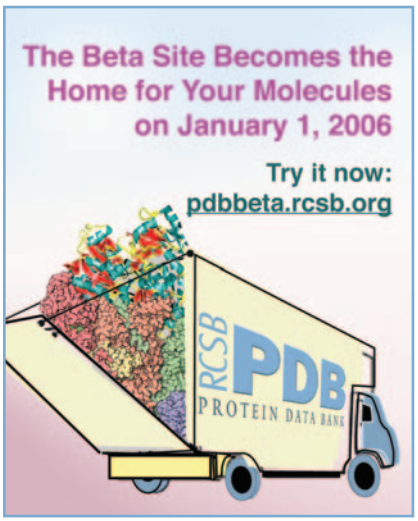
Contents

SNAPSHOT: OCTOBER 1, 2005
32823 released atomic coordinate entries

MOLECULE TYPE		EXPERIMENTAL TECHNIQUE	
29964	proteins, peptides, and viruses	28040	diffraction and other
1504	nucleic acids	4783	NMR
1342	protein/nucleic acid complexes	18200	structure factor files
13	carbohydrates	2647	NMR restraint files

PARTICIPATING RCSB MEMBERS:
RUTGERS: rutgers.rcsb.org
SDSC/UCSD: www.pdb.org
E-mail: info@rcsb.org
FTP: ftp.rcsb.org
The RCSB PDB is a member of the wwPDB **(www.wwpdb.org)**

Message from the RCSB PDB



To provide a powerful portal for studying the structures of biological macromolecules and their relationships to sequence, function, and disease, the RCSB has enhanced and revised the RCSB web and FTP sites.

We encourage you to test and review **http://pdbbeta.rcsb.org** & **ftp://ftpbeta.rcsb.org** before these resources replace the versions currently in production at **www.pdb.org** and **ftp.rcsb.org** on January 1, 2006.

The enhanced web design offers improved navigation for easily locating the resources and tools offered by the RCSB PDB. An always-present left hand menu makes features and resources related to structural genomics, education, and software easily attainable.

The search tab offers different ways of accessing the RCSB PDB database, including a new method for "browsing" through structures grouped in categories related to disease, molecular function, biochemical process, or cellular location.

The reengineered database is based upon the PDB data that has been remediated and standardized. Tools to examine these data include: improved ligand searching; a clear distinction between the reported primary and derived data; and the integration of external data resources currently not available on the production site.

A searchable help system with a glossary and user guide provides detailed information for accessing the website, database, and for understanding PDB data. A narrated presentation in Flash is linked from the beta site homepage to guide users through searching, navigating, generating reports, visualizing structures, and browsing PDB data.

The structural genomics portal offers target summary reports for each center, databases that track the progress of protein studies (TargetDB and PepcDB), and a tool to explore the distributions of functions found among structural genomics structures, PDB structures, genomes, and homology models.

Thanks to everyone who has contributed to the development of this resource through their feedback. Additional comments should be sent to **betafeedback@rcsb.org.**

RCSB PDB Beta Site:
http://pdbbeta.rcsb.org/

RCSB PDB Beta FTP Site:
ftp://ftpbeta.rcsb.org

Tutorial for using the new site:
http://pdbbeta.rcsb.org/pdbstatic/tutorials/tutorial.html

Data Deposition and Processing

PDB Focus: Releasing HOLD and HPUB structures

Structure data can be held for a maximum period of one year between the deposition and release of any entry, even if the paper associated with an HPUB structure is still unpublished.

Data can always be released before the paper is published. The citation information can then be updated at a later date.

When a structure has been on hold for a year, we ask the author to release or withdraw the structure. Withdrawn entries can be redeposited at a later time, at which point a new PDB ID will be issued.

We try to release HPUB files as close to the publication date as possible. The PDB receives publication dates and citation information from some journals. For other journals, the PDB scans the literature for publication information (using the deposited author names and deposited article title). We also greatly appreciate having citation information sent to **deposit@rcsb.rutgers.edu**.

PDB Focus: Tips for depositing multiple, related structures using ADIT

When depositing many structures that are related to one another, there are a few ways of making the ADIT deposition process simpler:

- Structures solved using X-ray crystallography should be prepared using pdb_extract before using ADIT. This will minimize manual typing and save time during deposition. pdb_extract takes information about data collection, phasing, density modification, and the final structure refinement from the output files and log files produced by the various applications used for structure determination. The collected information is organized into a file ready for deposition using ADIT. Information duplicated in all entries (author name, citation information, protein names, etc.) can be compiled into a single text file to run with pdb_extract for each entry. After pdb_extract has combined all the available information into a single file for each structure, ADIT can be used for quick deposition.

