

Contents

Message from the RCSB PDB..... 1

DATA DEPOSITION AND PROCESSING

Restarting ADIT Depositions 2

Weekly Deadlines for Release/Modify Entry Requests 2

Deposition Statistics 2

Depositing and Releasing Experimental Data..... 2

DATA QUERY, REPORTING, AND ACCESS

New Website Features 2

Time-stamped Copies of PDB Archive Available via FTP 3

RCSB PDB Focus: Saving Protein Workshop "States" for Future
Visualization Sessions 4

Website Statistics 4

OUTREACH AND EDUCATION

Citing Structures in the PDB: IDs, Citations, and DOIs. 4

New Information and Statistics Available at BioSync 4

Making Virus Models with Middle School Students 4

Molecules of the Quarter 4

EDUCATION CORNER: 2007 New Jersey

Science Olympiad 5

PDB COMMUNITY FOCUS: Angela Gronenborn, Ph. D. . . 6

**RCSB PDB PARTNERS, MANAGEMENT,
AND STATEMENT OF SUPPORT..... 8**

SNAPSHOT: APRIL 1, 2007

42474 released atomic coordinate entries

MOLECULE TYPE	EXPERIMENTAL TECHNIQUE
39000 proteins, peptides, and viruses	36086 X-ray
1713 nucleic acids	6159 NMR
1726 protein/nucleic acid complexes	144 electron microscopy
35 other	85 other
	25371 structure factor files
	3377 NMR restraint files

Participating RCSB Members:

Rutgers • SDSC/SKAGGS/UCSD

E-mail: info@rcsb.org

Web: www.pdb.org • FTP: <ftp.rcsb.org>

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

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

Message from the RCSB PDB

The RCSB PDB fosters communication with users to support access to the information contained within the PDB archive. At the same time, our outreach efforts are focused on soliciting input from the user community to help improve RCSB PDB services. Frequently, this dialog happens at workshops and professional society meetings.

Recent activity has included:

- The Keystone Symposia "Frontiers of NMR in Molecular Biology" (January 6-11 in Snowbird, Utah). A workshop entitled *The Future of Publicly-Accessible Databases for NMR Spectroscopy* was held to discuss issues surrounding deposition and data representation in the PDB and BMRB. These issues were further examined at a meeting of the PDB-BMRB joint NMR Task Force held after the workshop.



Visitors to the RCSB PDB exhibit booth at the Biophysical Society Meeting

- Annotators Jasmine Young and Monica Sekharan exhibited at the 51st Annual Meeting of the Biophysical Society (March 3-7 in Baltimore, Maryland).

- The RCSB PDB also exhibited at the Celebration of Teaching & Learning, an education-related professional development conference for teachers, administrators, and others (March 23-24 in New York City).

We hope to see many of you at our upcoming meetings, some of which are highlighted below:

- The Experimental Biology Annual Meeting for several professional societies, including the American Society for Biochemistry and Molecular Biology (April 28 - May 2 in Washington, DC).

- The 5th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB) & 6th European Conference on Computational Biology (July 19-25 in Vienna, Austria).

- The meeting of the American Crystallographic Association (ACA; July 21-26 in Salt Lake City, UT).

- The 21st Annual Symposium of The Protein Society (July 21-25 in Boston, MA).

- The 9th International Conference on Biology and Synchrotron Radiation (August 13-17 in Manchester, United Kingdom).

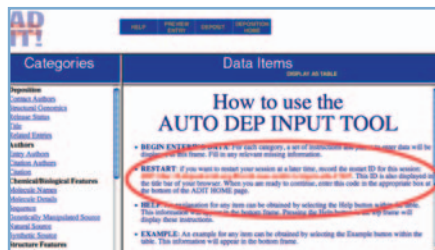
- The 234th American Chemical Society National Meeting (August 19-23 in Boston, MA).

The RCSB PDB Poster Prize will also be awarded at ISMB, ACA, and the the European Crystallographic Meeting.

Data Deposition and Processing

Restarting ADIT Depositions

A structure can be deposited in more than one Internet session by using ADIT's "Session Restart ID" feature. This identifier appears in red in the center of the browser window when ADIT's "deposit" step is first started. It is also seen in the title of the browser throughout the deposition session.



The ADIT Session Restart ID is circled.

the restart ID until the "DEPOSIT NOW" button is selected, for up to six months after the session has been last updated.

