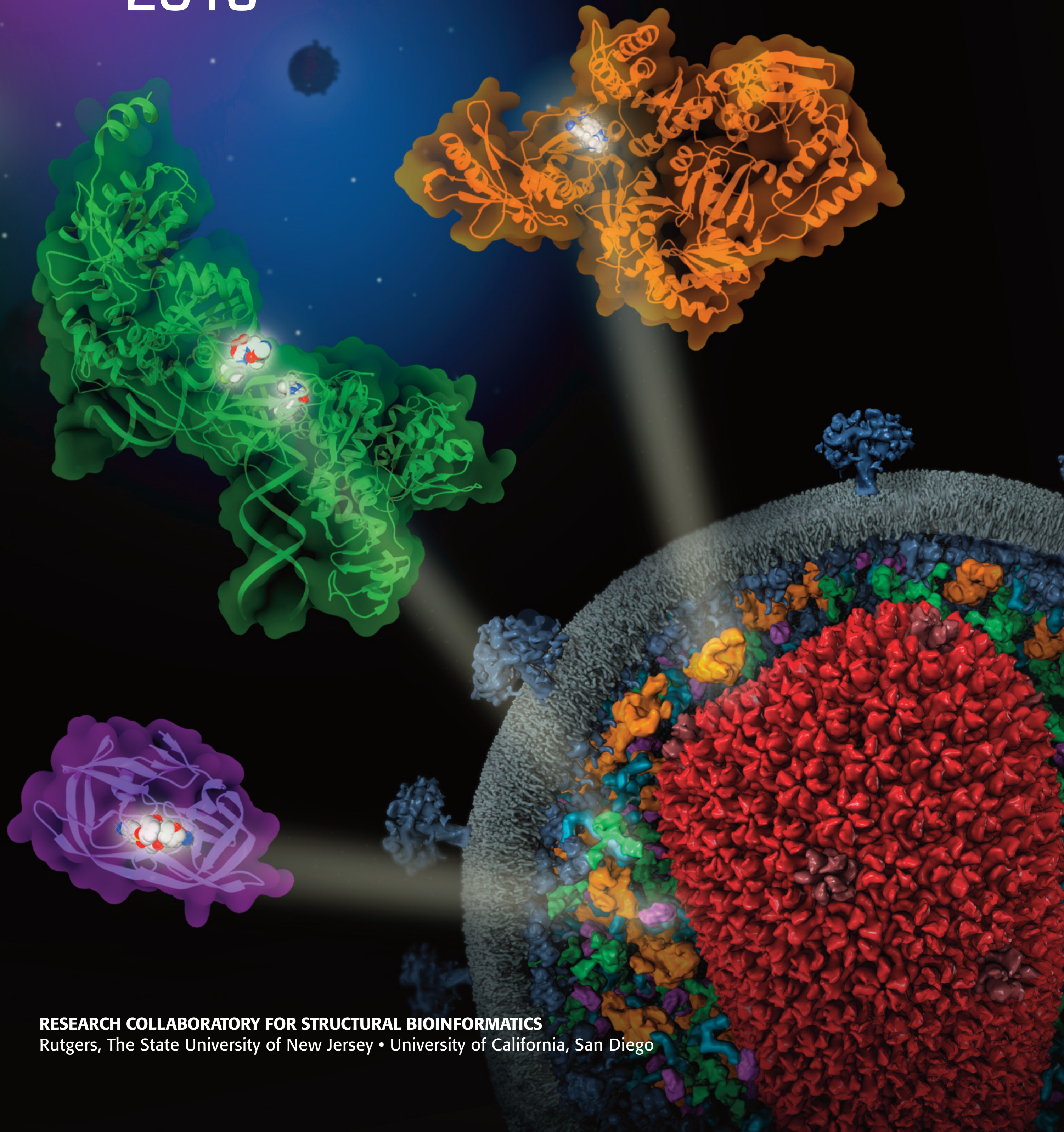
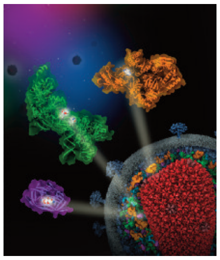
RCSB PDB
PROTEIN DATA BANK

ANNUAL REPORT 2013



RESEARCH COLLABORATORY FOR STRUCTURAL BIOINFORMATICS
Rutgers, The State University of New Jersey • University of California, San Diego

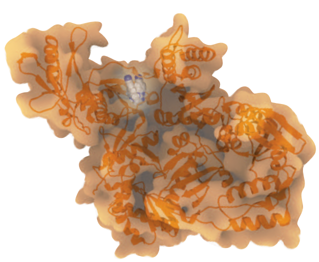


HIV (human immunodeficiency virus) is composed of two strands of RNA, 15 types of viral proteins, and a few proteins from the last host cell it infected, all surrounded by a lipid bilayer membrane. Together, these molecules allow the virus to infect cells of the immune system and force them to build new copies of the virus. Each molecule in the virus plays a role in this process, from the first steps of viral attachment to the final process of budding.

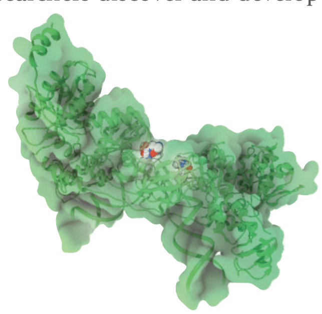
More than 25 years of research have revealed the atomic details of the structural biology of HIV, all publicly available in the Protein Data Bank (PDB) archive. As of this writing, the PDB contains more than 1,600 entries of HIV-related structures, bound and unbound to drugs. Using these data, researchers have designed new treatments for HIV infection, including effective drug regimens that halt the growth of the virus. PDB structures also provide new hope for development of a vaccine.

The RCSB PDB builds upon the data in the PDB archive to create resources that help explore the structural biology of the HIV virus. These include related *Molecule of the Month* articles, educational posters and animations, and specialized tools to help search, visualize, and learn about these protein-drug interactions.

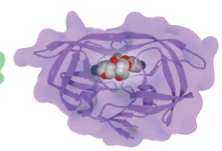
Highlighted on the cover is a representation of the HIV virus surrounded by examples of related protein-drug complexes. These include rilpivirne bound to HIV-1 reverse transcriptase; dolutegravir bound to the PFV virus, a protein similar to HIV integrase; and darunavir bound to HIV protease. Molecular snapshots like these help researchers discover and develop drugs for the fight against HIV.



Crystal structure of HIV-1 reverse transcriptase (RT) in complex with Rilpivirine (TMC278, Edurant), a non-nucleoside rt-inhibiting drug. PDB ID: 4g1q[60]



Crystal structure of the Prototype Foamy Virus (PFV) intasome in complex with magnesium and Dolutegravir (S/GSK1349572) PDB ID: 3s3m[61]



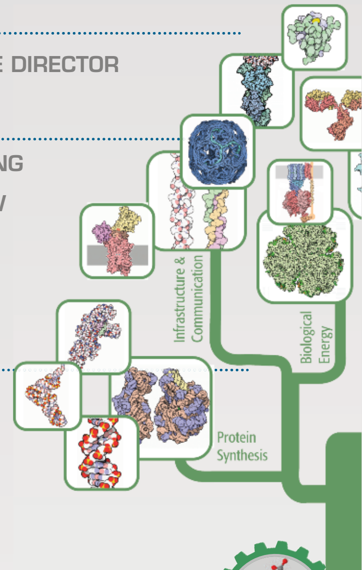
Crystal structure of wild type HIV-1 protease in complex with darunavir. PDB ID: 4hla[62]

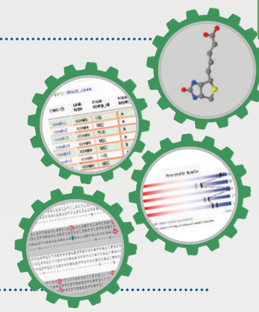
Video Challenge for High School Students

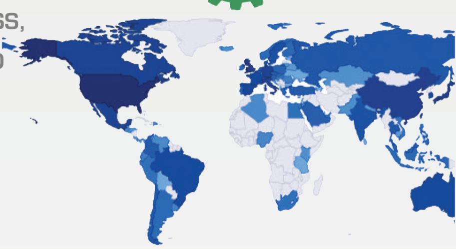
In 2014, the RCSB PDB will host a Video Challenge for High School Students focused on HIV/AIDS.

