Fall 2012 • Number 55

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

RCSB PDB PROTEIN DATA BANK NEWSLETTER

www.rcsb.org • info@rcsb.org

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

Contents

SNAPSHOT: OCTOBER 1, 2012

84846 released atomic coordinate entries

ENTRIES BY MOLECULE TYPE

78559	proteins, peptides, and viruses
3833	protein/nucleic acid complexes
2430	nucleic acids
24	other

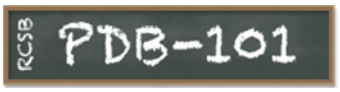
ENTRIES BY EXPERIMENTAL TECHNIQUE

74567	X-ray
9605	NMR
458	electron microscopy
51	hybrid
165	other

RELATED EXPERIMENTAL DATA FILES

63989	structure factors
6911	NMR restraints
672	NMR chemical shifts

Head Back to School with PDB-101



Click on the blackboard logo to visit PDB-101.

The PDB-101 website packages together resources that promote exploration in the world of proteins and nucleic acids for teachers, students, and the general public. Major elements include:

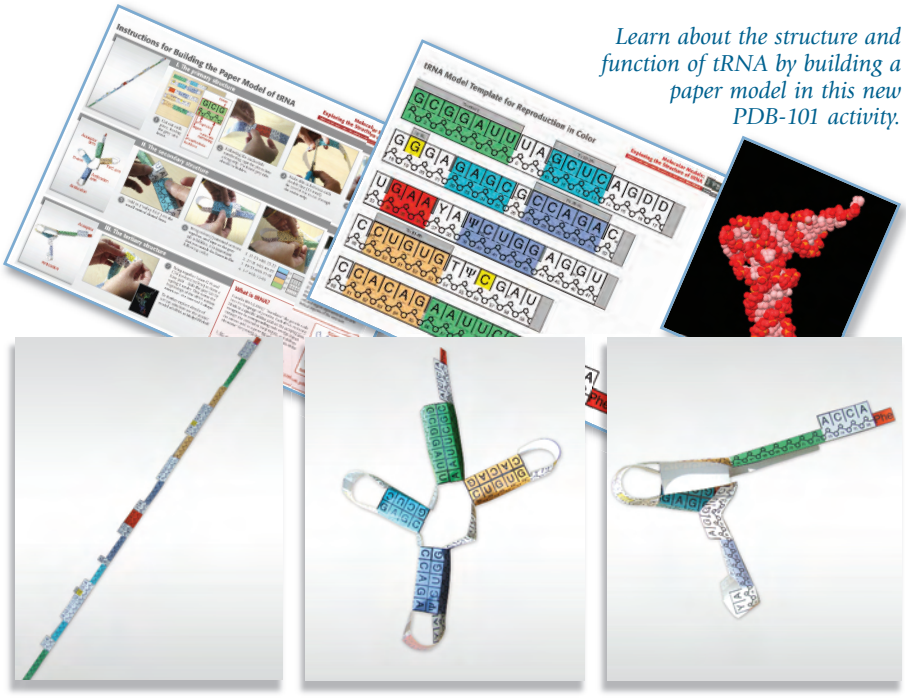
Author Profiles are a unique historical and educational tool that display a timeline of the structures associated with a particular researcher. Example profiles and searches for authors and structural genomics centers are available.

Educational Resources and materials, including posters, animations, and classroom lessons. Recently, a tRNA-building activity was added to the list of available activities.

Molecule of the Month articles describe the structure and function of a molecule, offer interactive views and discussion topics, and link to specialized pages to help explore specific example structures. Each feature can be accessed from a pull-down menu or from the archive.

Structural View of Biology. Built around the *Molecule of the Month* series, this feature promotes a top-down exploration of the PDB. Readers can travel through high-level functional categories (such as Protein Synthesis and Health and Disease) and descriptive subcategories (like Replication or Immune System) to access relevant articles that describe molecules in simple terms and provide specific examples.

Understanding PDB Data is a reference to help explore and interpret individual PDB entries. Broad topics include how to understand PDB data, how to visualize structures, how to read coordinate files, and potential challenges to exploring the archive.



Learn about the structure and function of tRNA by building a paper model in this new PDB-101 activity.

Data Deposition and Annotation

Prepare Data for Deposition with pdb_extract and SF-Tool

Image created at www.worlle.net

Tools are available to help prepare data files for deposition with ADIT. Using these resources can help minimize errors, validate data, and save time during the deposition process.