ADIT is available at the RCSB PDB and PDBj. ADIT-NMR can be used to deposit data to the PDB and BMRB at the same time.

ADIT tutorials (in English and Japanese) and example "in progress" deposition sessions are accessible from deposit.pdb.org.

Weekly Deadlines for Release/Modify Entry Requests

PDB entries are processed by three members of the wwPDB (RCSB PDB, MSD-EBI, and PDBj).

Each week, all files scheduled for release or modification are checked and validated one final time. Authors may be contacted to resolve any issues that may arise while preparing the entries for release.

When the release of a structure on hold for publication (HPUB) is requested, the wwPDB routinely confirms the primary citation. If this is not accomplished within that release cycle, the entry may be scheduled for release in a later update.

To be included in the next weekly update, any required author correspondence should be sent by 15:00 (local time) on Thursdays to the appropriate wwPDB member:

- RCSB PDB (help@deposit.rcsb.org)
- MSD-EBI (pdhelp@ebi.ac.uk)
- PDBj (adit@adit.protein.osaka-u.ac.jp)

All entries due for release are transferred to the RCSB PDB for final packaging into the master PDB ftp archive. These files are then released by 4:00 ET each Wednesday.

Requests received after these cutoff times will be processed during the next update cycle.

Deposition Statistics

In the first quarter of 2007, 2319 experimentally-determined structures were deposited to the PDB archive. Of these structures, 62.7% were deposited with a release status of HPUB, 22.3% with a release status of hold until a particular date, and 15.0% were released as soon as annotation

of the entry was complete.

80.8% of these entries were determined by X-ray crystallographic methods; 18.8% were determined by NMR methods. 80.3% of these depositions were deposited with experimental data.

92.8% of the crystal structures were deposited with structure factors; 27.7% of NMR structures were deposited with restraints.

Depositing and Releasing Experimental Data

The RCSB PDB strongly encourages depositors to follow the guidelines regarding the submission and release of coordinate and experimental data that have been set by the International Union of Crystallography, the National Institutes of Health, and the journals.

Deposition of experimental data (structure factor and/or NMR constraint files) is required by many journals, including *Acta Crystallographica*, *Biochemistry*, *Cell*, *Nature*, and *Science*.

These files can be uploaded during the ADIT deposition process.

Depending upon the hold status selected by the depositor, data release can occur when a depositor gives approval, the hold date has expired, or the journal article has been published. There is a one-year limit on the length of a hold period, including HPUBs. If the citation for a structure is not published within the one-year period, depositors will be given the option to either release or withdraw the deposition.

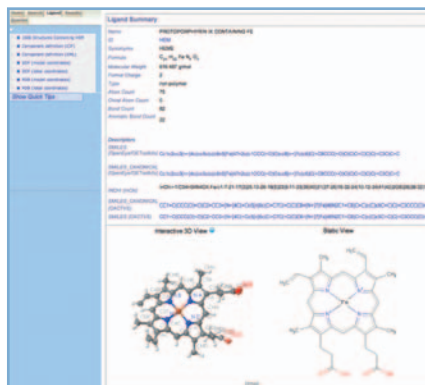
Detailed deposition and release information is at deposit.pdb.org.

Data Query, Reporting, and Access

New Website Features



The Ligand Hits tab from a search for "protein kinase".



The Ligand Summary page offers summary information, downloads (definitions and coordinates), and interactive and static views. HEM is shown here.

Many enhancements were made for data query and reporting this quarter.

• Improved Access to Ligand Data

The PDB chemical component dictionary (formerly the HET dictionary) has been remediated to better describe the components that interact with macromolecular structures. This new dictionary has been incorporated with the RCSB PDB database.

A new search results option is a tab called "Ligand Hits". This page lists the ligands known to interact with the structures that match the query.

For example, a search for "protein kinase" returns 2045 structures and 676 ligands. From the "Ligand Hits" page, users can find all

of the structures that contain that ligand or access information from the “Ligand Summary” page. This page offers summary information, downloads (definitions and coordinates), and interactive and static views.



A look in Ligand Explorer at a hydrophobic binding pocket for ATP in entry 1ATP.

• Ligand Explorer Tool for Viewing Protein-Ligand Interactions

Ligand Explorer is a Java-based program accessible from each Structure Summary page. Features include the ability to highlight ligand interactions based on conventional and user-defined thresholds, and a “contact map” that gives users the ability to see the details of each interaction.