FOR MORE INFORMATION, VISIT PDB-101 (**RCSB.ORG/PDB-101**).

CONTENTS







A new sculpture sits outside of the RCSB PDB offices at Rutgers, The State University of New Jersey. Julian Voss-Andreae's 20-foot, 3,200-pound polished stainless steel and colored glass sculpture *Synergy* is based on the triple-helical structure of collagen.

Collagen is particularly interesting. It is the most abundant protein in the body, and provides core infrastructure. Collagen's 3D structure is very simple and elegant: three strands, with a repeating sequence of a few amino acids. Subtle changes in this sequence are central to collagen's complex interactions with other molecules. In this sense, collagen symbolizes the ability to understand complexity through simplicity.

Similarly, the RCSB PDB aims to enable understanding of complex biological systems from biomedicine to agriculture by providing elegant resources and services for depositing, searching, and analyzing PDB data.

Many of these tools are developed using the core infrastructure provided by the PDBx/Macromolecule Crystallographic Information File (mmCIF) data dictionaries and format that describe the information content of PDB entries. The flexibility of the PDBx/mmCIF format allows the PDB archive to easily support different types of data, including a variety of experimental formats, very complicated structures, and extremely large complexes.

RCSB PDB resources are also built upon the synergies found in our collaborations with other institutions. A key collaboration involves our partners in the Worldwide Protein Data Bank (wwPDB), the organization responsible for the oversight and management of the PDB archive.

The development of the new Common Deposition and Annotation System, currently in testing, has utilized the knowledge and experience of the annotators and software developers of the wwPDB member data centers to produce an improved system that will support data deposition and annotation for the next decade.

Internal RCSB PDB synergies exist between our host institutions of Rutgers and the University of California, San Diego (UCSD). While Rutgers focuses primarily on data deposition and annotation ("Data In"), the UCSD group is similarly concerned with data access, query, and reporting ("Data Out"). At the same time, research and development projects are designed so that one process informs the other. For example, the release of the wwPDB's Biologically Interesting Molecule Reference Dictionary (BIRD), which contains information about the representation of peptide-like antibiotic and inhibitor molecules in the PDB archive, drove new RCSB PDB website features that utilize BIRD data to provide improved searching and visualization options for these molecules. And our education and outreach efforts are both informed by our experiences in "Data In" and "Data Out." In addition, the feedback we receive as we communicate new tools and resources to users drives important evolution of future improvements and features.

These synergies will continue to accelerate RCSB PDB development as we continue to promote a structural view of biology.

Helen M. Berman
Director, RCSB PDB
Board of Governors Professor of Chemistry and Chemical Biology
Rutgers, The State University of New Jersey

Enabling Research, Education, and Innovation Worldwide

The RCSB PDB (**rcsb.org**) is a vital resource for biological research and education worldwide.[1] It provides enhanced access to information about the 3D structures of nucleic acids, proteins, and large molecular machines contained in the Protein Data Bank (PDB) archive.[2] The RCSB PDB supports the development of standards for the representation, annotation, and validation of these structural data that are collected from different experimental methods. An online educational portal is enhanced by online and in-person outreach efforts targeted at promoting a structural view of biology.[3]

RCSB PDB resources are utilized by a variety of researchers, teachers, and students studying biology and its connections to molecular biology, structural biology, computational biology, pharmacology, and more.

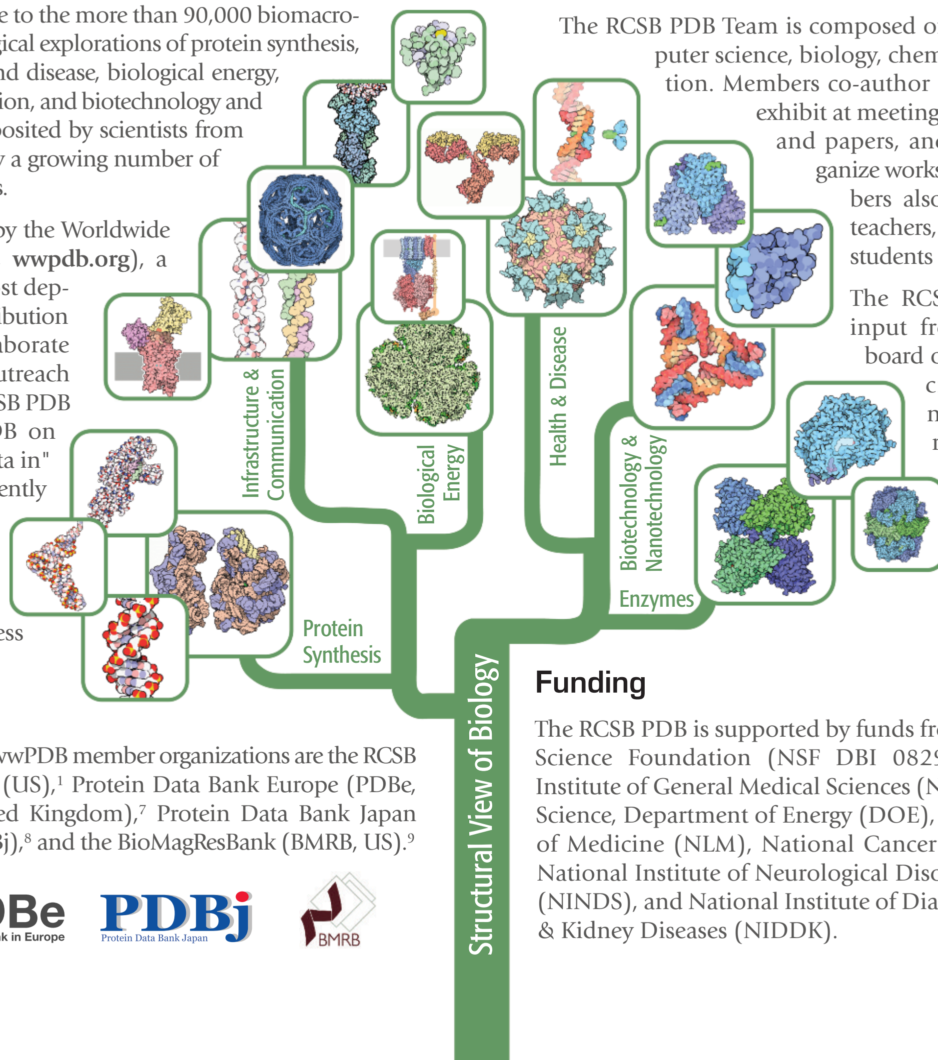
This annual report highlights recent advances made by the RCSB PDB.

Growth of the PDB Archive

The field of structural biology began in the late 1950s as scientists began to decipher the 3D shapes of proteins at an atomic level. The scientific community banded together as they envisioned how strong research advances could be made through a shared, public archive of data from these experiments.[4] In 1971, the PDB archive was established with seven structures.[5] Scientists would submit their coordinate data to the PDB, who would then mail them by post to interested users.