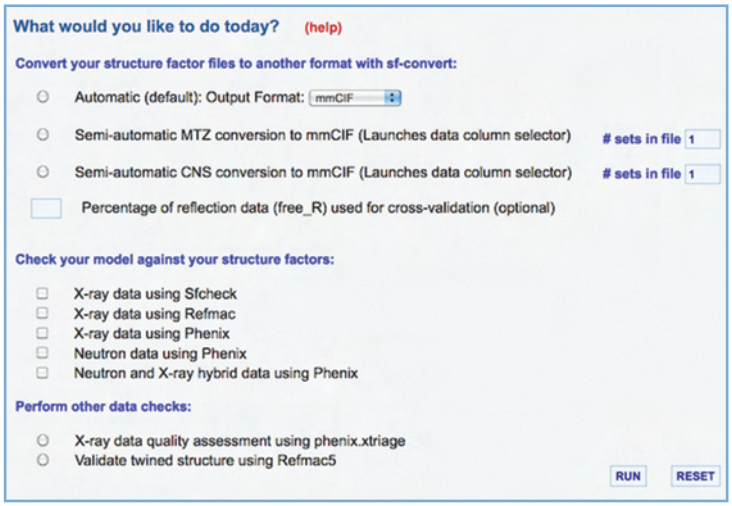
pdb_extract extracts key details from the output files produced by many X-ray crystallographic and NMR applications. The program merges these data into mmCIF data files that can be used with ADIT for validation and deposition.

pdb_extract features include support for:

- Data from hybrid method experiments
- NCS and TLS ranges in BUSTER and REFMAC
- Improved mtz-to-mmCIF conversion
- Quality assessment of X-ray data

Depositors can upload files into the pdb_extract webserver or download the latest workstation version at pdb-extract.rcsb.org.

SF-Tool is a streamlined, web-based tool for validating X-ray, neutron, and hybrid experimental data.



SF-Tool

Visit sf-tool.rcsb.org to:

- Validate model coordinates against structure factor data
- Easily convert structure factor files between different formats (mmCIF, MTZ, CNS/CNX, XPLOR, SHELX, TNT, HKL2000, SCALEPACK, D*Trek, SAINT, and more)
- Check for and validate twinned or detwinned data

The latest release of SF-Tool includes support for neutron and hybrid experiments; incorporates checks from REFMAC,[1] PHENIX,[2] and SFCHECK;[3] and converts multiple data sets into a single mmCIF file.

Deposition Statistics

From July 1-September 30, 2012, 2583 experimentally-determined structures were deposited to the PDB archive, and then processed and annotated by wwPDB teams.

Of the structures deposited in 2012 so far, 81% were deposited with a release status of hold until publication; 15% were released as soon as annotation of the entry was complete; and 4% were held until a particular date. 93% of these entries were determined by X-ray crystallographic methods; 6% were determined by NMR methods.

During the same time period, 2167 structures were released and made publicly available in the PDB.

wwPDB News

Special Symposium on PDB: Basis for Life Science and Drug Development

A special PDB symposium will be held on Saturday, October 13, 2012 at Hearton Hall in Osaka, Japan. The meeting is free and open to the public.

Presentations will include:

wwPDB and its Impacts on Science and Society, Haruki Nakamura, Osaka University

Impact of the Protein Data Bank on Drug Discovery, Stephen Kevin Burley, University of California at San Diego

Molecular Nanomachines in Living Organisms-Exquisite Structural Design Far Beyond State-of-the-Art Nanotechnology, Keiichi Namba, Osaka University

The symposium is organized by the wwPDB Foundation, and is sponsored by the National Bioscience Database Center-Japan Science and Technology Agency; Institute for Protein Research, Osaka University; Graduate School of Frontier Biosciences, Osaka University; The Biophysical Society of Japan; Protein Science Society of Japan; and the Osaka Pharmaceutical Manufacturers Association. For more information, see wwpdb.org.

Science as an open enterprise

The wwPDB's management of the PDB archive was highlighted as a major international data initiative with well-defined protocols for the selection and incorporation of new data and access to them in a report from The Royal Society on openness in scientific data.

The full report, entitled *Science as an open enterprise*, is online at royalsociety.org.

1. G. N. Murshudov, A. A. Vagin, E. J. Dodson. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr.* **D53**: 240-255.

2. P. D. Adams, R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. j. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica* **D58**: 1948-1954.

3. A. A. Vaguine, J. Richelle, S. J. Wodak. (1999) SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **55**: 191-205.

Data Query, Reporting, and Access

Smart Searching with the Top Bar Suggestion Box

Enter text in the top search bar to quickly find structures based on author, macromolecule name, sequence, ligand name or ID, and more.

The top search bar helps users easily and intuitively create simple text searches.

Typing text in the top search bar launches an interactive pop-up box that suggests possible matches that are organized by categories ranging from author name to ontology terms. The top search can be also be limited to quick searches on Author, Macromolecule name, Sequence, or Ligand by selecting the related icon. Using these categories helps to quickly differentiate between possible search results.