• **Access to Single Nucleotide Polymorphism (SNP), Pfam, and More** SNP information is now accessible from Structure Summary pages. Over 4000 PDB structures are linked to SNP information from the SNP database. This information is accessible from each entry’s “Biology and Chemistry Report” tab.

The Pfam database contains multiple alignments of protein domains. With each release of the Pfam data, files mapping Pfam domains to PDB structures are made available on the Pfam FTP site. This mapping is loaded into our database so that Pfam domain information for a protein structure is displayed on an entry’s Structure Summary page and Biology and Chemistry Report, when available.

The “External Links” option provides further information about the structure under study, such as biochemical pathway information, stereochemistry and ligand binding data. When looking at an entry’s Structure Summary page, the external links page is accessible from the left-hand menu.

SNP: www.ncbi.nlm.nih.gov/projects/SNP

Pfam: www.sanger.ac.uk/Software/Pfam

• New Advanced Search Options

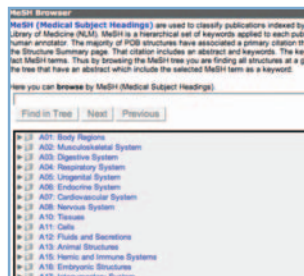
Simple searches of the RCSB PDB website can be performed using the keyword box at the top of each page. The “Advanced Search” feature makes more specific and complex searches possible.

New options can be used to search the PDB using keywords, phrases or a series of keywords:

– **Advanced Keyword Search:** This option can be used to search for keywords in the full text or by the author name. If you enter a phrase, you must place it in quotes otherwise it will be interpreted as a series of keywords. Advanced keyword search supports the lucene syntax for sophisticated string searching.

– **Medical Subject Headings (MeSH):** Searches for structures associated with particular MeSH terms from the National Library of Medicine (NLM). This option launches the MeSH Browser, which lets users either browse through the MeSH hierarchical tree or search the tree with keywords.

For example, the top folder “C11: Eye Diseases” can be opened to find structures indexed as relating to eye diseases, specifically cataracts (MeSH Number C11.510.245).



The MeSH browser.

– **Author Assigned:** Looks for structures based upon keywords used by the depositor.

– **PubMed:** Searches PubMed titles and abstracts for an entry’s primary citation (if it exists).

• **New Help Features**

A new set of Flash Tutorials, modeled on the popular guides on how to use the RCSB PDB overall site and the Advanced Search, are available.

These include tutorials for the MeSH Browser, Protein Workshop, KiNG, Jmol, and general navigation. They are accessible from the left-hand menu under “Site Tutorials”.

Quick Tips offer hints and quick links to exploring the RCSB PDB website. To view them, click on the “Show Quick Tips” in the left-hand menu. Clicking on the arrow button will scroll through these hints, and clicking on the “X” will close the box.

Please write to info@rcsb.org with any questions or comments about these new features.

Time-stamped Copies of PDB Archive Available via FTP

A time-stamped snapshot of the PDB archive as of January 2, 2007 has been added alongside time-stamped copies of the archive from January 2006 and 2005 at ftp://snapshots.rcsb.org. It is hoped that these snapshots will provide readily identifiable data sets for research on the PDB archive.

The directory 20070102 includes the 40,933 experimentally-determined coordinate files that were current (i.e., not obsolete) as of January 2, 2007. Coordinate data are provided in PDB, mmCIF, and XML formats. The date and time stamp of each file indicates the last time the file was modified.

Scripts can be used to automatically download data:

- ftp://snapshots.rcsb.org/rsyncSnapshots.sh
Makes a local copy of an annual snapshot or sections of the snapshot. Downloading the entire archive can be lengthy (more than 18 hours), but the time required to download data in a single format should be much less. Depending upon network speed, our tests show that all of the coordinate files in PDB format from a snapshot can be downloaded in about 2 1/2 hours.
- ftp://ftp.rcsb.org/pub/pdb/software/rsyncPDB.sh
Copies the current contents of the entire archive.
- ftp://ftp.rcsb.org/pub/pdb/software/getPdbStructures.pl
Copies portions of the current archive.
- ftp://ftp.rcsb.org/pub/pdb/software/getPdbUpdate.pl
Copies the data from the weekly updates.