Today, online access is available to the more than 90,000 biomacromolecules that promote biological explorations of protein synthesis, enzymatic processes, health and disease, biological energy, infrastructure and communication, and biotechnology and nanotechnology. Data are deposited by scientists from around the world, and used by a growing number of researchers in a variety of fields.

The PDB archive is managed by the Worldwide Protein Data Bank (wwPDB, **wwpdb.org**), a consortium of groups that host deposition, annotation, and distribution centers for PDB data and collaborate on a variety of projects and outreach efforts.[6] As a member, the RCSB PDB collaborates with the wwPDB on matters relating to getting "data in" to the archive, while independently developing tools and resources to get "data out" in the form of the RCSB PDB website, tools, and other resources that enable the access and analysis of PDB data.



WORLDWIDE PDB PROTEIN DATA BANK

The wwPDB member organizations are the RCSB PDB (US),[1] Protein Data Bank Europe (PDBe, United Kingdom),[7] Protein Data Bank Japan (PDBj),[8] and the BioMagResBank (BMRB, US).[9]

RCSB PDB PROTEIN DATA BANK · PDBe Protein Data Bank in Europe · PDBj Protein Data Bank Japan · BMRB

RCSB PDB Organization

The RCSB PDB is jointly managed at Rutgers, The State University of New Jersey and the University of California, San Diego (San Diego Supercomputer Center and the Skaggs School of Pharmacy and Pharmaceutical Sciences).

RCSB PDB Director Helen M. Berman was part of the team that first envisioned the PDB archive. She is a Board of Governors Professor of Chemistry and Chemical Biology at Rutgers. Professor Stephen K. Burley, Associate RCSB PDB Director and Director of Rutgers' Center for Integrative Proteomics Research joined the management team, which also includes Dr. Martha Quesada, Deputy Director (Rutgers) and Professor Philip E. Bourne, Associate Director (UCSD).

2013 Advisory Board Meeting

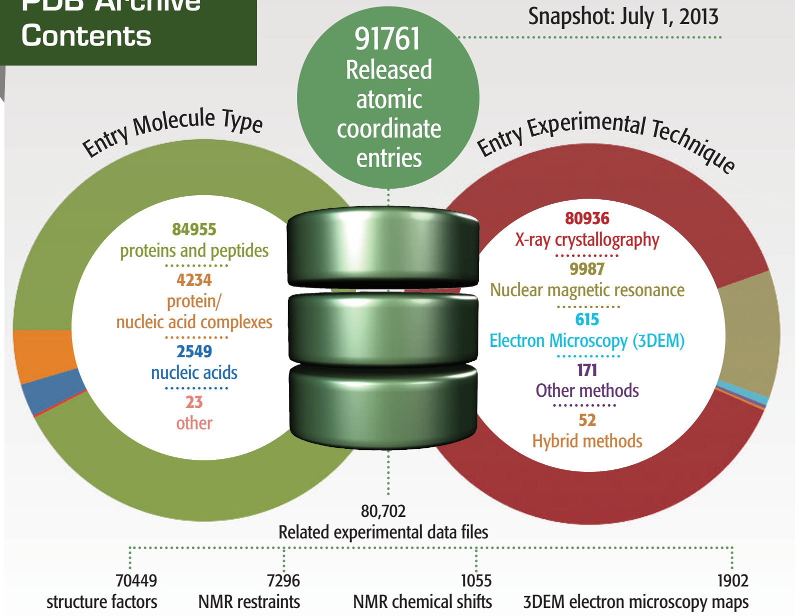
The RCSB PDB Team is composed of experts in computer science, biology, chemistry, and education. Members co-author scientific papers, exhibit at meetings, present posters and papers, and attend and organize workshops. Staff members also serve as tutors, teachers, and mentors to students of all ages.

The RCSB PDB receives input from an advisory board of experts in X-ray crystallography, nuclear magnetic resonance (NMR), 3D electron microscopy (3DEM), bioinformatics, and education.

Funding

The RCSB PDB is supported by funds from the: National Science Foundation (NSF DBI 0829586), National Institute of General Medical Sciences (NIGMS), Office of Science, Department of Energy (DOE), National Library of Medicine (NLM), National Cancer Institute (NCI), National Institute of Neurological Disorders and Stroke (NINDS), and National Institute of Diabetes & Digestive & Kidney Diseases (NIDDK).

PDB Archive Contents



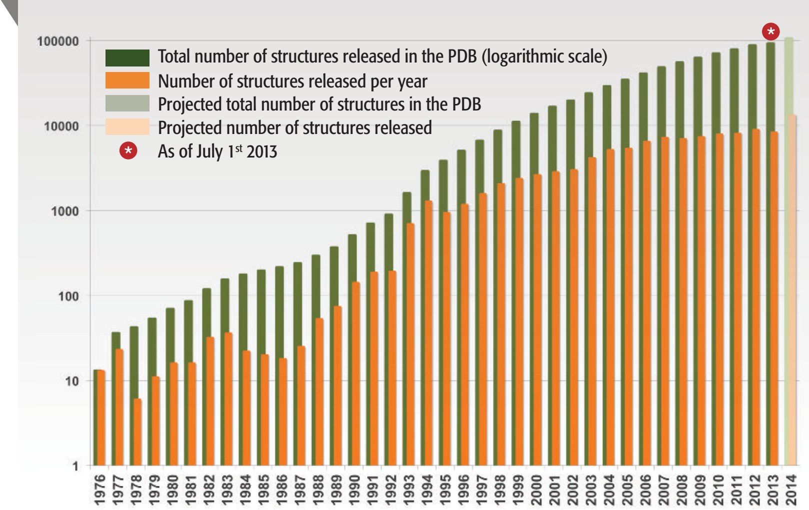
100,000 Structures in 2014: International Year of Crystallography

The United Nations General Assembly has proclaimed 2014 as the International Year of Crystallography (IYCr; **iycr2014.org**) in celebration of two important events that spawned the scientific field of X-ray crystallography: the 100th anniversary of the determination of the first atomic-resolution structure and the 400th anniversary of Kepler's fundamental observation that snow crystals formed symmetric shapes. IYCr is co-sponsored by the International Union of Crystallography (IUCr) and the United Nations Educational, Scientific and Cultural Organization (UNESCO).

As the majority of structures in the PDB archive are the result of X-ray crystallographic experiments, we are pleased to note the archive is on track to reach a milestone of 100,000 released entries during IYCr. We anticipate that the synergy of these events will introduce the power of structural biology to new communities.

In addition to the growing number of released entries, the PDB is experiencing a relentless rise in the complexity of structures determined. The past several years have seen marked increases in the number of ligands released, the size of the average polymeric molecular weights, and the number of chains per asymmetric unit.

Yearly Growth of Structures Released in the PDB archive



Data Pipeline

Structure Determination & Deposition

SCIENTISTS experimentally determine 3D structures of nucleic acids, proteins, and molecular machines, and submit data describing these structures to the PDB archive using specialized online tools.

10,368 structures were deposited during this report period

Data Annotation

wwPDB ANNOTATORS carefully curate and review each entry. Important validation checks are run and reported back to the depositor.

On average, structures are processed, validated, and fully prepared for release within two weeks.

Data Archiving

The PDB ARCHIVE is updated weekly. Time-stamped snapshots are also archived to provide readily identifiable data sets for research.

Structures released during this report period:

9,252	coordinate files
8,671	structure factor files
507	constraint data files
500	chemical shifts
494	EM maps

Data Access

DATA can be accessed via wwPDB-hosted FTP and websites.

Data were downloaded 365 million times from wwPDB FTP servers in 2012.