- A similar tool is being developed for structures solved by other experimental methods. For these structures, deposit one representative structure following the instructions provided at deposit.pdb.org. Then write to **deposit@rcsb.rutgers.edu** to let us know about the other related entries. Once the first entry has been annotated, processed and finalized, it can be used as a template for subsequent depositions. For each structure, replace the coordinates and update the information in the header section of the PDB file as necessary to prepare the related files for deposition.

- If the structures have ligands, drugs or inhibitors bound to them, please check Ligand Depot and match the 3-letter code in the file to the one used in the chemical component dictionary. If the ligand is not present in the dictionary, please email detailed information (complete chemical name, 2D figure showing connectivity, bond order and sterochemistry) along with the RCSB and PDB IDs of associated entries to expedite the processing of these files.

PDB Deposition Statistics

As of October 1, 4894 experimentally-determined structures have been deposited to the PDB archives this year.

The entries were processed by wwPDB team members at RCSB, MSD-EBI, and PDBj. Of the structures deposited, 69% were deposited with a release status of HPUB; 17% with REL; and 14% with HOLD.

81% of these entries were determined by X-ray crystallography; 16% were determined by NMR. 81% were deposited with experimental data. 56% released the sequence in advance of the structure's release.

Data Query, Reporting, and Access

PDB Focus: Beta FTP Site Organized by Experimental Type

The RCSB has introduced a new ftp site designed to address the increasing number and diversity of structure entries. Entries within **ftp://ftpbeta.rcsb.org** are organized by the method of structure determination. This new ftp structure distinguishes entries obtained from X-ray and neutron diffraction, NMR, electron microscopy, or theoretical modeling techniques.

Within each structure determination category, there are directories containing files for entries in mmCIF, PDB, and XML formats; biological unit files; and the structure factor or restraint files. Individual entries are stored in a collection of 2-character code subdirectories corresponding to the middle characters of the 4-character PDB identifier.

All entries in the new ftp site are compressed using the GNU gzip utility. For example, the PDB format file for entry 4HHB is found at **ftp://ftpbeta.rcsb.org/pub/data_by_method/x-ray_neutron_methods/entries/pdb/hh/pdb4hhb.ent.gz**.

This beta ftp site is updated weekly in conjunction with the current ftp site, **ftp://ftp.rcsb.org**. Both ftp sites will be maintained through December 2005. After this time, only the new organization will be supported.

For comments about the beta FTP send mail to **betaftp@rcsb.rutgers.edu.**

Website Statistics

	DAILY AVERAGE		MONTHLY TOTALS			
MONTH	HITS	FILES	SITES	KBYTES	FILES	HITS
Sep 05	287,624	208,009	144,490	230,808,396	6,032,283	8,341,119
Aug 05	203,087	149,599	99,488	241,077,817	4,487,998	6,092,621
Jul 05	215,619	160,129	100,218	272,765,261	4,803,890	6,468,598

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

Data Distribution on DVD

As of 2005, the PDB archives will be made available as an annual DVD product. The CD-ROM product has been discontinued.

The PDB has historically distributed coordinate data through the mail; starting in 1990, the data were produced on CD-ROM. In addition to the rapid increase in the number and size of structures in the archives, the number of formats to distribute has also increased (to PDB, mmCIF, and PDBML/XML). Currently, it would take over 50 disks to distribute the archives on CD-ROM.

To minimize the difficulty of producing (and keeping up with a large number of CDs), the RCSB PDB began distributing an incremental set of data each quarter in April 2003.

Starting in January 2005, **ftp://snapshots.rcsb.org/** was created to hold time-stamped yearly snapshots of the PDB archives. This was done as part of a wwPDB initiative (**wwpdb.org**). The first snapshot directory contains the exact and complete contents of the FTP archives as it appeared on January 6, 2005. These data, along with data in XML/PDBML format, were also made available as an 8 DVD release.

To order a DVD set, please email your postal address to **cddvd@rcsb.rutgers.edu**. New orders will be filled in the order that they are received until supplies are exhausted. Please allow several weeks for delivery.