RCSB PDB Focus: Saving Protein Workshop “States” for Future Visualization Sessions

Protein Workshop is a molecular viewer accessible from every PDB entry’s Structure Summary page. Its simple interface lets users quickly and easily select structural elements and change the coloring, labeling, and representation style (ribbons, cylinders, and more). Users can also color specific structural features such as conformation type and hydrophobicity.

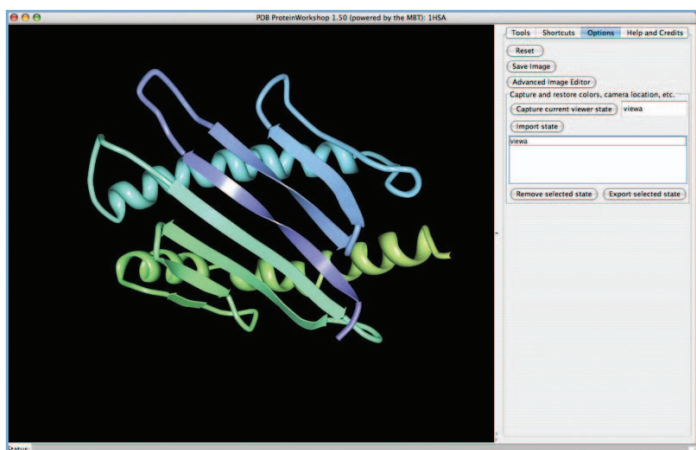
Protein Workshop is an excellent tool for generating high-resolution images in JPG, BMP, TIFF, WBMP, and PNG formats. A tutorial for using Protein Workshop and creating these images is available.

Protein Workshop offers a way to save the “state” of a session. Users can rotate and zoom a structure to a particular orientation and then capture

this view for later use. To save a state, enter a title next to the “Capture current viewer state” from the Options menu, and then select the adjacent button. The name of this state will be listed in the box below. The view of the molecule can then be changed around, but users can always go back to saved states by clicking on the state’s name.

These states can be saved in a XML file for later use by selecting the state and clicking the “Export selected state” button. States can be restored from a file by clicking the “Import state” button.

This tool uses the Molecular Biology Toolkit (mbt) and JOGL technology, and requires no installation other than the most recent version of Java.



A section of MHC as viewed in Protein Workshop. PDB ID: 1hsa. D.R. Madden, J.C. Gorga, J.L. Strominger, D.C. Wiley. (1992) *The three-dimensional structure of HLA-B27 at 2.1 Ångstrom resolution suggests a general mechanism for tight peptide binding to MHC.* Cell 70:1035-1048.

Website Statistics

Access statistics for the first quarter of 2007 are given below for the RCSB PDB website at www.pdb.org.

MONTH	UNIQUE VISITORS	NUMBER OF VISITS	BANDWIDTH
JANUARY	124,450	299,114	503.06 GB
FEBRUARY	118,164	282,552	449.78 GB
MARCH	125,905	298,491	472.08 GB

Outreach and Education

Citing Structures in the PDB: IDs, Citations, and DOIs

The contents of the PDB are in the public domain. Structures can be cited using their PDB ID and the published citation related to the structure.

- Structures may also be referenced using their Document Object Identifier (DOI). The DOIs for PDB structures all have the same format - 10.2210/pdbXXXX/pdb - where XXXX should be replaced with the

desired PDB ID. For example, the DOI for PDB entry 4HHB is “10.2210/pdb4hbb/pdb”.

This DOI can then be used as part of a URL to obtain the entry’s compressed data file in PDB format (<http://dx.doi.org/10.2210/pdb4hbb/pdb>), or can be entered in a DOI resolver (such as <http://www.crossref.org>) to automatically link to `pdb4hbb.ent.Z` in the PDB ftp archive (<ftp://ftp.rcsb.org>).

- The journal reference for the RCSB PDB is: H.M.Berman, J.Westbrook, Z.Feng, G.Gilliland, T.N.Bhat, H.Weissig, I.N.Shindyalov, P.E.Bourne. (2000) The Protein Data Bank, *Nucleic Acids Research* 28:235-242.

- The journal reference for the wwPDB is: H.M. Berman, K. Henrick, H. Nakamura. (2003) Announcing the worldwide Protein Data Bank. *Nature Structural Biology* 10:980.

Detailed information for citing the use of data, structures, and images is available from the RCSB PDB home page under “General Information”.