For example, autosuggestions for the input bird include authors whose names contain "bird" and structures from the organism bird.

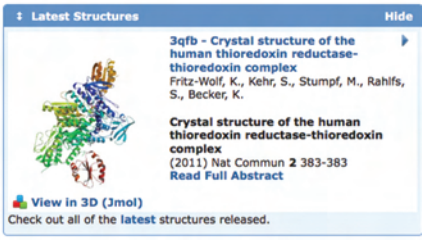
Users who want to perform a simple, non-categorized text search can click the magnifying glass icon or press return.

The top menu bar also recognizes particular types of syntax. Entering SMILES strings will suggest options to perform substructure, exact structure, or similar chemical structure searches, while typing in a sequence will offer different BLAST search options.

Simple text searches complement other RCSB PDB tools: the Explore Archive widget provides browsable data distribution summaries, Browse Database explores the PDB archive using different hierarchical trees, and Advanced Search combines multiple searches of specific types of data.

Different Ways To Explore New Entries

On average, the wwPDB releases 170 structures into the PDB archive each week. The RCSB PDB offers different ways of exploring these new entries:



The Latest Structures widget, located on the home page, cycles through newly released entries.

The **Latest Structures** widget on the home page provides a slideshow of individual entries. It displays the entry title, image, citation, and a link to the PubMed abstract, if available. Users can pause the show at any point to read the entire abstract, click on the entry title to view the entry's Structure Summary page, or go straight to the Jmol view of the entry.

The set of recently released structures can be launched in the Query Results Browser by **clicking on the date listed at the top of the page** and the link provided in the New Structures widget on the home page. From the Query Result Browser, users can drill through by category (organism, taxonomy, experimental method, and more), generate reports, and download structure and sequences files.

The **MyPDB** service can be set to run saved searches with each update. Email alerts (weekly or monthly) will be sent when new entries matching the search are released in the PDB archive.

Download RCSB PDB *Mobile*

The new app provides fast, on-the-go access. Search the entire PDB, view the latest weekly release of structures, access your MyPDB account, view the entire catalog of *Molecule of the Month* articles, and more using either a WiFi or cellular data connection and an iPad/iPhone device.

The latest release improves upon the 3D display of macromolecular biomolecules.

A version of the app for the Android platform is in development.

Search the PDB, access MyPDB, view molecules in 3D, and more with RCSB PDB Mobile.

Build Complex Queries with Advanced Search

Advanced Search provides the capability of combining multiple searches of specific types of data in a logical AND or OR. The result is a list of structures that comply with ALL or ANY of the search criteria, respectively.

Individual data items are organized by category; contextual help and examples are available by selecting the **question mark icon**. New options include quick searches by experimental and/or molecule type, searches based on structure determination method, and the ability to find structures containing interresidue connectivity (LINK records) that cannot be inferred from the primary structure.

Currently, users can build searches based on:

- Quick Search: retrieve all PDB entries or a subset based on experimental method or molecule type
- ID(s) and Keywords: PDB, PubMed, UniProtKB, Pfam IDs; text and keyword searching
- Structure Annotation: structure title, description; and macromolecule name
- Deposition: author name; deposit, release, and revision date; latest released and modified structures; Structural Genomics Project
- Structure Features: macromolecule type; number of chains (asymmetric unit or biological assembly), entities, models, and disulfide bonds; interresidue connectivity (LINK records); molecular weight; secondary structure content; secondary structure length; SCOP, CATH, taxonomy
- Sequence Features: sequence; translated nucleotide sequence; sequence motif; chain length; genome location
- Chemical Components: name; ID; InChi descriptor; SMILES/SMARTS; molecular weight; chemical formula; chemical component type; binding affinity; has ligands; has modified residues; sub-components
- Biology: Source; expression organism; Enzyme Classification; biological process; cell component; molecular function; Transporter Classification
- Methods: experimental method; X-ray resolution, R factor, diffraction source, structure determination method, reflections, cell dimensions, software, space group, crystal properties, detector; EM assembly
- Publication: citation; MeSH terms; PubMed abstract
- Misc: Has external links

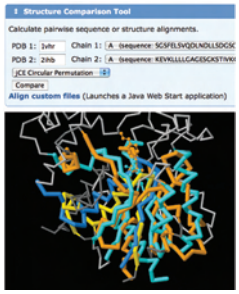
The number of entries matching each individual query can be shown before running the full Advanced Search. Searches can also be filtered by removing sequence similarity.

Advanced Searches can be stored in MyPDB to be run or modified at any time.