Outreach and Education

RCSB PDB at IUCr: wwPDB Booth, Presentations, and More

The new RCSB PDB website, along with tools for deposition, was demonstrated at this year's XX Congress & General Assembly of the International Union of Crystallography (IUCr, August 23-31 in Florence, Italy). The exhibit was part of the exhibit stand designed with wwPDB partners MSD-EBI and PDBj.

Philip E. Bourne at the wwPDB Roundtable.

Several talks were also presented during the Congress, including lead annotator Kyle Burkhardt's "Deposition and Validation using RCSB PDB Tools" as part of the CCP4 Workshop; director Helen M. Berman's "How the RCSB validates PDB Structures" during the microsymposium Improving Structures Using Bioinformatics; and software leader Zukang Feng's "Applications at the RCSB Protein Data Bank" during the COMCIFS Open Commission Meeting on the Current Status and Future Prospects of CIF.

Helen M. Berman gave an overview of the wwPDB at the Round Table on Data Mining chaired by Judith L. Flippen-Anderson. wwPDB representatives Kim Henrick (MSD-EBI), Haruki Nakamura (PDBj), and Philip E. Bourne (RCSB PDB) then showed how their different interfaces can be used to mine the data within the PDB archives.

The RCSB PDB Poster Prize was awarded to Sasa Jenko Kokalj for "Proline isomerisation in stefin B: A crucial step towards amyloid fibril formation" by Sasa Jenko Kokalj, Gregor Guncar, Eva Zerovnik, and Dusan Turk (Department of Biochemistry and Molecular Biology, Jozef Stefan Institute, Ljubljana, Slovenia). The judging committee was comprised of Maria-Arménia Carrondo (Chair), Carlos Frazao, Ramakumar Suryanarayanarao, Xiao-Dong Su, Edward Mitchell, and Marius Jaskolski.

Poster Prize winner Sasa Jenko Kokalj

mmCIF articles published in new edition of the International Tables for Crystallography (Volume G)

Several articles describing the macromolecular Crystallographic Information File format (mmCIF) and how it is used at the RCSB PDB are in the recently-published *International Tables for Crystallography Volume G: Definition and Exchange of Crystallographic Data*.

Included in this reference guide for programmers, data managers, and crystallographers are in-depth articles that will aid in the design of interoperable computer applications, including "Classification and use of macromolecular data", "Macromolecular dictionary (mmCIF)", "Specification of the Crystallographic Information File", "The Protein Data Bank exchange data dictionary", "The use of mmCIF architecture for PDB data management", and "Specification of a relational Dictionary Definition Language (DDL2)".

This volume was edited by S.R. Hall, and B. McMahon, and published for the International Union of Crystallography by Springer (Dordrecht, The Netherlands) in 2005. Ordering information is available at **journals.iucr.org/iucr-top/it/itg/itg.html**.

Molecule of the Month Features Available as Individual PDFs

The Molecule of the Month series presents short accounts that describe selected molecules from the PDB. These online chapters are now available as downloadable PDFs that can be easily shared, printed, or incorporated into classroom lessons. Users can download individual features by going to the Molecule of the Month contents page and clicking on the PDF link displayed after the molecule title.

Each installment includes an introduction to the structure and function of the molecule and relates the molecule to human health and welfare. Suggestions for viewing structures on the RCSB PDB website and for additional reading are also provided. Produced and illustrated by David S. Goodsell since 2000, the Molecule of the Month is a proven resource for the classroom.

Previous features have explored structures such as actin, collagen, DNA, green fluorescent protein, ribosome, and transfer RNA.

Molecules of the Quarter:

TATA-Binding Protein, Neurotrophins, Cholera Toxin

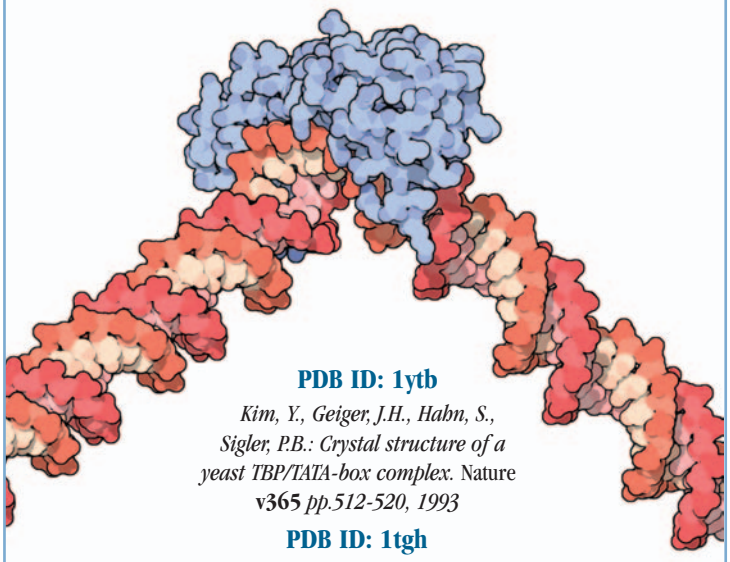
The **MOLECULE OF THE MONTH** *series explores the functions and significance of selected biological macromolecules for a general audience. This quarter, TATA-Binding Protein, Neurotrophins, and Cholera Toxin were highlighted. The July 2005 edition is exerpted here; the full Molecule of the Month features are available from the RCSB PDB website.*

TATA-Binding Protein (July)

The enzyme RNA polymerase performs the delicate task of unwinding the two strands of DNA and transcribing the genetic information into a strand of RNA. But how does it know where to start? Our cells contain 30,000 genes encoded in billions of nucleotides. For each gene, the cell must be able to start transcription at the right place and at the right time.

Specialized DNA sequences next to genes, called promoters, define the proper start site and direction for transcription. Promoters vary in sequence and location from organism to organism. In bacteria, typical promoters contain two regions that interact with the sigma subunit of their RNA polymerase. The sigma subunit binds to these DNA sequences, assists the start of transcription, and then detaches from the polymerase as it continues transcription through the gene. Our cells have a far more complex promoter system, using dozens of different proteins to ensure that the proper RNA polymerase is targeted to each gene. The TATA-binding protein is the central element of this system.

Our protein-coding genes have a characteristic sequence of nucleotides, termed the TATA box, in front of the start site of transcription. The typical sequence is something like T-A-T-A-a/t-A-a/t, where a/t refers to positions that can be either A or T. Surprisingly many variations on this theme also work, and one of the challenges in the study of transcription is discovering why some sequences work and others don't. The TATA-binding protein (sometimes referred to as TBP) recognizes this TATA sequence and binds to it, creating a landmark that marks the start site of transcription. When the first structures of TATA-binding protein were determined, researchers discovered that TATA-binding protein is not gentle when it binds to DNA. Instead, it grabs the TATA sequence and bends it sharply as seen on the right in PDB entries **1ytb, 1tgh** and **1cdw**.



PDB ID: 1ytb

Kim, Y., Geiger, J.H., Hahn, S., Sigler, P.B.: Crystal structure of a yeast TBP/TATA-box complex. Nature **v365** *pp.512-520, 1993*

PDB ID: 1tgh

Juo, Z.S., Chiu, T.K., Leiberman, P.M., Baikalov, I., Berk, A.J., Dickerson, R.E.: How proteins recognize the TATA box. J.Mol.Biol. **v261** *pp.239-254, 1996*

PDB ID: 1cdw

Nikolov, D.B., Chen, H., Halay, E.D., Hoffman, A., Roeder, R.G., Burley, S.K.: Crystal structure of a human TATA box-binding protein/TATA element complex. Proc.Natl.Acad.Sci.USA **v93** *pp.4862-4867, 1996*

PDB Education Corner:

Miriam Rossi, Vassar College

Q: What is your Structural Chemistry & Biochemistry course about? Tell us about its origin, goals and objectives.

A: Today, interdisciplinary research and areas of study are the norm in the sciences. The borderlines of research in chemistry, biology, physics, geology are vanishing, and increasingly molecular interactions and chemical transformations are found to be at the heart of biology, biochemical and biomedical phenomena as well as understanding the behavior of new materials, material sciences and geological sciences. For example, many diseases and their treatment are molecular in nature and medicinal chemistry and pharmaceutical companies have large components dedicated to drug design research that has structural information as its basis.