New Information and Statistics Available at BioSync

The BioSync website now contains updated beamline descriptions for operational US synchrotron beamlines as well as some basic information for almost all operational international beamlines.

PDB deposition statistics, grouped by site and beamline, can be found at biosync.rcsb.org. Galleries of structures, also grouped by site and beamline, are cross-linked to Structure Summary pages in the RCSB PDB. Tables of primary citations and some general information (phasing software, resolution, R-factors, etc.) are also provided. Most recently, similar tables and galleries have been added for structural genomics structures solved from synchrotron data.

BioSync (Structural Biology Synchrotron Users Organization) was formed in 1990 as a grassroots organization intended to promote access to synchrotron radiation. The BioSync resource, originally designed and hosted by UCSD/SDSC, has been updated and is now being maintained by the RCSB PDB.

Updates to beamline descriptions from local personnel, general comments, and suggestions should be sent to BioSync@deposit.rcsb.org.

Making Virus Models with Middle School Students



Annotators helped students create virus structures.

Annotators helped middle school students in New Jersey build 3D models of virus structures as part of Princeton University’s Science and Engineering Expo (March 22). For a copy of the template used to create paper virus models, please email info@rcsb.org.

Molecules of the Quarter

The *Molecule of the Month* series explores the function and significance of selected biological macromolecules for a general audience. The molecules featured this quarter were importins, exosomes, and zinc fingers. The complete features are accessible from www.pdb.org.

Education Corner

2007 New Jersey Science Olympiad

Science Olympiad tournaments, which take place across the country, consist of a series of individual and team events that students prepare for during the year. During this competition, teams demonstrate their diverse skills and knowledge in many different events. In *Forensics*, teams identify polymers, solids, and fibers at a crime scene, while in *Write It, Do It*, students compose a description of a structure that will be the only guide used by their other team members to recreate that structure (sight unseen) with raw materials.

High school teams at the New Jersey Science Olympiad (NJSO) demonstrated their understanding of structure and function in the 2007 *Protein Modeling* trial events that were sponsored by the RCSB PDB.

In this event, students identify key elements of a structure and demonstrate their knowledge of the protein by creating a three-dimensional model using Mini-Toobers, computer visualization tools, and RCSB PDB resources. The model is accompanied by a brief abstract that highlights the features shown in their model and discusses what the protein does. At the competition, teams also answer multiple choice and short answer-questions focusing on its structure and function.

As one team described, the students enjoyed “replicating protein molecules that are found in the body into real-life toober models”.

The entries are judged by the RCSB PDB annotators using a model built directly from the structure’s PDB file and a predetermined rubric that awards points for accurate depictions of the protein’s features. For example, judges look to see if the N- and C- terminus are labeled properly and



The Mini-Toober kit (center) was transformed into a section of the MHC structure.

AN EXCERPT FROM THE DIARY OF INSU LIN

I have been diagnosed with diabetes Type I! The doctor told me I would have to inject insulin into my body because I don't have enough. He also gave me a pamphlet explaining insulin. This helped me realize how important (and overlooked) this small hormone is in our bodies.

Insulin, a hormone secreted by the pancreas, regulates glucose levels. It is comprised of two chains (α and β). When glucose levels rise, insulin is secreted from the pancreas, and binds to receptors on the cell membrane of glucose requiring cells. This causes transport proteins in the cell membrane to take up glucose from the bloodstream. Essentially, insulin initiates blood glucose homeostasis.

The three disulfide bonds in insulin are critical to its function – two of which are between the two chains, while one is within the α chain. They connect the two chains of amino acids that make the insulin and help determine the tertiary and quaternary structure of insulin. If the chains did not fold into their natural positions, the hormone would not function properly, since the insulin molecule would not have the structure required to bind to receptors. Therefore, insulin would not be able to function.

The sequence of amino acids is the primary structure of the protein. According to the principles of protein folding, hydrophilic (polar) amino acids are positioned on the exterior of the protein, while hydrophobic (non-polar) amino acids are on the interior. If the primary structure was to be altered by even one amino acid in the residues that bind to the insulin receptor, insulin wouldn't fold into exactly the right shape. This would inhibit the insulin's ability to bind to cell receptors.

The structure of insulin directly impacts the hormone function. If there were to be any critical change in primary structure or disulfide bond formation, the folding pattern of insulin could be altered. It would not bind to cell receptors and thus not help manage cell metabolism.