Customized Home Page: Structure Comparison Tool

The RCSB PDB homepage is comprised of web widgets that can be moved around, minimized, or hidden so users can create a website that reflects their interests. Frequently used features can be moved to the top, while less popular items can be hidden or collapsed.

The Structure Comparison Tool widget can be added so it appears on every visit to the RCSB PDB home page. The tool calculates pairwise sequence (blast2seq, Needleman-Wunsch, and Smith-Waterman) and structure alignments (FATCAT, CE, Mammoth, TM-Align, TopMatch).



Add the Structure Comparison Tool widget to your home page to quickly calculate pair-wise sequence and structure alignments.

Comparisons can be made for any protein in the PDB archive and for customized or local files not in the PDB. Special features include support for both rigid-body and flexible alignments (via jFATCAT) and detection of circular permutations (via jCE).

To add this widget to the RCSB PDB home page, select the Customize This Page button from the left menu. Download Files, Sequence Search, and ADIT Deposition widgets can also be added or removed.

The Comparison Tool is also available from Structure Summary pages and as a stand-alone Java Web Start application, with additional information available online.

Website Statistics

Access statistics for the third quarter of 2012 are shown.

Month	Unique Visitors	Number of Visits	Bandwidth
JULY	184117	478089	697.83 GB
AUGUST	175961	439848	484.91 GB
SEPTEMBER	241088	572574	755.14 GB

Explore Different PDB Data Overviews

The RCSB PDB maintains several lists that display PDB data in interesting and useful ways. Highlights include:

Proteins Solved by Multiple Experimental Methods: Lists clusters of proteins (with greater than 95% sequence similarity) containing at least one structure solved by one method (e.g. X-ray) and one by a different method (e.g. NMR).

PDB Statistics: Content Distribution & Growth: View data distributions of the PDB archive based on characteristics like space group, journal, EC, and more. The growth of the number of PDB entries released per year is organized around experimental method, molecule type, and unique protein classifications. Data can be downloaded as Excel documents.

RSS Feeds: Subscribe to RSS feeds to access the latest structures (from the RCSB PDB) and *Molecule of the Month* articles (from PDB-101) as soon as they become available.

Summaries of PDB Data: Links to summary files available from the PDB FTP, including a file with all PDB sequences in FASTA format; one with all PDB IDs, molecule type, and experimental method; and one with all PDB IDs and authors.

Secondary Structure Files: A FASTA formatted file generated using DSSP displays sequences and secondary structure for all entries. A separate file includes notation of regions which have not been experimentally observed in addition to the secondary structure.

Latest weekly release and all released entries: The top right of every RCSB PDB page header links to the full archive (select the number of total structures) and to the entries in the most recent update (select the date listed).

PDB-101 Author Profiles: Author Profiles vertically display structures associated with a particular researcher. Profiles are also available for Structural Genomics centers.

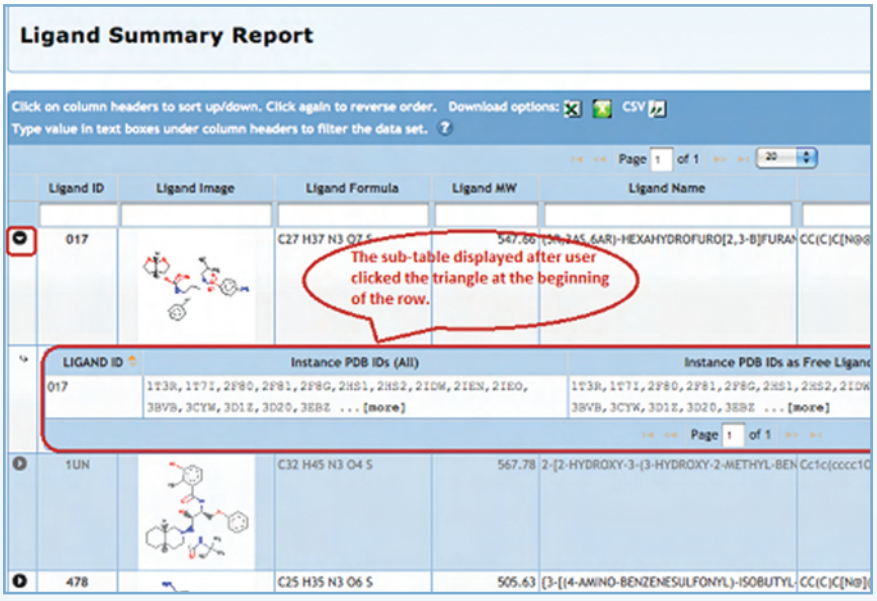
PDB-101 Molecule of the Month: Articles are indexed by title, date, and category, and accessible from an alphabetical pull-down menu.

