Winter 2012 • Number 52

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

RCSB PDB PROTEIN DATA BANK NEWSLETTER

www.pdb.org • info@rcsb.org

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

Contents

SNAPSHOT: JANUARY 1, 2012
78237 released atomic coordinate entries

ENTRIES BY MOLECULE TYPE

72468	proteins, peptides, and viruses
3416	protein/nucleic acid complexes
2330	nucleic acids
23	other

ENTRIES BY EXPERIMENTAL TECHNIQUE

68424	X-ray
9210	NMR
396	electron microscopy
48	hybrid
159	other

RELATED EXPERIMENTAL DATA FILES

57821	structure factors
6515	NMR restraints
284	NMR chemical shifts

This newsletter is printed on recycled paper

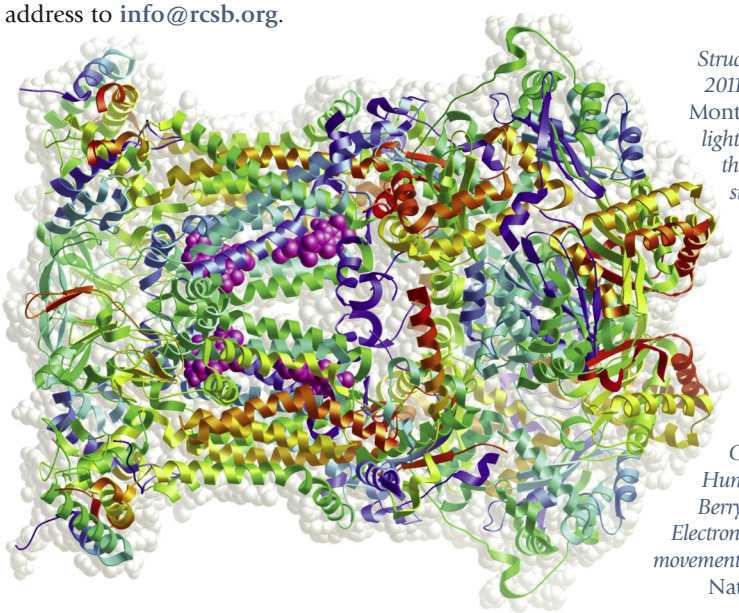
Message from the RCSB PDB

The recently published Annual Report highlights the RCSB PDB's accomplishments for 2011. Milestones covered include the ways in which searching has been simplified. Complex queries can be made by simply clicking through the distribution charts provided for major data categories. The top bar of every web page also offers easy and intuitive searching. After typing a few characters, an interactive pop-up box appears with suggestions of common PDB search terms, organized in different categories to provide more precise results.

The Annual Report demonstrates the features of the new PDB-101 educational resource. This view of the RCSB PDB website packages together resources of interest to teachers, students, and the general public to promote exploration in the world of proteins and nucleic acids. To promote top-down exploration of the archive, PDB-101 offers an interface that lets readers browse from high-level functional biological categories down to specially selected related molecules.

wwPDB efforts, such as the PDB40 anniversary celebration and the publication of the X-ray Validation Task Force's recommendation report, are also highlighted.

These bulletins provide a yearly snapshot of RCSB PDB activities and the state of the PDB archive. This edition, the RCSB PDB's twelfth, is available as a PDF download from www.rcsb.org/ar11. If you would like a printed copy, please send your postal address to info@rcsb.org.



Structures highlighted in 2011's Molecule of the Month *features are highlighted on the cover and throughout the report, such as this image of Cytochrome bc1.*

PDB ID: 3h1j. Z. Zhang, L.S. Huang, V.M. Shulmeister, Y.I. Chi, K.K. Kim, L.W. Hung, A.R. Crofts, E.A. Berry, S.H. Kim (1998) Electron transfer by domain movement in cytochrome bc1. Nature 392: 677-684.