I built an insulin model to better understand insulin and the non-polar amino acids (white tape) are in the center, and polar amino acids (blue tape) surround them. As represented in my model, acidic and basic amino acids remain close to neutralize each other. This is shown in the central helix of the beta chain.

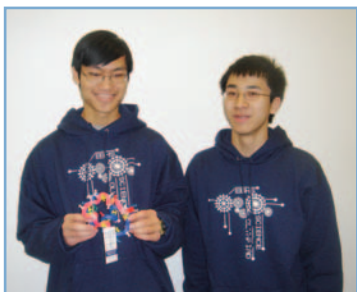
This excerpt is part of the abstract submitted by West Windsor-Plainsboro's team – Aleesha Shaik, Linda Maa, Ilya Podkopaev – at the regional competition. It has been edited slightly.

carefully consider the helices of the model. They also consider if the main functional and structural features of the protein are described in the written abstract. The written exam asks questions based upon the entry's Structure Summary page, the Molecule of the Month entry, and beyond. In 2007, teams built an insulin structure (PDB ID 4hiu) for the regional competitions held in January, and a section of a major histocompatibility complex (MHC) structure (PDB ID 1hsa) for the state competition in March. The hand-built models were really impressive, and the written abstracts and exams exhibited that many teams were quite scientifically literate.

At the Central New Jersey regional, East Brunswick High School (First Place and the 2006 State Champions in this event), West Windsor-Plainsboro South High School (Second), and West Windsor-Plainsboro North High School (Third) created very strong models.



The judges use a model of the insulin structure and a rubric to help score the entries.



*Central New Jersey regional champions
Leebyn Chong and Anthony Sin of East
Brunswick High School.*



*Northern New Jersey regional champions
Marina Mainescu, Benjamin Yang, and
Edward Hong of Bergen County Academy.*



*Ola Hadaya, Sarah Goodman, and Yong
Kim from Princeton High School.*

At the Northern New Jersey regional, Bergen County Academy (First Place), Westfield High School (Second), and New Providence High School (Third) exhibited very strong skills.

At the state finals, students from all over the garden state competed. The highest ranked teams were Princeton High School (First Place), Montgomery High School (Second), and The Lawrenceville School (Third).

The Science Olympiad is an international nonprofit organization devoted to improving the quality of science education, increasing student interest in science and providing recognition for outstanding achievement in science education by both students and teachers. The 2007 NJSO (www.njscienceolympiad.org) was presented by the New Jersey Science Teachers Association and the New Jersey Science Education Leadership Association. Special thanks to the Center for BioMolecular Modeling at the Milwaukee School of Engineering (www.rpc.msos.edu/cbm) for the design of this event. Kits similar to those provided for this event may be purchased from www.3dmoleculardesigns.com. Questions about the NJSO Protein Modeling trial event should be sent to buildmodels@deposit.rcsb.org.

A website with information and resources for participating in the protein modeling event can be found at education.pdb.org/olympiad.



PDB Community Focus:

**Angela Gronenborn,
University of Pittsburgh**

Q: How would you compare X-ray and NMR methods for determining structure?

A: I truly believe both methods are complementary. Each provides a model for the 3D structure of a molecule and as such, each presents a picture of the spatial architecture, one in the solid state and the other in solution. Naturally, the environment and conditions in which the structural studies are conducted will influence the outcome to a certain degree. pH, temperature, and ionic strength are rarely identical if both methods have yielded structures, and details may vary accordingly. For example, sidechain orientations may differ depending on the protonation state, and loop regions may get 'locked in' in the crystalline state. In addition, since a much larger degree of order is required for crystals to form, the oligomerization state may be different in solution and the crystal. Indeed, there are numerous examples of proteins for which dimers and higher oligomers are observed by X-ray crystallography, but the solution NMR structures are monomeric.

In terms of methodological maturity, it is evident that X-ray crystallography is 25 years ahead of NMR as a structural method, thus it is a robust method. This is reflected in the significantly larger numbers of X-ray structures in the PDB compared to NMR structures. If one looks at the growth rate, however, I believe NMR follows exactly the trend that was seen 25 years ago in the crystallographic field. Structural NMR is still evolving, with novel and advanced approaches being introduced all the time. A case in point was the introduction of R(esidual) D(ipolar) C(coupling)-based methodologies that led to better defined structures and allows for unambiguous positioning of relative structural elements.