Improved Ligand Reports

Ligand Summary Reports include information about the selected chemical components such as formula, molecular weight, name, SMILES string, which PDB entries are related to the ligand, and how they are related.

All tabular report features are also available, including sorting, filtering, export to other report formats, and column customization.

Ligands can also be displayed as an image collage.



For each ligand included a Ligand Summary report, a sub-table can be revealed to show lists of all related PDB entries that contain the ligand, the entries that contain the ligand as a free ligand, and entries that contain the ligand as part of a larger, polymeric ligand.

Outreach and Education

High Resolution Images at the RCSB PDB

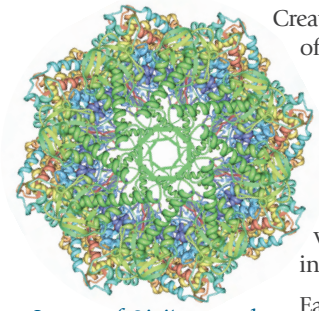


Image of 3izi[4] created using Protein Workshop

Create or download publication-quality pictures of biomacromolecules with the RCSB PDB.

Several interactive, Java-based tools[5] can be used to visualize PDB data and create pictures. **Protein Workshop** offers easily customized views; **Simple Viewer** utilizes a quick ribbon display; and **Ligand Explorer** visualizes the interactions of bound ligands in protein and nucleic acids structures.

Each program can be used to create and save high-resolution images in JPEG, PNG, and TIFF formats. Using the Save Image dialog box from the File menu, users can specify the width and height of an image in pixels, inches, or millimeters.

From the educational PDB-101 site, high resolution TIFFs used to illustrate *Molecule of the Month* articles can be downloaded and used in presentations and publications (see citation and usage information).

PDB-101 posters, including *The Structural Biology of HIV* and *Molecular Machinery: A Tour of the Protein Data Bank*, can be saved as high or low resolution PDFs.

Author Profiles: Timeline Display of Structures

Author Profiles vertically display structures associated with a particular researcher. Profiles are now available for Structural Genomics centers. The timeline can be sorted by deposition date, and the first instance of a protein or complex in the author's profile is highlighted. Images are linked to the Structure Summary page for the entry.

To access this feature, select the link from the left hand RCSB PDB menu, or from the PDB-101 Features pull-down menu. A search box is used to search by author or Structural Genomics center.

Image from the Author Profile for the Center for Structural Genomics of Infectious Diseases

Meeting News

ISMB and 3DSIG: The International Society for Computational Biology recently celebrated the 20th year of the ISMB–Intelligent Systems for Molecular Biology–conference (July 15-17, Long Beach, CA).

RCSB PDB presentations included *Internal pseudo-symmetry in proteins* (Andreas Prlić) and *Efficient searching and mining of the RCSB Protein Data Bank* (Peter Rose).

At the ISMB Special Interest Group meeting Bioinformatics Open Source Conference (BOSC), Andreas Prlić also presented *How to use BioJava to calculate one billion protein structure alignments at the RCSB PDB website*.

At the 3Dsig: Structural Bioinformatics & Computational Biophysics satellite meeting, John Westbrook discussed *Format déjà vu: PDBx/mmCIF, the new data format for the wwPDB*. *Molecule of the Month* author David Goodsell described *Communicating and Interacting with the Molecular Cell* (along with Arthur Olson).

Associate Director Phil Bourne was one of the 3Dsig Program Chairs.

ACA: From July 28-August 1, the RCSB PDB was at the 2012 Meeting of the American Crystallographic Association in Boston, MA. Attendees visited the exhibition booth to meet with members of the RCSB PDB and the PSI Structural Biology Knowledgebase and learned about new features such as RCSB PDB *Mobile*, Author Profiles, the wwPDB Common Tool for Deposition and Annotation, and much more. Lead annotator Jasmine Young presented an update on *The Worldwide Protein Data Bank: Current Projects during the poster session*. During the Structure-Guided Drug Discovery session, *Data Management of Small Molecule Ligands, Antibiotics, and Peptide Inhibitors in the Protein Data Bank* was described by RCSB PDB Director Helen Berman.

Protein Society: At the 26th Annual Symposium of The Protein Society (August 5-8, San Diego), Helen Berman was the 2012 Awardee of the Carl Brändén Award. The award, sponsored by the Rigaku Corporation, is given to an outstanding protein scientist who has also made exceptional contributions in the areas of education and/or service to the science. This award recognizes Dr. Berman's accomplishments toward enabling a freely available and uniform worldwide archive of 3D structural information for biomedical research and education. She presented *Trendspotting from the Protein Data Bank* as part of the Plenary Awards Session.

Poster Prizes Awarded

The RCSB PDB Poster Prize is awarded for the best student poster presentations at selected meetings. Recipients receive an educational book and a subscription to *Science*.