Data Deposition and Annotation

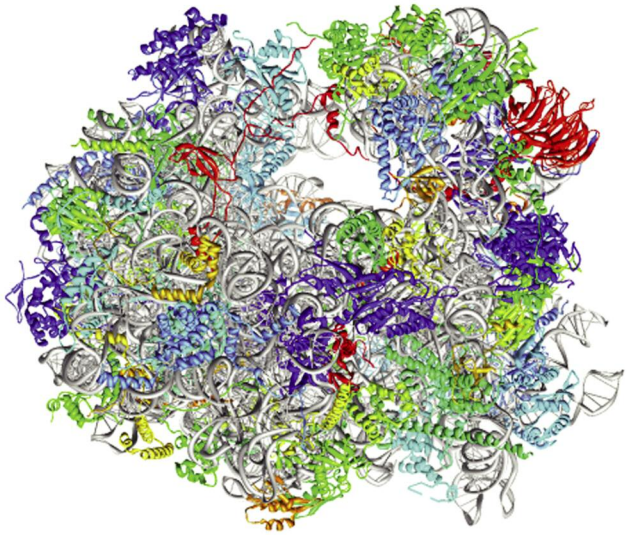
2011 Deposition Statistics

In the fourth quarter of 2011, 2295 experimentally-determined structures were deposited to the PDB archive for a total of 9250 entries deposited in the year. 8865 entries were deposited in 2010.

Of all structures deposited in 2011, 81% were deposited with a release status of hold until publication; 17% were released as soon as annotation of the entry was complete; and 2% were held until a particular date.

93% of these entries were determined by X-ray crystallographic methods; 6% were determined by NMR methods.

8122 structures were released in the PDB archive in 2011. They account for 10% of the current total holdings of 78237.



A record 30 ribosome structures were deposited in the fourth quarter of 2011, for a total of 49 ribosomes submitted in the year. These ribosomes include the crystal structure of the 80S ribosome from the yeast Saccharomyces cerevisiae *shown here from PDB ID 3u5b.*[1]

PSI | nature StructuralBiology Knowledgebase

Structural Genomics News

New Target Registration Database

TargetTrack (sbkb.org/tt), a new experimental data tracking database, offers the latest information on the progress of structural studies on registered protein targets.

TargetTrack provides information about protein complexes, membrane proteins, cryoEM studies, and other advances in structural biology along with the data previously available in TargetDB and PepcDB. Provided by worldwide structure genomics centers, these data can help many biological researchers with their projects' experimental design.

WORLDWIDE PDB PROTEIN DATA BANK

wwPDB News

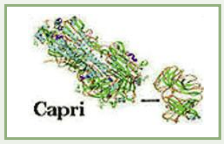
X-ray Validation Task Force Report Published

To improve validation methods in the PDB, the wwPDB has convened method-specific Task Forces to collect recommendations and develop consensus on additional validation that should be performed, and to identify software applications to perform validation tasks.

The first report has been published by the X-ray Validation Task Force. These recommendations will be incorporated into the wwPDB data processing procedures and tools as part of the Common Deposition & Annotation Tool project.

A new generation of crystallographic validation tools for the Protein Data Bank. R.J. Read, P.D. Adams, W.B. Arendall III, A.T. Brunger, P. Emsley, R.P. Joosten, G.J. Kleywegt, E.B. Krissinel, T. Lutteke, Z. Otwinowski, A. Perrakis, J.S. Richardson, W.H. Sheffler, J.L. Smith, I.J. Tickle, G. Vriend, P.H. Zwart (2011) *Structure* **19**: 1395-1412.doi: 10.1016/j.str.2011.08.006

PDB Depositions Can Be Targets for CAPRI



Depositing a protein-protein, protein-DNA, protein-RNA or protein-peptide complex? Consider submitting your structure as a target for CAPRI after depositing your entry to the PDB.

CAPRI (Critical Assessment of PRedicted Interactions) is a community-wide, double-blind experiment aimed at assessing the performance of protein docking algorithms.

Submitting a target to CAPRI will help advance protein methods calculations, may provide new information on the quaternary structure of your complex, and should increase the visibility of your work. Moreover, CAPRI is designed to maintain strict target confidentiality, and imposes little delay on its publication.

More information about target submission is available at the CAPRI site at http://bit.ly/vDLjWq and from the CAPRI management team (joel.janin@u-psud.fr).

Notice for REFMAC Users

Depositors using REFMAC with TLS should make sure that the ATOM records contain full (and not residual) B factors before depositing an entry to the PDB.

When TLS is involved in refinement using a script, the keyword "tlsout addu" should be added to produce a coordinate file containing the full B factor. The latest REFMAC interface (in CCP4i) has a button "Add TLS contribution to XYZOUT" in the folder of TLS Parameters for adding this keyword.

If depositors have finished the final round of refinement without including "tlsout addu" keywords, and do not see ANISOU records in the coordinates, only the residual B factors have been generated in the coordinate file. The residual B factor needs to be updated to the full B factor in order to properly run validation checks.