Overview

Biocurators annotate each PDB entry to ensure accurate representation of both the structure and experiment. The annotation team reviews polymer sequences, small molecule chemistry, cross references to other databases, experimental details, correspondence of coordinates with primary data, protein conformation (Ramachandran plot), biological assemblies, and crystal packing. Annotators communicate with depositors to make sure the data are represented in the best way possible.

Depending on the hold status selected by the depositor, data release occurs when a depositor approves the annotated entry (status: release immediately/REL), the hold date has expired (HOLD), or the corresponding journal article has been published (hold for publication/HPUB). Data files describing each entry are released in different file formats (PDB, PDBx, PDBML-XML)[10,11] along with related experimental data.

10,368 entries were deposited to the PDB archive and prepared for release by the wwPDB during the period of this report. 7,850 of these entries were submitted to the RCSB PDB. On average, structures are processed, reviewed by the author, and finalized for release in two weeks.

Sequences for 40% of these depositions were made publicly available prior to the release of the coordinate entry. This helps prevent duplication of structure determination efforts and promotes blind testing of structure prediction and modeling techniques.

In June 2013, the number of structures determined by NMR available in the PDB archive passed the 10,000 mark.

Shown: PDB ID 2m8k. Pyrimidine motif triple helix in the Kluyveromyces lactis *telomerase RNA pseudoknot is essential for function in vivo (2013) D. D. Cash, O. Cohen-Zontag, N. K. Kim,* et al. Proc Natl Acad Sci USA ***110****: 10970-10975.*

Data Dictionaries and PDBx/mmCIF

PDB entries are curated and annotated using standard PDBx/macromolecular Crystallographic Information File (mmCIF)[12] data dictionaries for the macromolecular structures and small chemical components found in PDB entries.

PDBx/mmCIF

The PDB Exchange (PDBx) data dictionary[13] consolidates content from a variety of crystallographic data dictionaries and includes extensions describing NMR, 3DEM, and protein production data. Internal data processing, annotation, and database management operations rely on the PDBx/mmCIF dictionary content and corresponding file format.

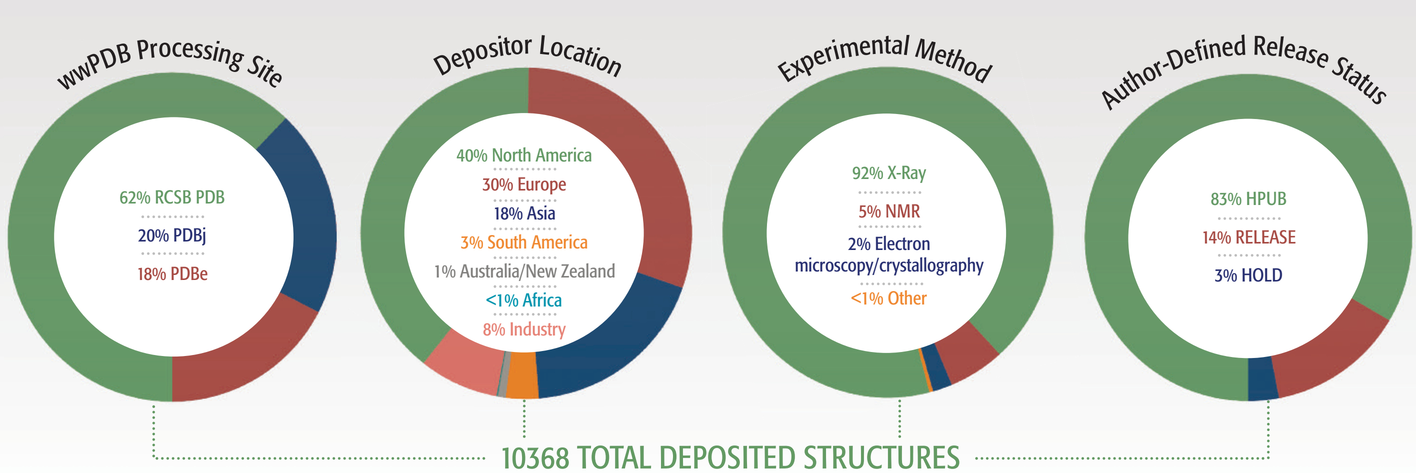
As the PDBx/mmCIF file format is very extensible, it can expand and grow to support new types of information. It is also able to accommodate large and complex entries in a single file. The original "PDB file format" has size limitations that require the large molecular machines to be "split" among several different entries.

In 2011, the wwPDB established the PDBx/mmCIF Working Group of software developers to enable direct use of PDBx/mmCIF in the major macromolecular crystallographic refinement and analysis tools. PDBx/mmCIF files suitable for deposition can now be created with recent versions of the CCP4[14] (REFMAC[15] 5.8) and Phenix[16] (1.8.2) software packages. The Working Group also made recommendations about essential extensions required for large structures that have been incorporated. These extensions have enabled the release of several large structures as single files.

With the support of the PDBx/mmCIF Working Group, the wwPDB announced that the PDBx/mmCIF file format will become the working format for PDB structure data. For the first time in the history of the PDB archive, structures will be deposited, annotated, and released in a single format. To ease this transition, the wwPDB will continue to work with refinement and visualization software developers to ensure that their software will both output and accept PDBx file formats.

Details about the PDBx/mmCIF file format, example files, and related documentation are available at **wwpdb.org**.

Deposition Statistics, July 1, 2012 – June 30, 2013



Small Molecule Dictionaries

All residue and small molecule components found in PDB entries, including standard and modified amino acids/nucleotides, small organic molecules, and solvent molecules, are described in the wwPDB's Chemical Component Dictionary. Each chemical definition describes properties such as stereochemistry, aromaticity, idealized coordinates representing the molecular connectivity, bond order, and chirality that is energetically most favorable for the component, chemical descriptors (SMILES & InChI), and systematic chemical names. This dictionary grows as new components are deposited as part of new PDB entries. Of the ~17,000 small molecules currently available in the dictionary, more than 1,800 were released during this report period.

The RCSB PDB's **Ligand Expo (ligand-expo.rcsb.org)** provides tools for accessing, visualizing, and viewing reports about the information in the Chemical Component Dictionary.[17] This tool is used internally, and by depositors preparing data submissions. Many of the dictionary search and reporting features available in Ligand Expo have been integrated with the other query capabilities available at **rcsb.org**.

The **Biologically Interesting molecule Reference Dictionary (BIRD)**[18] supports the representation of peptide-like antibiotic and inhibitor molecules. This dictionary contains chemical and functional information for more than 650 small polymers and is updated on the PDB FTP site with each weekly release.

Data Reviews and Remediation

The wwPDB regularly reviews the archive to correct errors and inconsistencies. These remediation efforts involve the creation of new versions of the data files, and are documented in detail at **wwpdb.org**. Past remediation efforts have improved the representation of sequences, ligand chemistry and nomenclature, biological assemblies, residual B-factors, peptide inhibitors and antibiotics, and entries in nonstandard crystal frames, and other improvements.[17,19,20] Since the 2011 remediation effort, detailed descriptions of any changes made to data are recorded in the data entry.

In this report period, X-ray crystallographic structure factor data files in the PDB archive have been standardized, with corrections made to older entries. 43,800 structure factor files have been updated and re-released in 2013.

A multi-year review, analysis, and remediation effort is underway to address the representation of carbohydrates, protein modifications, and metal-containing ligands in the archive.