Sergei Kalynych

At ACA, the award went to Sergei Kalynych for *Crystallographic studies of closely related lipopolysaccharide O-antigen chain length regulators* (Sergei Kalynych,[1] Deqiang Yao,[2] James Magee,[1] Mirek Cygler;[2] [1]McGill University, [2]University of Saskatchewan)

An honorable mention was awarded to Rebecca Goldstein for *A possible mechanism for the regulation of the PI-PLC from* S. aureus (Rebecca Goldstein, Jiongjia Cheng, Mary Roberts; Boston College).

Many thanks to the judges: Barry Finzel (Chair, University of Minnesota), Patrick Loll (Drexel University), and David Rose (University of Waterloo), and to Marcia Colquhoun and the ACA.

Alan Barber

At ISMB, Alan Barber was recognized for *Evolution of function in the alkaline phosphatase superfamily* (Alan Barber,[1] Jonathan Lassila,[2] Helen Wiersma-Koch,[2] Michael Hicks,[1] Daniel Herschlag,[2] Patricia Babbitt;[1] [1]University of California, San Francisco, [2]Stanford University).

Many thanks to Steven Leard, the International Society for Computational Biology, and the judges: Yana Bromberg (Chair, Rutgers), Hannah Carter (Johns

4. N.R. Douglas, S. Reissmann, J. Zhang, B. Chen, J. Jakana, R. Kumar, W. Chiu, J. Frydman (2011) Dual action of ATP hydrolysis couples lid closure to substrate release into the group II chaperonin chamber *Cell* **144**: 240-252.

5. J.L. Moreland, A.Gramada, O.V. Buzko, Q. Zhang and P.E. Bourne (2005) The Molecular Biology Toolkit (MBT): A modular platform for developing molecular visualization applications BMC *Bioinformatics* **6**:21.

Hopkins), Jeroen deRidder (Delft University of Technology), Iddo Friedberg (Miami University, Oxford, Ohio), Tatyana Goldberg (Technical University Munich, Germany), Ben Jelen (Rutgers), Hande Kucuk (University of Miami), Biao Li (Buck Institute, Novato), Magali Michaut (Netherlands Cancer Institute), Yanay Ofran (Bar Ilan University), Susanna Repo (EMBL-EBI), Venkata Satagopam (EMBL-Heidelberg), Andrea Schafferhans (Technical University Munich), Avner Schlessinger (UCSF), Stefan Senn (Rutgers), Janita Thusberg (Buck Institute, Novato), Tobias Wittkop (Buck Institute, Novato), Chengsheng Zhu (Rutgers).

Jonas Lindholt

At the European Crystallographic Meeting (August 7-11, Bergen, Norway), Jonas Lindholt received the prize for *Cardiotonic Steroids and the Na/K-ATPase* (Jonas Lindholt, Linda Reinhard, Poul Nissen; Centre for Membrane Pumps in Cells and Disease (PUMPkin) and Department of Molecular Biology and Genetics, Aarhus University, Denmark).

Many thanks to Andreas Roodt (ECA Prize Committee and University of the Free State, South Africa), the European Crystallographic Association, and the judges: Udo Heineman (Max-Delbrück Center for Molecular Medicine, Berlin), Linda Shimon (Weizmann Institute of Science), and Ute Krengel (University of Oslo).

All poster prize awardees are listed on the RCSB PDB website.

wwPDB News

Future Funding of the BioMagResBank: White Paper and Community Discussions

A white paper on the future funding of the BMRB and related articles have been published in the September issue of *Nature Structural & Molecular Biology*:

Editorial: The BMRB matters (2012) *NSMB* **19**: 853 **doi:10.1038/nsmb.2387**

Community Comments: In support of the BMRB

From the BMRB: *White Paper on Future Funding* (2012) *NSMB* **19**: 854–860 **doi:10.1038/nsmb.2371**

This issue has also been discussed in the *Nature, The Scientist, and Corante*

The BMRB, housed at the University of Wisconsin–Madison, collects, annotates, archives, and disseminates the important spectral and quantitative data derived from NMR spectroscopic investigations of biological macromolecules and metabolites.

Current funding from the National Library of Medicine is set to expire in 2014. It appears critical that a funding plan be developed within the coming year.

Education Corner by John Rose, Wendy Dustman and Julie Kittleson

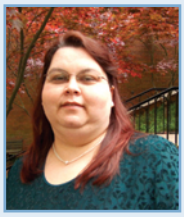
The UGA Biotech Boot Camp: Introducing Georgia High School Teachers to Biotechnology and the PDB

The UGA Biotech Boot Camp is a one-week (M-F) resident workshop in biotechnology for Georgia high school science teachers aimed at increasing the teacher's content knowledge and skill proficiency in the rapidly growing field of biotechnology. Each day of the Boot Camp consists of morning sessions addressing content knowledge, afternoon sessions focused on laboratory skills, along with lunch and evening activities aimed at improving teaching effectiveness (for more information, see **biotechbootcamp.uga.edu**).