This course was developed to make students aware that the same fundamental principles that govern the molecular architecture for metals and small salt compounds also explain the structures of macromolecules and molecular assemblies, such as viruses. General texts in chemistry, biology,

DR. MIRIAM ROSSI *has been at Vassar since 1982 after she worked as a Research Associate at The Institute for Cancer Research of The Fox Chase Cancer Center in Philadelphia. She received her Ph.D. at The Johns Hopkins University. Her work is concerned with the relationship between the structure and function of molecules, mainly those having biological activity. These include natural plant products that show anti-tumor activity as well as others that are active against some of the proteins in HIV. The technique she uses is single crystal X-ray crystallography, and she is co-author of a leading text in this area. She has received grants from the National Science Foundation, the Petroleum Research Fund of the American Chemical Society, and the Camille and Henry Dreyfus Foundation. Besides the U.S., she has taught courses in Australia, Italy, and most recently under the auspices of the Rotary Foundation, in Chile. Her work has appeared in the Journal of Medicinal Chemistry, Inorganic Chemistry, Organometallics, Archives of Biochemistry and Biophysics, and the Journal of Natural Products, among many others. Her teaching interests include general chemistry, inorganic chemistry, and structural chemistry, and she particularly enjoys conducting research with undergraduates.*

This interview about one of her courses will also appear in a Vassar College website that highlights courses that incorporate information technology in teaching activities at **computing.vassar.edu/news/faculty/facultyfocus.**

geology, biochemistry and solid-state physics are full of molecular structure representations obtained from the three-dimensional atomic coordinates determined by X-ray diffraction. Unfortunately, there is difficulty in interpreting these spatial arrangements, especially as the structures of increasingly larger molecules become available. This problem is not new. In 1925, Sir William Bragg wrote in the Preface to his book, *Concerning the Nature of Things*[1]: "…There are some who think this difficulty is incurable, and that it is due to the want of some special capacity, which only a few possess. I am persuaded that this is not the case: we should have nearly as much difficulty in grasping events in two dimensions as in three were it not that we can so easily illustrate our two-dimensional thoughts by pencil and paper. If one can turn over a model in one's hand, an idea can be seized in a mere fraction of the time that is required to read about it, and a still smaller fraction of the time that is required to prepare the description."

The goals of this course involve familiarizing students with basic concepts of molecular structure and geometry, chemical bonding and intermolecular interactions; to introduce students to X-ray crystallographic methods for determining molecular structure; how to read a crystal structure paper; to study the structures of molecules of chemical and biological interest; interpreting the complex images that accompany structural papers. The overall aim is to see how the molecular structure is one of the determinant features responsible for chemical and/or biological activity. It is an advanced level course open to chemistry and biochemistry majors or students declaring a chemistry or biochemistry correlate sequence (frequently called a "minor" subject area).

I am able to attain these goals by using different computer programs and databases that can be accessed in a computer classroom containing 15 desktop computers. An Academic Computing Consultant who is a science computation specialist maintains this computer classroom; his presence has ensured the successful outcome of this course.

Q: What were the technologies used and how did they change or enhance your course?

A: I use two main computer tools: the PDB, a large repository for the processing and distribution of 3-D biological macromolecular structure data, and the Cambridge Crystallographic Structural Database (CCSD, **www.ccdc.cam.ac.uk**), which is a database of bibliographic, chemical and crystallographic information for organic molecules and organo-metallic compounds. The CCSD is not freely available, but fortunately the college has a campus-wide license for it. The RCSB PDB is freely and widely available on the web.

The addition of these tools has made a huge difference in teaching this course. For example, the various visualization capabilities available in the two databases make the content easier to teach, since many structural features become self-evident when they are viewed. The ability to manipulate, rotate and edit structures allows instructors to convey these structural "rules" that are not easy to visualize or understand otherwise. While this is true for all three-dimensional structural data, it is essential for understanding macromolecular data; the PDB is an indispensable resource to achieve this objective.

Q: How have your students responded to your use of technology?

A: Students are very receptive to acquire information through interactive and visual experience. I hear students praise the software all the time; they like being able to see and manipulate molecules; it is intuitive, easy and fun. Frequently, in courses where the concepts of molecular shape and chemical properties derived from molecular shapes are introduced, for example, organic chemistry, students use model kits as a teaching aid. While this is a useful exercise, being able to rotate the structure on the screen and see how a molecule interacts with others is especially valuable. The use of these structural databases reinforces material that they learn about in their textbooks. It becomes clear that the connections between atoms to make molecules and how molecules are grouped together to make molecular assemblies all have similar foundations. Geometrical details can be calculated easily and displayed. These resources have made describing molecular features much easier, and since students can access the PDB online at anytime students can access information for problem sets at their convenience.