If needed, the TLSANL page at deposit.rcsb.org/adit/ REFMAC.html provides web form server and scripting information for converting files.

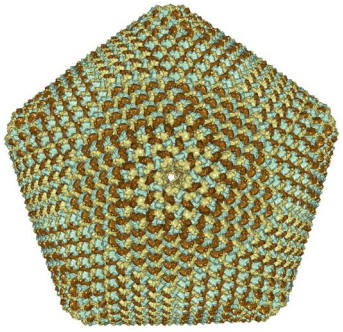
Data Query, Reporting, and Access

2011 RCSB PDB Website Statistics

Month	Unique Visitors	Number of Visits	Bandwidth
JANUARY	210,986	522,651	982.26 GB
FEBRUARY	213,862	526,896	1,114.34 GB
MARCH	239,640	598,545	1,004.99 GB
APRIL	230,825	557,911	763.04 GB
MAY	233,123	561,880	814.08 GB
JUNE	206,456	507,798	620.89 GB
JULY	187,344	470,491	589.52 GB
AUGUST	189,681	467,477	642.00 GB
SEPTEMBER	242,775	567,725	683.56 GB
OCTOBER	274,335	646,682	800.58 GB
NOVEMBER	280,058	660,068	945.80 GB
DECEMBER	233,065	547,841	673.20 GB

Latest Website Release

New and enhanced features have been added to the RCSB PDB website, including:



Surface of PBCV-1 virus capsid with 5040 chains; one of the largest assemblies in the archive (PDB ID 1m4x).[2] Image created using Protein Workshop's new Surfaces option.

• **Visualizing molecular surfaces**. Protein Workshop now supports molecular surfaces to aid in the display of quaternary structure, protein-protein interactions, and binding sites. The program is one of the molecular viewing options available from every Structure Summary page.

• **Exploring ligands.** Improved Jmol visualization, information about subcomponents, and links to DrugBank are some of the features added to Ligand Summary pages. Additionally, binding affinity data from PDBbind (**www.pdbbind-cn.org**) have been added to the External Ligand Annotations widget on Structure Summary pages.

• **Access pre-released sequences via Web Services.** The RCSB PDB supports RESTful Web Services, which can be used to help software developers access data more easily. A new service for accessing pre-released sequences in FASTA format has been added; information about all possible services is available at **http://bit.ly/tiqkzs**.

The What's New page at **http://bit.ly/rJPaA2** has complete descriptions of all the new features.

Looking for Ligands

All residues and small molecule components found in PDB entries, including standard and modified amino acids/nucleotides, small molecule ligands, ions, and solvent molecules, are described in the wwPDB's Chemical Component Dictionary.

The information stored in this dictionary can be easily searched using any of these query options:

• **Simple top bar searching** (with autocomplete suggestions). Click on the Ligand icon (located above the top search box) to limit your search. Enter a ligand name or the 3-character chemical component ID.

• **Chemical Component Search.** This powerful form can be accessed by selecting the top bar ligand icon and then **[additional ligand options]**. Launch the chemical structure editor to draw a structure, or paste in a SMILES or SMARTS string, and then perform a substructure, exact, similarity, or superstructure search. In addition, ligands can be searched by name, identifier, formula, and molecular weight. A tutorial and screencast demonstrating the chemical structure search is available at **http://bit.ly/ruxFLI**.

• **Advanced Search.** Supports the types of searches mentioned above, plus options to query by chemical component type, binding affinity, sub-component, and more. These searches can be combined and even amended with searches for the associated structures.

These searches will return a Ligand Summary page that contains an overview of the chemical component, 2D and 3D images, links to other resources, and links to related ligands and PDB entries.

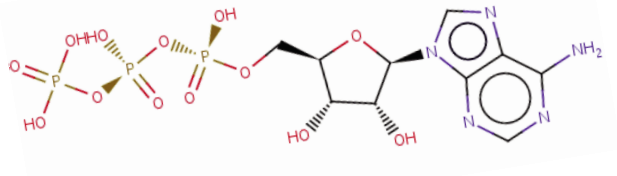
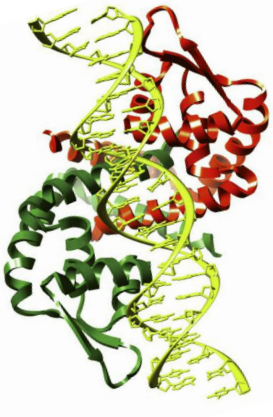