A New System for Deposition and Annotation

The number of depositions to the PDB archive continues to rise. Structures are also increasing in size and complexity. The members of the wwPDB have collaborated on an important project to ensure that the wwPDB can effectively support these increases while continuing to consistently provide high quality data. The result of this partnership is the next generation of PDB deposition and annotation tools.[21]

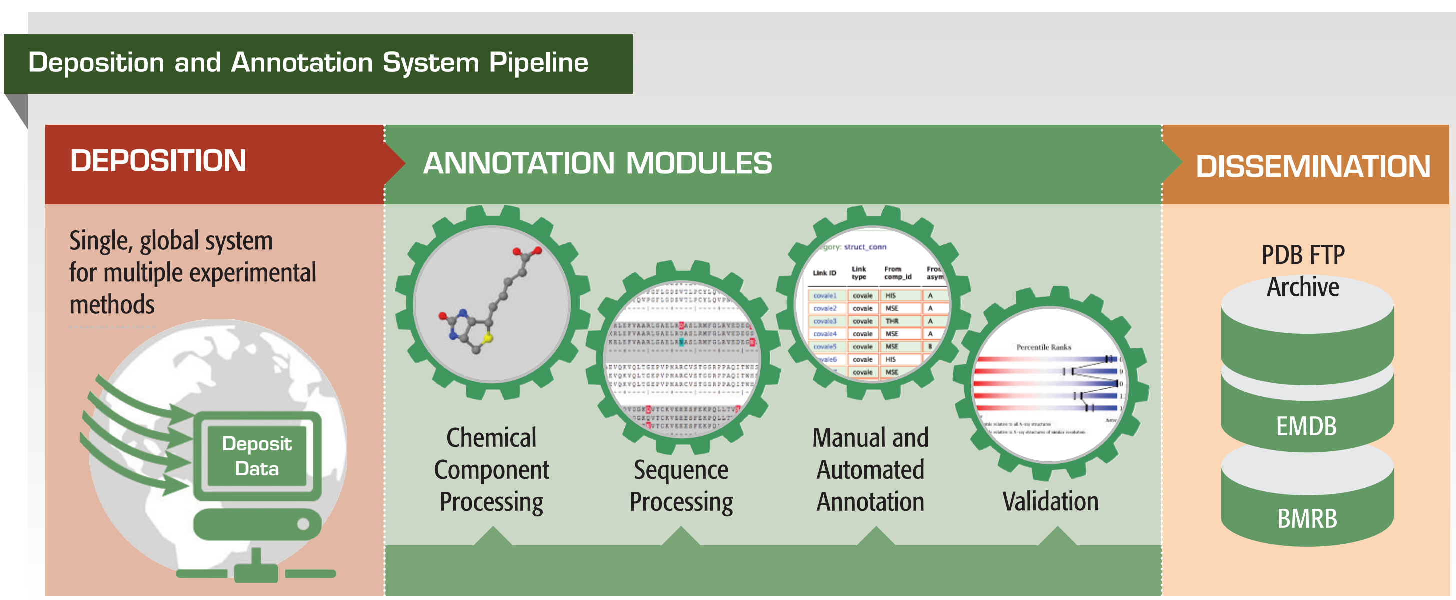
The wwPDB Common Deposition & Annotation (D&A) system provides uniformity across all wwPDB sites for the first time. Currently in community testing, the new and improved system will go into full production in early 2014.

Deposition Improvements

The new interface supports structures of any size that have been determined using X-ray diffraction, 3DEM, and/or NMR methods. Automated upload and extraction of PDBx formatted files and the availability of pdb_extract for PDB files has significantly minimized the need for manual entry. New depositions can also be based on existing submissions, saving time and effort.

The HIV capsid was the first PDB entry deposited, annotated, and released entirely in the PDBx/ mmCIF format.

PDB IDs 3j3q (shown) and 3j3y. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics (2013) G. Zhao, J. R. Perilla, E. L. Yufenyuy, et al. Nature ***497****: 643-646.*



As some depositors make improvements to their entries based upon feedback received during the deposition and annotation process, the new system supports easy coordinate and experimental data file replacement at any point before entry release. Data files can be previewed and downloaded before and after deposition.

The deposition system provides community-defined validation results prior to submission, and includes improved checking for ligand chemistry and polymer sequence consistency.

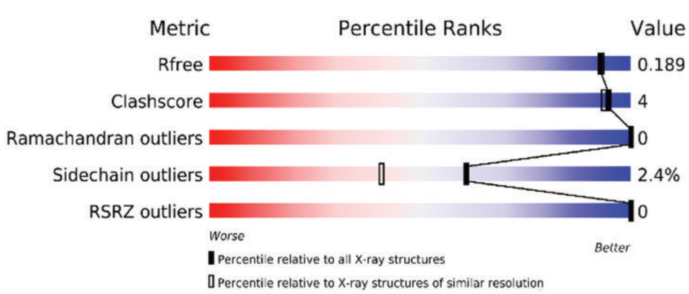
IDs for PDB, EMDB, and BMRB entries are assigned in a single deposition session.

Annotation Improvements

The new annotation system is in use at the wwPDB member sites as a parallel production system during the last phase of system testing. The D&A system pipeline involves four major modules designed to improve annotation quality and productivity:

- **The Chemical Component Annotation Tool**[22] is the module used for ligand processing. In a major improvement to previous manual systems, this tool automatically compares deposited ligands to wwPDB small molecule reference dictionaries. Possible matches are then displayed in 2D and 3D views for visual chemical comparison and ligand ID assignment.
- **The Sequence Processing Module** compares sequences as entered in the deposition, as captured in the coordinates of the entry, and as represented in cross-references in third-party sequence databases.[23,24] It also supports multiple sequence alignments and interactive annotation of discrepancies.
- **The Manual and Automated Annotation Module** derives data for secondary structure, binding sites, linkages and predicted biological assemblies. 3D visualization is available for review and editing as needed.
- **The Validation Module**[25] produces validation reports that incorporate recommendations of method-specific community Task Forces. A PDF summarizing model quality, fit of the model to data, and residue geometry is created for depositor review and journal submission. A visual "slider"image provides an at-a-glance summary of key statistics for overall quality. The report also tabulates diagnostics for the entry's macromolecule and ligands, and summarizes key data and refinement statistics.

Other system improvements include improved status tracking and file versioning, which will greatly facilitate future remediation efforts.



The visual "slider" graphic provides a visual summary of structure quality. In this example, the boxes show assessment of the entry relative to percentile scores (ranging between 0-100) for all X-ray structures (solid box) and all X-ray structures of similar resolution (box outline).[25, 26]

Following an extensive testing period, the new annotation system is now in use at the wwPDB data centers.

Validation Reports and Community Task Forces

To help ensure the accuracy of entries in the PDB, deposited data are compared with community-accepted standards during the process of validation. As part of annotation, wwPDB members provide depositors with detailed reports that include the results of geometric and experimental data validation.[25] These reports, available as PDFs, provide an assessment of structure quality while maintaining the confidentiality of the coordinate data.

wwPDB validation reports are required by several journals for manuscript review, including *eLife*, *The Journal of Biological Chemistry*, and the journals of the International Union of Crystallography. The wwPDB encourages all journal editors and referees to incorporate these reports in the manuscript submission and review process.

To improve validation methods, method-specific Validation Task Forces (VTF) in X-ray Crystallography,[26] NMR,[27] 3DEM,[28] and Small Angle Scattering[29] have been convened to collect recommendations and develop consensus on additional validation that should be performed, and to identify software applications to perform validation tasks. The X-ray VTF recommendations have been incorporated into validation reporting. All recommendations will be incorporated into the wwPDB data processing procedures and tools as part of the wwPDB Common Deposition & Annotation Tool development.