The use of interactive computer graphics software is especially indispensable when teaching macromolecular structure. The level of complexity in macromolecular structure is astounding. Proteins can have diverse shapes or motifs that make up their final 3-D structure. It is here that the PDB becomes such a useful teaching tool: it permits students to look at the complex protein or other large macromolecular assembly using options that permit a different perspective by zooming in to a particular site and rotating the structure. It becomes possible to view the same structural motif in many proteins, allowing for interesting discussion to take place in the classroom: the PDB reinforces material that they learn about in their textbooks (The text I use is *Introduction to Protein Structure*[2], 2nd Edition by Branden and Tooze). Geometrical details can be calculated easily and displayed. The Molecule of the Month feature is outstanding and almost forms the basis for the macromolecular part of this course; it features an in-depth

[1] Dover Phoenix Editions

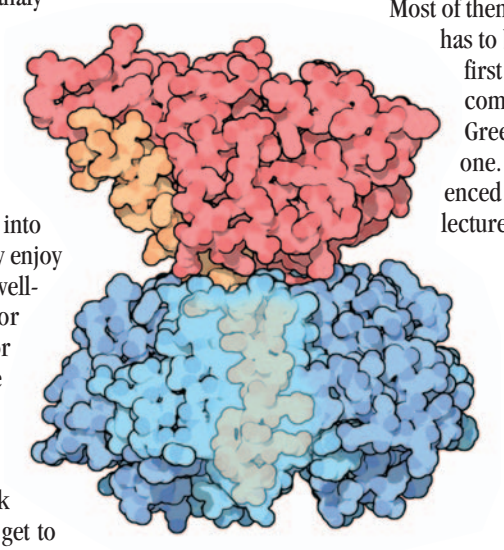
[2] Garland Publishing

look at a collection of well-studied molecules. It provides a well-written introduction to a molecule, its biological importance, how the structural features are utilized by the molecule to attain its function, and highlights the use of common structural motifs by related compounds. My students use this feature as a starting point in many of their in-depth analyses that are required for homework assignments. What I have found is that because the RCSB PDB is easy to use, students literally can spend hours manipulating these complex structures. And, every instructor knows that the more time a student spends on trying to understand course material, the more attracted they become to the subject.

Every year, I try to shape the course contents taking into account the particular interests of my students. They really enjoy being able to use these databases to find and visualize well-known structures they learn about in other contexts; for example, they can see how toxins such as cholera or anthrax toxins behave at an atomic level, as well as the common cold and influenza viruses. Some of the topics we discuss are how DNA modifies its shape as it interacts with a variety of proteins; the mathematical description of the morphology of virus particles, etc. Since students pick a topic for an end-of-the-year presentation, sometimes I get to see really interesting things happening in the classroom. I remember that a student who was an Art History major gave a presentation comparing the structural features that provide stability to molecules to those that are commonly used when building macroscopic structures such as bridges or buildings.

Q: What were the challenges you faced when teaching this course?

A: Even though these computer programs are very well built, and easy to use, there is always a learning curve that students must go through. Most of them are very quick learners, but with a few one has to be more patient and guide them through the first couple of sessions. For example, seeing complex protein structural patterns, such as the Greek key motif, need to be explained one-on-one. A good idea would be having an experienced student assistant helping students during lecture time.



Professor Rossi uses the Molecule of the Month feature in her classroom.

Shown here:

Cholera Toxin (September 2005)
PDB ID: 1xtc

Zhang, R.G., Scott, D.L., Westbrook, M.L., Nance, S., Spangler, B.D., Shipley,G.G., Westbrook, E.M.: The three-dimensional crystal structure of cholera toxin.
J.Mol.Biol. **v251** *pp.563-573, 1995*

PDB Community Focus:

William L. Duax, International Union of Crystallography and Hauptman-Woodward Medical Research Institute

Q: You were a prolific small molecule crystallographer (with ~250 entries in the Cambridge Structural Database, CSD). Why did you decide to shift your research efforts to focus on macromolecules?