Image of ATP from the Ligand Summary page

Search the PDB for Homo- and Hetero-multimer Structures



Protein homodimer bound to DNA as seen in PDB ID 3Q5F.[3]

The following Advanced Search options (listed under Structure Features) can be combined to find homo- and hetero-multimer structures in the PDB:

• Number of Chains (Biological Assembly): specifies the number of polymeric chains in a multimer

• Number of Entities: limits the number of unique polymeric units (protein, DNA, or RNA) in a multimer

• Structure Features/Macromolecule Type: selects the type(s) of macromolecule to be included in the multimer (protein, DNA, RNA, and/or DNA/RNA hybrid).

Detailed examples of these types of searches can be found in the RCSB PDB news at **http://bit.ly/tdNlan**.

Customize Your RCSB PDB Homepage

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 removed from the page or collapsed by selecting *Hide*.

Widget boxes with an arrow can be dragged up and down on the page. Select the Customize This Page button from the left menu to choose which sections are displayed.

Options in the default view include:

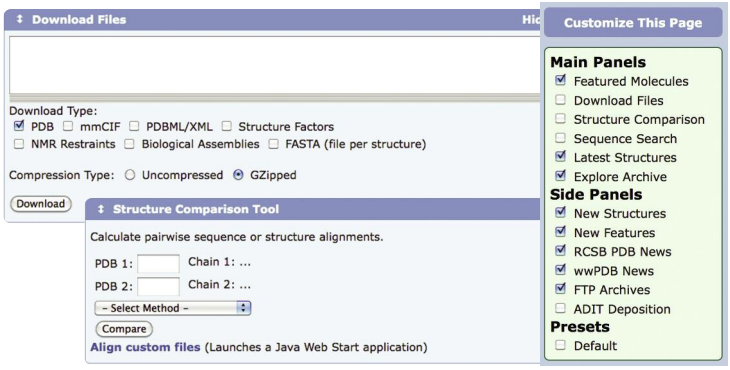
- **Featured Molecules**: Read the latest *Molecule of the Month* and the PSI SBKB's Featured System.
- **Latest Structures**: Scrolls through a slideshow of individual entries, with links to the related abstract, Structure Summary page, and Jmol view.
- **Explore Archive**: Tour the PDB archive by "drilling-down" on significant properties of structures like *Organism and Polymer Type*. This widget also gives a quick statistical overview of the archive.
- **New Structures**: Links to the latest release, structure papers included in a release, and to the Unreleased Entry Search.
- **New Features**: Scrolls through the latest website features and improvements with links to detailed descriptions.
- **RCSB PDB News**: Read the latest updates and highlights, *Newsletter* and more.
- **wwPDB News**: Updates from wwPDB.org are linked in this widget.
- **FTP Archives**: Links to the current FTP archive and to archival snapshots.

Widgets not included in the default view that can be added are:

- **Download Files**: Provides a form on the home page to easily download any number of structures in a variety of formats.
- **Structure Comparison**: Run pairwise structural and sequence alignments.
- **Sequence Search**: Enter a PDB ID or sequence to query for similar sequences.
- **ADIT Deposition**: Provides a quick interface to the RCSB PDB ADIT deposition services. Users can start a new deposition, or continue an existing session.

Select a combination of widgets to display, and then move the widgets around to customize the display.

Structure Summary pages and query result pages can be similarly customized.



Customize the RCSB PDB homepage with different widgets.

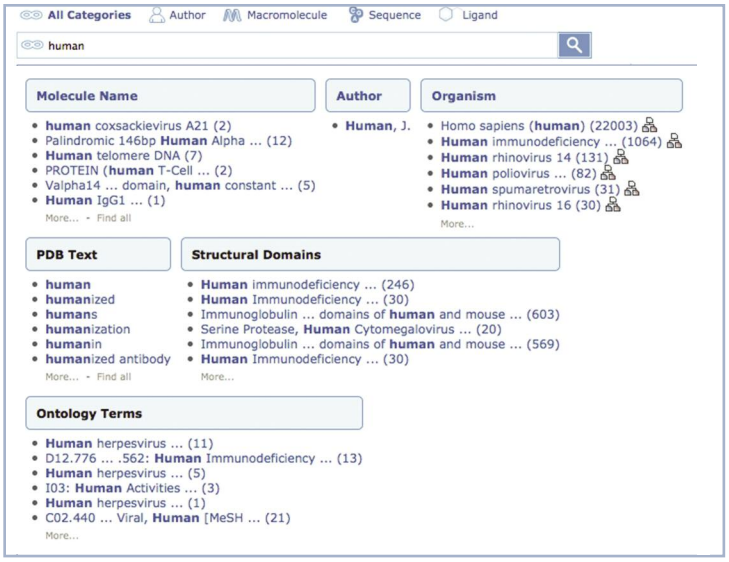
Use the Top Bar Suggestion Box for Quick and Precise Searching

The top search bar has been redesigned to help users easily and intuitively create precise searches.