Afternoon sessions allow teachers to gain hands-on laboratory experience and learn new skills. Teachers are exposed to DNA analysis (including agarose gel electrophoresis and DNA finger printing), DNA transformation (producing green fluorescent protein (GFP)), protein purification (using ion exchange chromatography and polyacrylamide gel electrophoresis to purify and characterize GFP), protein crystallization (hen egg white lysozyme), X-ray diffraction (via remote access to SER-CAT[6]), bioinformatics, and protein structure-function.

These afternoon sessions include participation in several hands-on, low or no-cost classroom activities and simulations associated with the above experiments. Computer-based activities, such as the bioinformatics and protein structure-function exercises, are very appealing to the teachers, since schools usually have a computer lab and Internet access. Teachers are shown how to access and use: (1) *UniProt* Knowledgebase for protein sequences,[7] (2) *BLAST* for sequence comparison,[8] (3) the *RCSB PDB* for 3D structures of proteins and nucleic acids, and (4) *PubMed*[9] and *Google Scholar*[10] for literature searches. Exercises and ideas about how to incorporate these resources into their classroom instruction are provided. For example, in the

DR. JOHN ROSE *is an Associate Professor in the Department of Biochemistry & Molecular Biology at the University of Georgia, where he teaches* Introduction to Biochemistry *to undergraduates and several graduate level courses in biochemistry. He serves as Assistant Director of SER-CAT at the Advanced Photon Source, Argonne National Laboratory, one of the world's most productive facilities for X-ray structural biology.*



DR. WENDY DUSTMAN *is an Assistant Professor in the Department of Microbiology at the University of Georgia, where she teaches* Introduction to Microbiology *and other undergraduate courses in microbiology. She was a 2009 Lilly Teaching Fellows finalist, and received the 2011 UGA Sandy Beaver Excellence in Teaching Award for the Division of Biological Sciences.*

DR. JULIE KITTLESON *is an Assistant Professor in the Department of Mathematics and Science Education at the University of Georgia. She teaches several undergraduate and graduate level science education classes, including* Philosophy and Leadership in Science Classroom Practice. *She currently serves on the Board of Reviewers for* Science Education.

Classroom Activity *Connecting DNA to Disease Using BLAST*, a DNA sequence belonging to a disease related protein is given to the student. The student then transcribes and translates the DNA sequence to produce the corresponding amino acid sequence. Using the amino acid sequence, the student then uses *BLAST* to compare sequences and UniProt to identify and characterize the protein and the disease to which it is associated. Based on this information and using other tools, the student then is asked to describe the protein, its chromosome location and function, the symptoms of the disease, the prognosis and possible treatments.

To help illustrate structure-function in the classroom, teachers are shown how to download and install the free Chimera molecular modeling tool from UCSF[11] and are provided a library of amino acid structures and other common molecules found in biology such as aspirin, ATP/ADP, glucose, sucrose, Tylenol® and vitamin B12 for display in Chimera. Teachers are instructed on using basic Chimera commands to display and manipulate (i.e., rotate, translate, color and label atoms) in wireframe, ball and stick, and spacefilling representations of these molecules. The RCSB PDB is then introduced, and the teachers learn how to navigate the website, search the archive and download protein coordinates for display in Chimera. Education resources such as PDB-101 are highlighted, followed by a discussion as to how these resources could be used in the classroom. The Classroom Activity *Using Chimera in the Classroom* illustrates how the PDB structures and Chimera can be adapted to the classroom (computer lab) setting. Teachers use a simple leucine zipper structure (PDB ID 1ZIK) and Chimera running on their own laptop to probe the 1ZIK structure to understand protein structure-function, including how hydrogen bonding stabilizes the α helix and the role of electrostatics. In this exercise, teachers increase their skills in using Chimera features, and learn how to produce simple movies and publication-quality figures. Teachers are also introduced to Proteopedia[12] and its excellent articles on hemoglobin, enzymes, ion channels, and other proteins. The hemoglobin article provides an excellent teaching tool for protein structure-function relationships, since most students know that it is the red protein found in blood but unaware as to how it binds and releases oxygen. The Proteopedia article describes the $(\alpha\beta)_2$ hemoglobin tetramer, how oxygen binds to the heme group and the allosteric nature of the R to T transition that accompanies oxygen binding and release.

The Boot Camp is generally held in early June, when teachers are out of class and beginning to develop lesson plans for the next school year. The workshop hosts between 12-15 high school science teachers

The 2012 UGA Biotech Boot Camp participants.