New Task Force Reports

NMR VALIDATION TASK FORCE REPORT

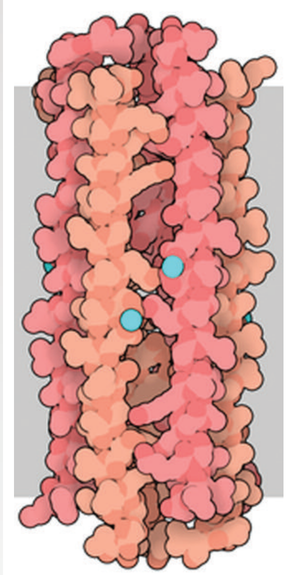
Recommendations of the wwPDB NMR Validation Task Force

Gaetano T. Montelione, Michael Nilges, Ad Bax, Peter Güntert, Torsten Herrmann, Jane S. Richardson, Charles D. Schwieters, Wim F. Vranken, Geerten W. Vuister, David S. Wishart, Helen M. Berman, Gerard J. Kleywegt, John L. Markley *Structure* (2013) **21**: 1563-1570
doi: 10.1016/j.str.2013.07.021

SMALL ANGLE SCATTERING TASK FORCE REPORT

Report of the wwPDB Small-Angle Scattering Task Force: Data Requirements for Biomolecular Modeling and the PDB Structure

Jill Trewhella, Wayne A. Hendrickson, Gerard J. Kleywegt, Andrej Sali, Mamoru Sato, Torsten Schwede, Dmitri I. Svergun, John A. Tainer, John Westbrook and Helen M. Berman (2013) *Structure* **21**: 875-881
doi:10.1016/j.str.2013.04.020



Dermcidin: June 2013 Molecule of the Month

PDB ID 2ymk. C. Song, C. S.Weichbrodt, E Salnikov, et al. (2013) Crystal structure and functional mechanism of a human antimicrobial membrane channel. Proc Natl Acad Sci USA *110: 4586- 4591.*

PDB Archiving and Distribution

RCSB PDB services and PDB data are freely available online. As the archive keeper for the wwPDB, the RCSB PDB maintains the PDB archive at **ftp://ftp.wwpdb.org**. Weekly updates of the PDB archive and the RCSB PDB website are coordinated with the other wwPDB sites.

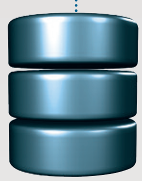
A total of 9252 coordinate files, 8671 structure factor files, 507 constraint data files, 500 chemical shifts, and 494 EM maps were released during this period.

Monthly activity at **ftp://ftp.wwpdb.org** via FTP and rsync protocols typically amounts to about 18.7 million downloads by about 12,600 unique users for a total of approximately 4.1 Terabytes of data. PDB data are also downloaded from the wwPDB member FTP sites and individual web portals at PDBe and PDBj. In 2012, data from the PDB archive were accessed 365 million times from wwPDB FTP and websites–for an average of 1 million a day.

Online RCSB PDB Access

RCSB PDB RESOURCES ARE AVAILABLE VIA

Website (rcsb.org)



Separate FTP server for derived data

Web Services

Mobile devices[3]

The website is accessed by

~286,000 UNIQUE VISITORS PER MONTH

(up 18% from last year's 241,000) from about

190 COUNTRIES.

Approximately

1.3 TERABYTES

of data are transferred each month from the website.

The RCSB PDB website offers tools for query, reporting, and visualization that are integrated with a database that contains data from the PDB archive, data and links from external resources, and pre-calculated data. New features and enhancements are continually being developed, with major new features first introduced through the public beta test site (**betastaging.rcsb.org**). Improvements that support the internal infrastructure and automate tasks are made regularly to improve performance and reliability.

THE RCSB PDB UNIVERSAL APP FOR IOS

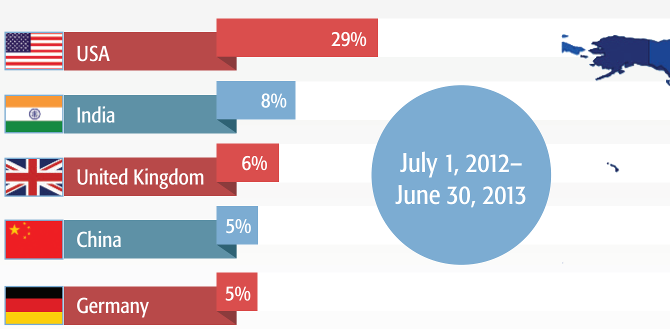
has been downloaded from the Apple App store more than

10,000 TIMES.

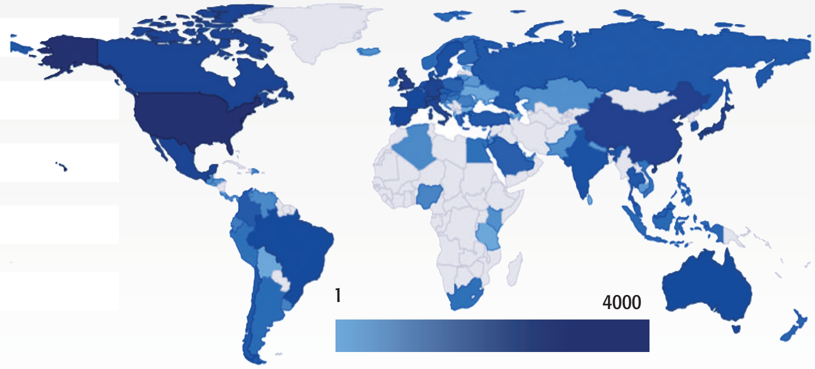
RCSB PDB *Mobile* enables users to search the entire PDB, view the latest weekly release of structures, access MyPDB accounts, view the entire catalog of *Molecule of the Month* articles, and more via WiFi or cellular data connection.

An Android version is in beta testing.

TOP FIVE COUNTRIES USING RCSB.ORG



RCSB PDB *MOBILE* DOWNLOADS BY COUNTRY



Search Tools

Quick Searching:

The top search bar on every web page facilitates easy, intuitive, and precise queries. Typing text in the top search bar launches an interactive pop-up box containing search suggestions organized by category. Support for new searches have been added, including use of synonyms, multi-component protein naming, short names, and gene names.

This search function can be used to quickly find a particular structure or set of structures, such as a search by PDB ID (4hhb), molecule name (collagen), or author (Hendrickson).

Advanced Searching:

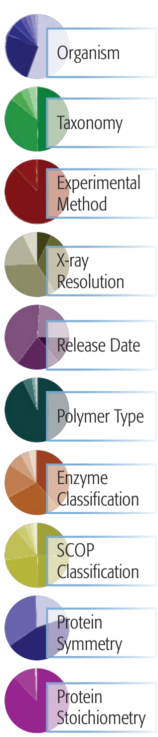
Multi-parameter searches on specific data items or annotations can be performed using the Advanced Search tool. A logical AND or OR can be used to return a list of structures that comply with ALL or ANY of the search criteria, respectively.

Advanced Search can be used to build very focused queries (for example, all structures containing N-terminal histidine tags or structures with resolution better than 2.0Å complexed with RNA).

Browse Database:

The PDB archive can be explored using different hierarchal trees organized around Gene Ontology (GO) Terms,[30] Enzyme Classification,[31] Source Organism, Genome Location, Medical Subject Headings (MeSH) terms,[32] protein classifications from SCOP[33] and CATH,[34] Anatomical Therapeutic Chemical (ATC) Classification System,[35] Transporter Classification,[36] and Protein Symmetry.

Browse Database options can find structures using a search constraint specific to a given hierarchy (such as cell chemotaxis (GO ID:60326)) or more open-ended queries (*what structures are related to MeSH term Plant Roots?*).



Drill-down Pie Charts:

Standard characteristics of PDB entries–resolution, release date, polymer type, organism, and more–are used to create searchable data distribution summaries. Users can tour the PDB archive by drilling down through combinations of these significant properties either directly from the home page (starting with all entries in the archive) or from a query results set (to further refine a search).