Q: What aspects of structural biology are more accessible by NMR than X-ray methods?

ANGELA GRONENBORN, PH. D. is one of the country's leading structural biologists and an internationally renowned specialist in the application of nuclear magnetic resonance (NMR) spectroscopy for investigating structure, dynamics and folding of biological macromolecules. She joined the faculty of the University of Pittsburgh as a Professor in the School of Medicine in 2004. In 2005, the Department of Structural Biology was established with Prof. Gronenborn holding the Rosalind Franklin Professorship and Chair. The department is located in the new Biomedical Science Tower, housing state of the art equipment devoted to NMR spectroscopy, X-ray crystallography, and cryo-electron microscopy.

Prior to her move to Pittsburgh, Prof. Gronenborn was a member of the Senior Biomedical Research Service and Chief of the Structural Biology in the National Institute of Diabetes and Digestive and Kidney Diseases at the National Institutes of Health (NIH). She received both her undergraduate and Ph.D. degrees from the University of Cologne, Germany. After post-doctoral training she joined the Scientific Staff in the Divisions of Molecular Pharmacology and Physical Biochemistry at the National Institute for Medical Research, Mill Hill, London. In 1984, she moved to the Max-Planck Institute in Munich as head of the Biological NMR Group, and in 1988 to the NIH.

Prof. Gronenborn's research harnesses the power of NMR in two major areas: understanding biochemical mechanisms and the structural basis of cellular regulation as well as HIV pathogenesis. She has authored more than 350 publications, including structural studies on interleukins, chemokines, the tumor suppressor protein p53, various transcription factors and enzymes, and a number of HIV-encoded proteins including integrase and protease. She also is noted for her contributions to advancing technology on how best to apply NMR to elucidate important problems in the biosciences.

A: As we all know, the rate-limiting step in X-ray crystallography frequently is the time it takes to obtain well-diffracting single crystals – NMR solution structural work is not hampered by this requirement. Crystallization may be prevented if, for instance, a protein is very flexible

or contains mobile regions, but NMR can investigate such ‘floppy’ proteins. Examples of this type are folding intermediates or partially folded proteins, for which NMR is probably the only method that allows one to carry out structural characterizations (see for example ref.¹).

In addition, structures of weakly interacting systems are another area where NMR excels. Tight binding is often required for complexes to be amenable to crystallization, and exchanging systems present major challenges (sometimes overcome by cross-linking the components). NMR can deal with exchanging systems and structures of “weak” complexes can be determined (see for example ref.²). This property of NMR was exploited early on in studies of protein-ligand complexes and the transferred NOE methodology has been widely used in pharmaceutical applications.

Q: *What were the most exciting projects in which you have been involved?*

A: There have been numerous exciting projects all along the way – and I still can get thrilled about seeing a new structure for the first time or coming up with some crazy idea.

One exhilarating period that comes to mind was the late eighties/early nineties when we were all in the bowels of Building 2 at the NIH. Marius Clore and myself had just moved from the Max-Planck. Together with Ad Bax, who already was working there, and a combined group of congenial post-docs, we developed and implemented 3- and 4D NMR and its application for protein structure determination.

Also, my work on cyanovirin (CVN) – starting with the initial structure of a protein whose sequence had no relatives in any database via dissecting its folding and domain-swapping, to carbohydrate binding and the structural basis of its anti-HIV activity has kept me captivated for years. Indeed, it inspired me to embark on a fishing expedition – for its gene – which in turn has now led to the discovery of CVN homologs in truffles and plants, whose structures we are currently working on.

Q: *You are the chair of the BMRB Advisory Committee. How do you see the interaction of the BMRB with the PDB?*

A: The BMRB is the arm of the PDB that deals with NMR structure determinations. Its mission is to collect, archive, and disseminate data derived from NMR spectroscopic investigations of biological macromolecules. Initially structures, irrespective of their origin (*i.e.*, the method with which they were determined) were all deposited in the PDB. Likewise, the constraints that were used for generating these structures were archived there, whereas other NMR parameters such as chemical shifts were deposited at BMRB. Now that BMRB is part of wwPDB, it is becoming more of an integral part of the overall depository and it is envisaged that no matter at which site a deposition is made, all data will be retrievable in identical format from any of the sites. I strongly believe that an open and collegial atmosphere is essential for science to thrive, unencumbered by national interests – to that end the PDB data is maintained as a single global archive whose contents is freely and publicly available. The NMR community and BMRB has a stake in contributing to the international and interdisciplinary nature of the PDB by providing and archiving high quality and consistent NMR data, thereby playing an active role in the continued success of this global resource.