Expression and purification of GFP. (A) Induced cells glowing green under UV light, (B) GFP colonies illuminated by UV, (C) GFP eluting from the ion exchange column illuminated by UV and (D) Purified GFP glowing green under UV light. Teachers shown (from left to right): Cindy Kay Plumly, Debbie Thomlinson, Laurie Diffie, Denita Alonso.

and is held at the NESPAL[13] facility located on the University of Georgia Tifton campus. The program provides teachers with meals and lodging through the nearby Abraham Baldwin Agricultural College. The workshop provides teachers with (1) a DNA model kit for use in their classroom, (2) a 300 page bound set of workshop (lecture/lab/activity) notes, (3) a USB memory stick containing over 900 Mbytes of workshop related materials (notes, reference papers, experiments and a variety of other classroom activities), (4) Ellyn Daugherty's text book *Biotechnology Science for the New Millennium*, and (5) a wonderful poster provided by the RCSB PDB.

The workshop has been designed to foster sustained contact between the teachers and workshop personnel via a Science Facilitator who travels to the teacher's classroom and helps the teacher conduct one of the Boot Camp lab activities. Since each teacher attending the workshop represents between 80-130 students each semester, we estimate that over 2500 students per year benefit from skills and knowledge his/her teacher brings back from the workshop. The Boot Camp is funded by an *Improving Teacher Quality* grant to the University of Georgia School of Education from the U.S. Department of Education. For more information please see **www.biotechbootcamp.uga.edu**.

6. Southeast Regional Collaborative Access Team (SER-CAT) Sector 22 at the Advanced Photon Source, Argonne National Laboratory, **www.ser-cat.org**.
7. UniProt Consortium (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* **40**: D71-75.
8. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410 blast.ncbi.nlm.nih.gov.
9. E. W. Sayers, T. Barrett, D. A. Benson, E. Bolton, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. Dicuccio, S. Federhen, M. Feolo, I. M. Fingerman, L. Y. Geer, W. Helmberg, Y. Kapustin, S. Krasnov, D. Landsman, D. J. Lipman, Z. Lu, T. L. Madden, T. Madej, D. R. Maglott, A. Marchler-Bauer, V. Miller, I. Karsch-Mizrachi, J. Ostell, A. Panchenko, L. Phan, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, M. Shumway, K. Sirotkin, D. Slotta, A. Souvorov, G. Starchenko, T. A. Tatusova, L. Wagner, Y. Wang, W. J. Wilbur, E. Yaschenko, J. Ye. (2012) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **40**: D13-25. **www.ncbi.nlm.nih.gov/pubmed/**.
10. Butler, D. (2004). Science searches shift up a gear as Google starts Scholar engine. *Nature* **432**(7016): 423. **scholar.google.com**.
11. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin. (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612. **www.cgl.ucsf.edu/chimera**
12. Hodis, E., J. Prilusky, et al. (2008). Proteopedia - a scientific 'wiki' bridging the rift between three-dimensional structure and function of biomacromolecules. *Genome Biol* **9(8)**: R121. Web site: **www.proteopedia.org**.
13. NESPAL: the National Environmentally Sound Production Agriculture Laboratory, a unit University of Georgia's College of Agricultural and Environmental Sciences. Web site: **nespal.cpes.peachnet.edu**.

RCSB PDB PROTEIN DATA BANK NEWSLETTER

Web: www.rcsb.org
Email: info@rcsb.org
FTP: ftp.wwpdb.org

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
Center for Integrative Proteomics Research
174 Frelinghuysen Road
Piscataway, NJ 08854-8087

San Diego Supercomputer Center and the Skaggs School of Pharmacy and Pharmaceutical Sciences
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0537

The RCSB PDB is a member of the Worldwide Protein Data Bank (www.wwpdb.org)

RCSB PDB Management

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

DR. MARTHA QUESADA, Deputy Director
Rutgers, The State University of New Jersey
mquesada@rcsb.rutgers.edu

DR. PHILIP E. BOURNE, Associate Director
San Diego Supercomputer Center and the Skaggs School of Pharmacy and Pharmaceutical Sciences,
University of California, San Diego
bourne@sdsc.edu

A list of current RCSB PDB Team Members is available from www.rcsb.org.

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 Institute of Neurological Disorders and Stroke, and the National Institute of Diabetes & Digestive & Kidney Diseases.*

Return Service Requested

RCSB PROTEIN DATA BANK
www.rcsb.org • info@rcsb.org
Rutgers, The State University of New Jersey
Center for Integrative Proteomics Research
174 Frelinghuysen Road
Piscataway, NJ 08854-8087
USA