Typing just a few characters launches an interactive pop-up box with suggestions of common PDB search terms, organized in different categories. These categories return more precise results than simple text searches. For example, entering the word *human* presents several options organized by category (such as Molecule Name, Author, Organism). Suggestions include the number of results and link to the set of matching structures. Results for organism *Homo sapiens* (human) will not include entries of author Human, J.

The redesigned top search can be also be limited to quick searches on Author, Macromolecule name, Sequence, or Ligand by selecting the related icon.

The order of results of a PDB text search or a sequence search is now based on the relevance of the term (for a text search) or the alignment score (for a sequence search). Search results can be further refined using Advanced Search or by drilling down through the pie charts organized by major categories.



Pop-up box showing suggestions for human

EMDB to Join the PDB Archive

The EM Data Bank (EMDB), the primary archive for experimentally-determined maps obtained using 3D electron microscopy methods, will join the PDB archive (ftp://ftp.wwpdb.org) on March 7, 2012. For more information, see wwpdb.org/em.

Time-stamped Copies of the PDB Archive

A snapshot of the PDB archive (ftp.wwpdb.org) as of January 2, 2012 has been added to ftp://snapshots.wwpdb.org/. Snapshots have been archived annually since January 2005 to provide readily identifiable data sets for research on the PDB archive.

The directory 20120102 includes the 78,237 experimentally-determined coordinate files and related experimental data that were available at that time. Coordinate data are available in PDB, mmCIF, and XML formats. The date and time stamp of each file indicates the last time the file was modified.

The script at ftp://snapshots.wwpdb.org/rsyncSnapshots.sh may be used to make a local copy of a snapshot or sections of the snapshot.

Outreach and Education

Papers Published

The December issue of *Nature Structural & Molecular Biology* contains a special essay collection called *Celebrating Structural Biology* (paywall). Coinciding with the 40th anniversary of the PDB, it features personal accounts that describe the history and predict the future of structural biology. Included are the reflections of RCSB PDB director Helen Berman on *The evolution of the Protein Data Bank*.

> Celebrating structural biology. Tom L Blundell, Stephen C Harrison, Robert M Stroud, Shigeyuki Yokoyama, Lewis E Kay, Michael G Rossmann, Helen M Berman, Alexander Wlodawer, Elena Conti, Brian Kobilka, Janet M Thornton, David Cowburn, Nenad Ban, Olga Boudker (2011) *Nature Structural & Molecular Biology 18*: 1304-1316 **doi: 10.1038/nsmb1211-1304**.

Other papers published:

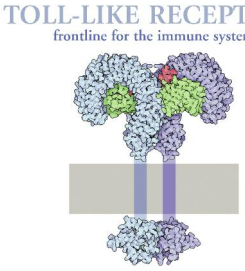
> Target highlights in CASP9: Experimental target structures for the critical assessment of techniques for protein structure prediction (2011) *Proteins: Structure, Function, and Bioinformatics* **79**: 6-20 **doi:10.1002/prot.23196**

> The Protein Data Bank: Evolution of a key resource in biology in *Leadership in Science and Technology: A Reference Handbook* (William Sims Bainbridge, editor) SAGE Publications, Inc., 2012.

Molecule of the Month News

The RCSB PDB's *Molecule of the Month* and related resources help promote a structural view of biology. Visit the latest features at PDB-101, including:

• **RSS Feed.** Subscribe to the PDB-101 RSS (Really Simple Syndication) Feed to access *Molecule of the Month* columns as they are published. This feed pushes the articles to RSS readers so users can easily view the latest articles. Access this feed by selecting the RSS icon PDB-101 header.



In their research, Beutler and colleagues discovered the role of a TLR that recognizes lipopolysaccharide. A downloadable flyer (PDF) highlights an example of this TLR structure using PDB IDs 3fxi[4] and 2j67[5].