Q: *How has NMR changed and/or advanced as a technique over the course of your career?*

A: In one word – tremendously! Just considering instrumentation alone, we have seen unbelievable advances. During my thesis work I recorded spectra on 60MHz spectrometers, with the HA100 our most advanced instrument. When I started in “biological NMR” during my post-doctoral training, 2D NMR had not been invented and protein NMR was in its

infancy. We were thrilled to see individual resonances for histidine residues in 1D proton spectra! High field magnets were 270 MHz; and we would read in pulse programs from paper tapes. These days we routinely record heteronuclear, multidimensional (2-, 3- and 4D) spectra on uniformly (¹³C/¹⁵N/²H) labeled biological macromolecules at 600MHz, with 700, 800 and 900 MHz magnets available in regional and national NMR facilities for the most demanding applications and samples. Whereas initially it was feared that NMR would be limited to studying proteins of less than ~100 amino acids – certainly true if one were restricted to proton 2D NMR only – we now can tackle much larger systems. Indeed, structure determinations of systems up to ~100kDa are possible and good quality spectra have been obtained for much bigger systems like the 800-kDa tetradecameric GroEL chaperone. Using selective or segmental labeling approaches in conjunction with clever spectroscopy no doubt will lead to further advances – we haven’t reached our limits yet.

Q: *What do you see as the future of NMR?*

A: I believe we will see continuous advances in the development of NMR methodology as well as its application to more and more challenging systems. As alluded to above, NMR is particularly powerful to look at dynamic systems, such as unfolded or partially folded proteins, exchanging systems or regions in multi-component systems that are flexible. I predict that we will see increasing activities in the NMR community directed at characterizing protein dynamics and relating motions to biological activities.

Q: *You recently moved to the University of Pittsburgh to chair the structural biology department there. What were the challenges involved in setting up that program?*

A: As is often the case when one embarks on a new venture, one is excited about the opportunities. Challenges and opportunities usually go hand in hand, and setting up Structural Biology at the University of Pittsburgh Medical School presented a tremendous opportunity. I firmly believed that it was necessary to embrace the continuum of structural methodologies if one wanted to understand biological systems at the most fundamental level – cryo-electron microscopy, X-ray crystallography and NMR – and needed to assemble faculty and instrumentation of the highest caliber. Naturally, this comes at considerable cost. Fortunately for Pittsburgh and myself, the institution and the Senior Vice Chancellor for the Health Sciences, Art Levine, not only shared this view, but had already embraced structural biology as being crucial for medical research. As a result, he supported me all along the way. We have now completed the first round of hiring, and the new Structural Biology Department already has a number of faculty who moved to Pittsburgh from all over the US and Europe. We are housed in a state-of-the-art new research building, BST3, that was designed to optimally site all our instrumentation. Despite its location on one of the busiest streets in Pittsburgh, 5th Avenue, the NMR spectrometers and electron microscopes are shielded from vibrations and other environmental influences. Working with an extremely competent and responsible architectural firm made this possible – as well as getting natural light into the NMR facility, a pleasure for those of us who have worked underground for most of our careers. We are now looking forward to carrying out exciting new biomedical research and entering into adventurous collaborations with our clinical and biological colleagues.

¹ Byeon, I. J., Louis, J. M. & Gronenborn, A. M. (2004) A captured folding intermediate involved in dimerization and domain-swapping of GB1. *J Mol Biol* 340:615-25.

² Garrett, D. S., Seok, Y. J., Peterkofsky, A., Gronenborn, A. M. & Clore, G. M. (1999) Solution structure of the 40,000 Mr phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. *Nat Struct Biol* 6:166-73.

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 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. PHILIP E. BOURNE, Co-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.pdb.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 Center for Research Resources, the National Institute of Biomedical Imaging and Bioengineering, the National Institute of Neurological Disorders and Stroke, and the National Institute of Diabetes & Digestive & Kidney Diseases.*

RCSB PROTEIN DATA BANK www.pdb.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