Pie charts can help answer general questions (What is the distribution of resolution in the archive?) or used to build more complicated queries based on the results found (homomeric structures with cyclic symmetry from a human source).

Web Services.

RESTful Web Services offer programmatic access to PDB data including a generic option to execute queries similar to those available on the website; custom report services that return data in CSV, Excel, and XML format; and several specialized services for retrieval of information about the status and description of PDB entries, chemical components, and sequence-based mappings.[37]

MyPDB:

This online service provides personalized access to PDB data. It can store any type of RCSB PDB structure search, including a particular keyword, sequence, or Advanced Search composite query, to be run at any time. MyPDB can be customized to provide email alerts when new entries matching a saved search are released in the PDB archive. Personal annotations and notes can be saved on any entry's Structure Summary page, together with a bookmark list of favorite structures.

Search Results

The Query Results Browser lets users refine the set of structures returned by a search, access related *Molecule of the Month* features, and review individual entries. This browser can be modified to display the default view of query results that includes information about each entry, such as authors, compound, citation, classification, and residue count; a condensed view (title and macromolecule name); a gallery of images; and a timeline display of structure images.

Search results can be also be refined by editing the query using Advanced Search or by drilling down through the data distribution pie chart options. The final results set can be used to create a variety of tabular reports, download all sequences, or access all coordinate files.

For every entry in the PDB, an RCSB PDB Structure Summary page provides an overview of the structure; derived data from external resources; tools to examine the sequence, sequence domains, and sequence similarity; detailed information relating to the entry's citation, biology and chemistry, experiment, and geometry; and links to related resources. Several molecular viewers, including Jmol[38] and the RCSB PDB's Protein Workshop and Ligand Explorer,[39] help users view the molecule interactively.

Ligand Summary pages are available for each chemical component in the PDB to provide access to 2D and 3D visualization, subcomponent information, and information from DrugBank.[40] Ligand Summary Reports that include information about selected ligands (formula, molecular weight, name, SMILES string,[41] which PDB entries are related to the ligand, and how they are related) can be created for a search results set.

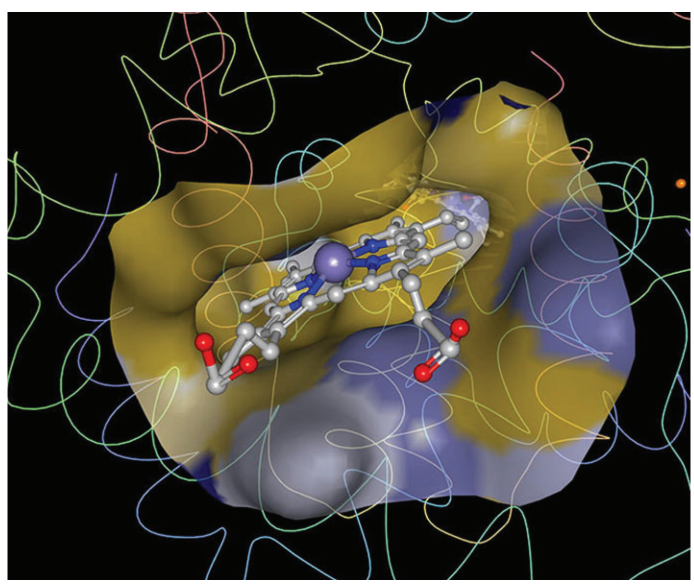
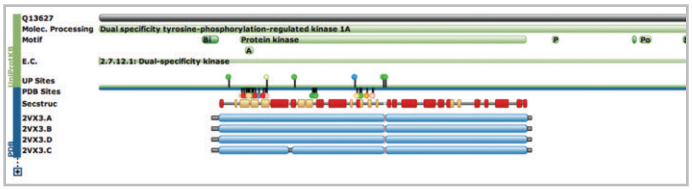


Image of a heme in hemoglobin created with Ligand Explorer.[45,46]

New Features

Protein Feature View:

This improved sequence view provides a mapping of protein chains from PDB entries onto the full-length sequences from UniProt.[23,42] All structures or structure fragments related to one full-length protein can be viewed. For example, for the HIV Gag-Pol polyprotein, the structures of different domains such as HIV protease, reverse transcriptase, and integrase provide a 'catalog' of all HIV related proteins and show the structural coverage of a sequence. This view is annotated with UniProt features, Pfam[43] and SCOP domains, protein secondary structure, and available homology models from the Protein Model Portal.[44]



Protein Feature View from the Structure Summary page for 2vx3.[63]

Drug View:

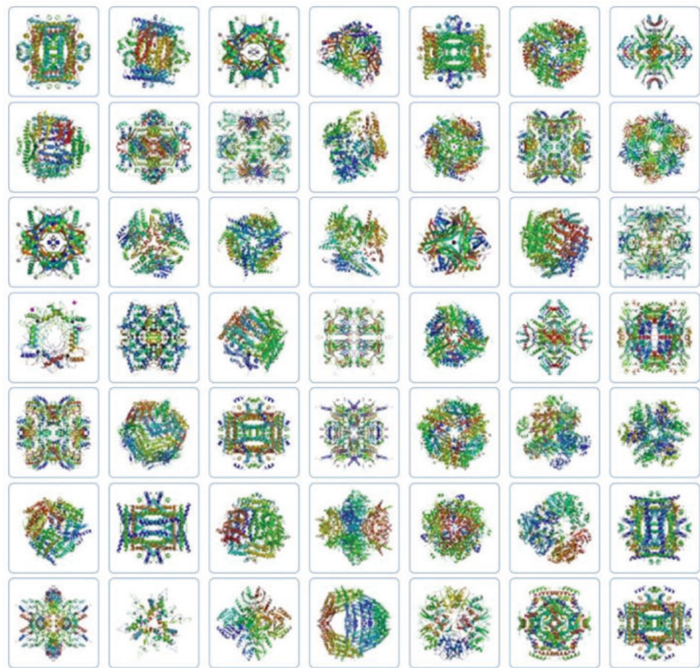
Drugs in DrugBank have been mapped to ligand structures in the PDB. Drugs are searchable by generic and brand names, and drugs can be browsed using the Anatomical Therapeutic Chemical (ATC) classification scheme by the World Health Organization.[35] Ligand Summary pages display drug annotations such as generic and brand names, mechanism of action, indication, and drug target information.

Support for the Biologically Interesting molecule Reference Dictionary (BIRD):

BIRD contains information about the representation of peptide-like antibiotic and inhibitor molecules in the PDB archive. BIRD molecules can be searched by name from the top search bar, links from Structure Summary pages, or queried by type and class through the Advanced Search. BIRD annotations and 2D chemical diagrams are displayed on Structure Summary pages. The three most highly used 3D viewers (Jmol, Ligand Explorer, Protein Workshop) have been modified to properly display these molecules as a single entity and to provide binding site visualization in Ligand Explorer.

Protein Stoichiometry and Symmetry:

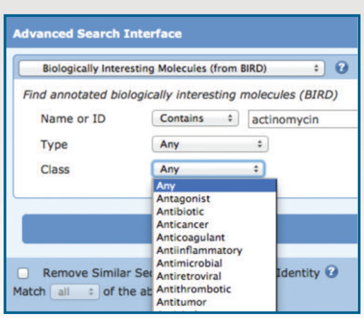
The stoichiometry and point group symmetry of protein complexes is now available as a drill-down on the home and query results pages and a constraint in the Advanced Search. These options help answer frequently asked questions, such as how to retrieve all tetramers with D2 symmetry. Special features have been added to the Jmol page to assist with the analysis of protein complexes, including the default alignment of protein complexes with their symmetry axes, visualization of symmetry axes, and a polyhedron color scheme that highlights the symmetry of a protein complex.