• **Toll-like Receptors Flyer.** November's *Molecule of the Month* feature explores TLRs and the innate immune system in honor of the 2011 Nobel Prize in Physiology or Medicine. The award was divided, one half jointly to Bruce A. Beutler and Jules A. Hoffmann for their discoveries concerning the activation of innate immunity and the other half to Ralph M. Steinman for his discovery of the dendritic cell and its role in adaptive immunity.

• **ePubs** (Electronic Publication): Articles are now available as a downloadable ePub document, which can be viewed offline in ePub readers on mobile devices (such as iBooks for Apple mobile devices, and Aldiko for Android devices). To download a *Molecule of the Month* ePub, click on the link at the top of any feature.

• **Interactive Molecular Views for Mobile Devices:** PDB-101 Structure Focus pages, which highlight the PDB entries highlighted in *Molecule of the Month* articles, feature interactive molecular views.

Molecule of the Month *ePubs.*

The ability to rotate the molecules on these pages has now been enabled for mobile devices. For an example, select one of the Discussed Structures from the top of any article to view its Structure Focus page, or go to **http://bit.ly/uCXmTS** for ribonuclease A entry 5rsa and rotate the molecule about the Y-axis by dragging on the image left and right.

PDB40: A Special Celebration of the 40th Anniversary of the Protein Data Bank

On October 28-30, 2011, Cold Spring Harbor Laboratory hosted a unique meeting to commemorate the 40th anniversary of the PDB. PDB40's distinguished speakers described structural biology's past as well as current research in related fields. Close to 100 posters were presented. Travel awards were made for 34 early career scientists from around the world.

Selected presentations and links to a photo gallery have been posted to the PDB40 website at **wwpdb.org**. Future updates related to this meeting will be posted to this site.

The wwPDB organizers of this event are grateful to our industrial sponsors and funding agencies for their support of this meeting. Thanks to everyone who made PDB40 such a success.

Photos by Constance Brukin.

Education Corner

by the 2011 RCSB PDB Poster Prize Awardees

The RCSB PDB awards a poster prize for the best student poster presentation at selected meetings. This Education Corner looks at the educational experiences that influenced the 2011 awardees.

For more information on the RCSB PDB Poster Prize, please see ***http://bit.ly/rCZxOZ.***

Learning to Become a Structural Biologist

Briony Yorke

CURRENT POSITION: *PhD student, University of Leeds*

PAST EDUCATION: *Combined BSc and Masters, University of Leeds*

RCSB PDB Poster Prize Awardee at American Crystallographic Association (ACA) for New Approaches to Time-Resolved Structural Studies of Macromolecules

I am a Wellcome Trust-funded graduate student at the University of Leeds working in Arwen Pearson's laboratory on time-resolved X-ray crystallography. I have a collaborative supervisorial team including Michael Webb, Emanuele Paci (University of Leeds) and Robin Owen (Diamond Light Source). I am trying to develop methods to enable time-resolved X-ray crystallography on irreversible systems using the microfocus beamline at Diamond Light Source (I24). For my masters project at the University of Leeds, I worked on novel methods for ultra-fast time-resolved spectroscopy with Godfrey Beddard. I was introduced to macromolecular crystallography by Dr. Pearson who asked me to analyze radiation damage in myoglobin crystals using spectroscopic data for a rotation project.

When looking at the plethora of structures in the PDB, it is often easy to forget how much work goes into each experiment leading up to the point of structure determination. In the case of X-ray crystallography, the difficulty of growing crystals in the first place is a common complaint, as is the process of obtaining high-resolution data. However, the work that goes into the software packages used to process the data is something that now, more than ever, goes unnoticed. The analysis and processing of X-ray crystallography data requires the use of many algorithms which are more often than not hidden behind a user friendly GUI (graphical user interface). As a young crystallographer I have found myself wanting to understand how crystallographic software is written and also how it is designed to ensure it is user-friendly and intuitive.

The 2011 IUCr Crystallographic Computing School was held in Oviedo, Spain, with the theme Algorithms Development and Implementations. *The school is run by the Computing Commission immediately before the triennial IUCr Congress and General Assembly.*

The 2011 IUCr Crystallographic Computing School provided me with an opportunity to learn about the processes involved in software design. It was also instructive in the utilization of currently available software and libraries for performing novel analysis on X-ray data. I found that many of the tools I will need for my thesis project already exist behind the scenes of current software and I learned the techniques necessary to utilize them. Tutorial sessions with groups of around 10 people provided an opportunity to ask questions relating to specific projects but also to observe how an experienced computational crystallographer would work out the answer.