A: There were two main reasons, one scientific and the other practical. The majority of our structures in the CSD are steroids. We were trying to understand how subtle changes in the shape and intermolecular interactions of steroids influenced their control of sugar metabolism, salt balance, sexual development, and fertility. Although we were able to develop empirical models for steroid structure and function, it became clear that we needed the three-dimensional structures of their protein targets to adequately test our models. Therefore, we developed collaborations with biochemists who were isolating steroid hormone receptors and enzymologists studying the biosynthesis and metabolism of steroids. We were fortunate that our first protein structure determination was of an ancient member of the short-chain oxidoreductase (SCOR) enzyme family. SCORs control the balance of active and inactive steroids involved in normal biological processes. Using that first structure, Debashis Ghosh was able to propose a model for the mechanism of action of the enzyme. This model provided the basis for understanding the activities of dozens of additional SCORs in the PDB and thousands of homologous proteins in the Swiss-Prot/TrEMBL databases (**www.expasy.org/sprot**). The practical reason for shifting our focus to macromolecules was that we had to go where the money was. When we published our first steroid structure in 1969, there were only a few hundred crystal structures of organic compounds in the literature. In the 1960s and 1970s, small molecule crystallography enjoyed significant support from the US National Institutes of Health (NIH). Currently, there are over 325,000 entries in the CSD, and NIH support has gradually shifted to macromolecular crystallography. Today, the macromolecular crystallographic community is facing a challenge similar to when small molecule crystallographers were unwisely relegated to the role of service crystallographers in the 1980s.

DR. WILIAM L. DUAX *received his Ph.D. in physical chemistry from the University of Iowa in 1967. He came to the Hauptman-Woodward Medical Research Institute in 1968 where he was head of the Molecular Biophysics Department from 1970-88, Research Director from 1988-93, Executive Vice President 1993-1999, and is currently H.A. Hauptman Distinguished Scientist. He is a Professor of the Department of Structural Biology at the State University of New York at Buffalo and the Chief Executive Officer of the American Crystallographic Association. He was a Fulbright Scholar in Yugoslavia in 1987, and received the Distinguished Scientist award from the Clinical Ligand Assay Society in 1994 and an Honorary Doctoral Degree from the University of Lodz Poland in 1999. He recently ended his term as President of the International Union of Crystallography.*

Dr. Duax uses X-ray crystallographic analysis to determine the structure of biologically active molecules and correlates structures with biological activity. Early in his career he determined the structures of hundreds of steroids, published a two volume Atlas of Steroid Structure and developed an empirical model for steroid receptor binding and hormone action, (the A-ring binding/D-ring acting model) that proved valuable in the design of antiprogestational agents. More recently, together with Debashis Ghosh, he began a series of studies of steroid dehydrogenases. They determined the first crystal structure of a short chain oxidoreductase (SCOR) enzyme, elucidated of the molecular mechanism of action of SDHs, the basis for inhibition of 11β-HSD by glycyrrhizic acid and the role of licorice in hypertension.

Recently he has begun combining information from the 3D structures of SCOR enzymes with sequence analysis of all putative SCOR genes in the gene bank to predict fold, function, cofactor, and substrate for 6000 genes. He is also engaged in a genomics project to trace the origin and evolution of the genetic code and the evolution of the amino acid composition of proteins.

Crystallographers are well-equipped to make sense of the mass of data that the genome project has brought to light. Anticipating yet another sea change in federal funding, we have begun to utilize the explosion of macromolecular sequence data in combination with macromolecular structural data in an effort to try to predict the structure function relationship and ligand binding of thousands of hypothetical gene products.

Q: As a relatively new user of the RCSB PDB, have you found it easy to use and beneficial in your current research?

A: The PDB is absolutely essential to our current efforts to predict the structure and function of the 6,000 putative SCOR enzymes in the Swiss-Prot/TrEMBL databases. I find it very easy to access and use the PDB. This is all the more remarkable because I am totally incompetent when it comes to using computers. We use the PDB to examine the details of a billion years of molecular evolution and interactions between proteins and their cofactors and substrates in search of patterns that have predictive power. We also use the PDB to examine the nature of the interfaces among monomers in multimeric structures that are often their active forms. These studies give us insight into protein-protein interactions and allosteric behavior.