Tammy Cheng

CURRENT POSITION: *Postdoctoral Research Fellow at Cancer Research UK London Research Institute, UK*

PAST EDUCATION: *PhD in Biochemistry, University of Cambridge, UK; MSc in Chemistry, National Tsing-Hua University, Taiwan; BSc in Life Science, National Tsing-Hua University, Taiwan*

RCSB PDB Poster Prize Awardee at the Annual International Conference for Intelligent Systems for Molecular Biology (ISMB) for Structural biology meets systems biology: Gauging the systemic impact of non-synonymous single nucleotide polymorphisms

My enthusiasm in Structural Biology was first uncovered while an undergraduate student reading the classic textbook *Introduction to Protein Structures* by Carl Brandon and John Tooze. I found it fascinating that proteins implement their functions by delicately coordinating their 3D conformations, and the beautiful protein structures illustrated in the book left a deep impression on me. Since then I have been studying proteins ceaselessly, and that passion continues into my current post-doctoral research.

During my undergraduate study, I realized that my interest lay in studying proteins through a computational approach rather than wet-lab experiments. Before I finished my final undergraduate year, I decided to take a summer studentship at the Scripps Institute working on a protein structure viewer written in Python. As a result of this invaluable experience, I continued to explore computational algorithms for protein study. This led to a MSc Degree in Chemistry focusing on bioinformatics, and then a Ph.D. in Biochemistry specializing in computational biology.

With a deep interest rooted in protein structures, my Ph.D. research was concerned with the functional impact of proteins on a biological system. I studied non-synonymous single nucleotide polymorphisms (nsSNPs) and developed an algorithm, Bongo (**www.bongo.cl.cam.ac.uk/Bongo2/Bongo.html**), which analyzes the structural effect of nsSNPs on proteins. As an indirect approach to help estimate the possible phenotypic effect of nsSNPs, I was also involved in developing pyDock (**life.bsc.es/servlet/pydock/**), which predicts com-

plex protein structures assuming rigid-body protein-protein interactions.

In addition to focusing on the behavior of proteins in isolation, I am now taking a further step to study the impact of nsSNPs at the cellular level. My current research project simulates the systemic effect of nsSNPs by considering both protein structure and pathway dynamics. Complemented by experimental data, my colleagues and I have shown that the impact of a nsSNP on a specific pathway can be effectively quantified as a combination of protein stability change and pathway perturbation. I am grateful to all the collaborators involved in this project, which won the 2011 RCSB PDB Poster Award at the ISMB conference.

I have always thought that it is exciting to facilitate our understanding of the human genome by utilizing the knowledge of proteins, and this continues to motivate the direction of my research. I would like to take this opportunity to say 'thank you' to everyone who has ever worked on the PDB, which has always been an essential part of my research. I have no doubt that PDB will continue to play an important role in a wide range of research in the future.

Serah Kimani

CURRENT POSITION: *PhD student in Molecular and Cell Biology, University of Cape Town, South Africa*

PAST EDUCATION: *MSc in Structural Biology, University of Cape Town and the University of the Western Cape; BSc in Biomedical Sciences and Technology, Egerton University, Kenya*

RCSB PDB Poster Prize Awardee at the Congress and General Assembly of the International Union of Crystallography for Unexpected reactions resulting from mutating catalytic residues in an amidase reveal the role of the catalytic unit. *Serah also received the award at the Asian Crystallographic Association Meeting in 2007, and at ACA in 2010.*

I was born and raised in Gacharage, a small village in central Kenya. I did a BSc (Hons) in Biomedical Sciences and Technology at Egerton University (Kenya); a course that not only turned out to be over 98% theory due to lack of resources in our public universities, but also lacked a few key aspects of biochemistry like protein biochemistry. It is therefore not surprising that when I completed my degree I was completely ignorant of macromolecular structures. I saw a pamphlet advertising a new master's course in structural biology at the University of Cape Town (UCT, South Africa) and it promised a bursary, making it quite attractive even though for me it was clearly an adventure into a world unknown to me.