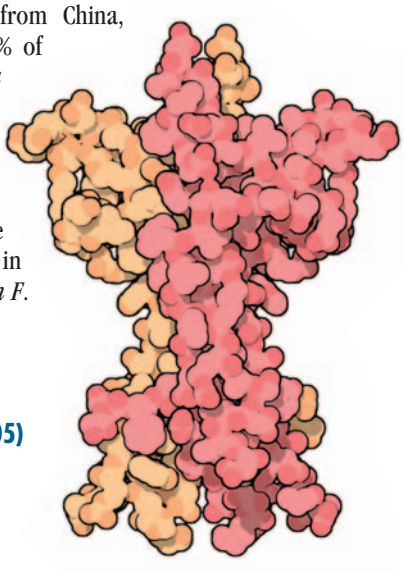
TargetDB (**targetdb.pdb.org**) is another reason that the RCSB PDB is critical to our research program. We use it as a resource for information on the most recently determined SCOR proteins as well as to identify new candidates to test predictions based on our proteomic research.

Q: The RCSB PDB has been involved in the creation, maintenance, and extension of the mmCIF dictionary since its commission by the IUCr in the early 1990s. The collaboration was further strengthened this year when the IUCr launched Acta Crystallographica Section F for the rapid publication of communications on structural genomics, protein structure and crystallization. What is the IUCr's perspective on this collaboration?

A: The collaboration between the RCSB PDB and the IUCr has been vitally important and extremely productive. One of the IUCr's goals is to extend the powerful technique of crystallography throughout the world. The IUCr journals are meant to be international archival journals for publication of the most accurate structure determinations. It is also essential to the mission of the IUCr to make the data available to the world community as economically and effectively as possible. The fact that the PDB is available to everyone in the world at no cost makes it an ideal resource for the IUCr community.

Q: You have traveled extensively during your tenure as IUCr President. Where have you seen the most potential for growth in new depositions to the PDB?

A: The future growth of the field of crystallography with respect to new structure determinations of all types of matter and the next generation of crystallographers is in the emerging nations of Latin America, the Asia-Pacific region, and Africa. Most emerging countries have indigenous plant, insect, and bacterial species with unique properties. Studies of the structures in these materials will provide new leads for rational drug design and disease control. We can already see the impact of the growth of a crystallographic infrastructure in these countries in the increased numbers of structure reports in *Acta Crystallographica Section E* that are coming from China, Malaysia and Turkey. Over 50% of the papers submitted to *Acta Crystallographica Section E* in 2004 came from China. It is only a matter of time before we see a similar explosion in macromolecular structure reports from these countries in *Acta Crystallographica Section F*.



Neurotrophins (August 2005)

PDB ID: 1bet

McDonald, N.Q., Lapatto, R., Murray-Rust, J., Gunning, J. Wlodawer, A., Blundell, T.L., New protein fold revealed by a 2.3 Å resolution crystal structure of nerve growth factor. Nature **v354** *pp.411-414, 1991*

RCSB PDB Partners

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

RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY
Department of Chemistry and Chemical Biology
610 Taylor Road • Piscataway, NJ 08854-8087

SAN DIEGO SUPERCOMPUTER CENTER, UCSD
9500 Gilman Drive • La Jolla, CA 92093-0537

The RCSB PDB is a member of the **Worldwide PDB (www.wwpdb.org)**

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**

STATEMENT OF SUPPORT: *The RCSB PDB is supported by funds from the National Science Foundation, the National Institute of General Medical Sciences, the Office of Science, 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 Leadership Team

The overall operation of the PDB is managed by the RCSB PDB Leadership Team.

DR. HELEN M. BERMAN, Director
Rutgers, The State University of New Jersey
berman@rcsb.rutgers.edu

DR. PHILIP E. BOURNE, Co-Director
San Diego Supercomputer Center
University of California, San Diego
bourne@sdsc.edu

ALLISON CLARKE, Operations Coordinator
Rutgers, The State University of New Jersey
aclarke@rcsb.rutgers.edu

JUDITH L. FLIPPEN-ANDERSON, Outreach Coordinator
Rutgers, The State University of New Jersey
flippen@rcsb.rutgers.edu

DR. JOHN WESTBROOK, Co-Director
Rutgers, The State University of New Jersey
jwest@rcsb.rutgers.edu

A list of current RCSB PDB Team Members is available at www.rcsb.org/pdb/rcsb-group.html

Job openings available at the RCSB PDB are listed at www.rcsb.org/pdb/jobs.html

RCSB PROTEIN DATA BANK
www.pdb.org
9650 Rockville Pike
Bethesda, MD 20814
Return Service Requested