When I started at UCT, the learning curve was obviously quite steep given my background. I had to begin with the very basics of mastering the structures and properties of the 20 amino acids as well as learning how to use a computer, since I had never owned or worked at one before except to use email. Google and the internet in general became a friendly companion that slowly but surely helped to shorten the gap. I came face-to-face with a lot of challenging mathematics, physics and computer programming that I would never have thought were behind the structure-determining processes. But despite the difficulties, every unveiled detail was exciting, and I was amazed by the cool insights that one could obtain from structures. The idea of 'structure-based drug design,' for example, held great interest and appeal. In addition to traditional classroom lectures, web-based course materials from other universities (especially Birkbeck College, London) have formed a major part of my understanding of structural biology methods. I have also had the privilege of learning from experts and specialists in the field, who were flown into South Africa to give short courses and conduct workshops on various techniques, as well as share their work during the early days of the structural biology program at UCT.

Regular email correspondence with distinguished crystallographers and software developers has been a great resource in my research work. For instance, in my MSc dissertation project I had a difficult task of solving the crystal structure of a bacterial amidase[6,7] that had roughly 20% sequence identity with a few homologous structures in the PDB using molecular replacement phasing method. With extensive reading and expert guidance through endless email communications with scientists like Prof. Randy Reed (Cambridge, UK) and Prof. Thomas Terwilliger (Los Alamos National Laboratory, USA), among others, I was able to overcome problems related to phasing and structure refinement, and to solve one of the first crystal structures of an aliphatic amidase from the nitrilase superfamily of enzymes.

The CCP4 bulletin board (ccp4bb) has also been quite instrumental in my work. In my current PhD project, which is concerned with understanding the role of the catalytic residues in amidases, I have frequently run into the problem of protein misfolding and instability on mutation of some of the active site residues. The crystallographers on ccp4bb have willingly provided many handy tips on how to overcome this and numerous other problems. My attendance at both local and international conferences and workshops has been a means of keeping in touch with the current developments in structural biology, as well as sharing my work, while consultations with my colleagues have enhanced my research and continually shaped me to be a better scientist.

It is unfortunate that protein structures remain unknown and thus unutilized in most parts of Africa. It is my hope that after my training, I will go back to Kenya and contribute towards raising awareness of structures. An introduction of a protein biochemistry module in undergraduate biological courses at universities would be a great start.

References

[1]A. Ben-Shem, N. Garreau de Loubresse, S. Melnikov, L. Jenner, G. Yusupova, M. Yusupov (2011) The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* 334: 1524-1529.

[2]N. Nandhagopal, A. Simpson, J.R. Gurnon, X. Yan, T.S. Baker, M.V. Graves, J.L. Van Etten, M.G. Rossmann (2002) The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. *Proc.Natl.Acad.Sci.USA* **99**: 14758-14763.

[3]K.T. Dolan, E.M. Duguid, C. He (2011) Crystal structures of SlyA protein, a master virulence regulator of *Salmonella*, in free and DNA-bound states. *J.Biol.Chem.* **286**: 22178-22185.

[4]B.S. Park, D.H. Song, H.M. Kim, B.S. Choi, H. Lee, J.O. Lee (2009) The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **458**: 1191-1195.

[5]T. Nyman, P. Stenmark, S. Flodin, I. Johansson, M.Hammarstrom, P. Nordlund (2008) The crystal structure of the human toll-like receptor 10 cytoplasmic domain reveals a putative signaling dimer. *J.Biol.Chem.* **283**: 11861-11865.

[6]V.B. Agarkar, S.W. Kimani, D. A. Cowan, M. F.-R. Sayed, B. T. Sewell (2006) The quaternary structure of the amidase from *Geobacillus pallidus* RAPc8 is revealed by its crystal packing. *Acta Cryst.* F62, 1174-1178.

[7]S.W. Kimani, V.B. Agarkar, D.A. Cowan, M.F.-R. Sayed, B. T. Sewell (2007) Structure of an aliphatic amidase from *Geobacillus pallidus* RAPc8. *Acta Cryst.* D63, 1048–1058.

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

Rutgers, The State University of New Jersey
Center for Integrative Proteomics Research
174 Frelinghuysen Road
Piscataway, NJ 08854-8087

UCSD SDSC SKAGGS SCHOOL of PHARMACY and PHARMACEUTICAL SCIENCES

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

WORLDWIDE PDB PROTEIN DATA BANK

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.*

RCSB PROTEIN DATA BANK
www.pdb.org • info@rcsb.org
Department of Chemistry and Chemical Biology
Rutgers, The State University of New Jersey
610 Taylor Road
Piscataway, NJ 08854-8087
USA

Return Service Requested