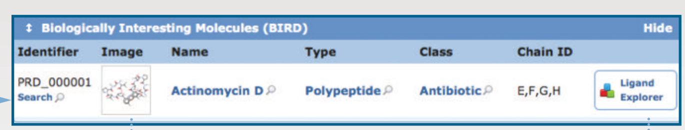
PDB structures with tetrahedral symmetry. Structures were selected using the data distribution pie charts for symmetry, and the image was generated using the "Gallery View" option for exploring search results.

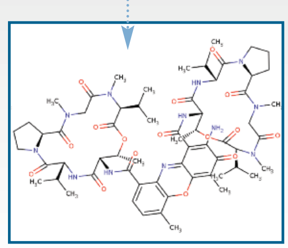
Search, display, and visualization of peptide-like molecules annotated in the Biologically Interesting molecule Reference Dictionary (BIRD)

BIRD molecules can be queried by TYPE and CLASS through the Advanced Search Interface

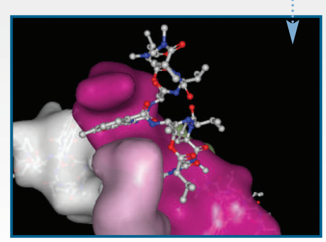


BIRD annotations are found on Structure Summary pages





2D chemical diagram



3D binding site visualization

Actinomycin D bound to DNA in PDB ID 1i3w.[46]

The RCSB PDB supports a variety of users. Our outreach efforts aim to inform different user communities about the RCSB PDB while collecting feedback to help develop a powerful resource for science, medicine, and education.[47,48]

Users include biologists from a variety of specialties, scientists from other disciplines, students and educators at all levels, authors and illustrators, and the general public. An 2012 survey of our website users gave some additional insight, as highlighted below.

While the website serves as the primary tool for outreach, staff interact directly with users at local, national, and international meetings, workshops, presentations, festivals, and more. Prizes are awarded for student posters at selected meetings.

Electronic help desks, newsletters, and flyers are also used to communicate with users and solicit detailed feedback about the resource.

Journal articles covering a diverse array of subjects are published regularly. Recent articles have described *Creating a community resource for protein science*,[49] *Trendspotting in the PDB*,[50] *Illustrating the machinery of life: Viruses*,[51] *Putting proteins in context*,[52] *BioJava: an open-source framework for bioinformatics*,[53] and *New resources for research and education*.[3]

Questions?
Email info@rcsb.org

RCSB PDB's help desk manages around 1000 conversations each year with users from all over the world. One query could be from a teacher searching for classroom materials, and the next a user wondering how to construct a specific search for finding proteins found in algae.

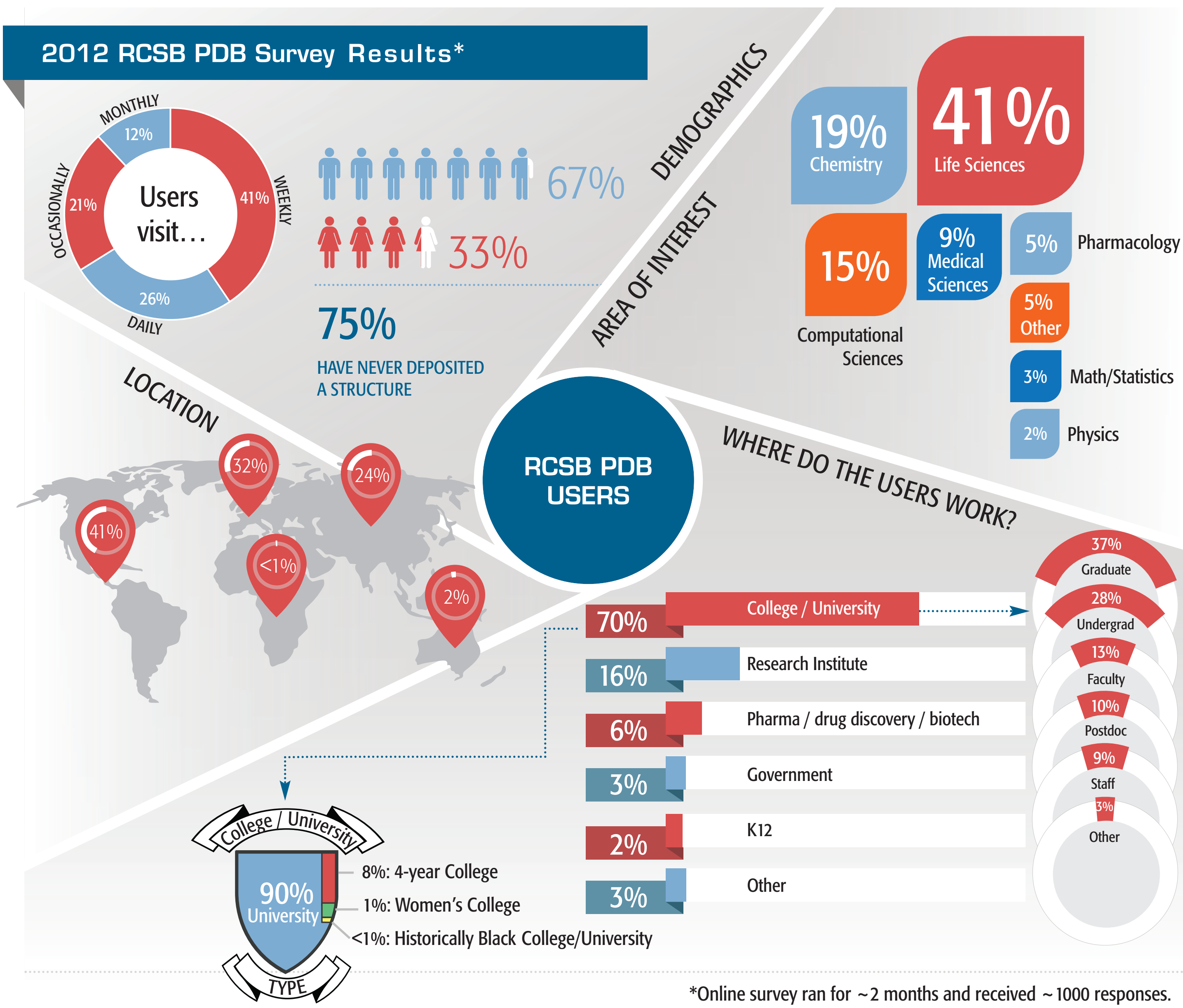
All questions–including homework help requests–are loaded into a tracking system for monitoring. Many of these help desk conversations lead to enhancements made to the website.

Social Media

FOLLOW and LIKE the RCSB PDB

@BUILDMODELS
twitter.com/buildmodels

RCSB PROTEIN DATA BANK
www.facebook.com/RCSBPDB

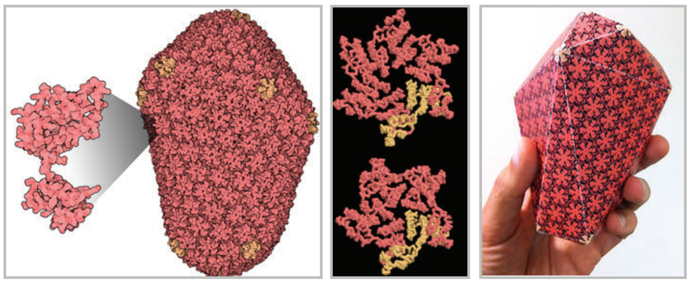


*Online survey ran for ~2 months and received ~1000 responses.

PDB-101 and Other Educational Activities

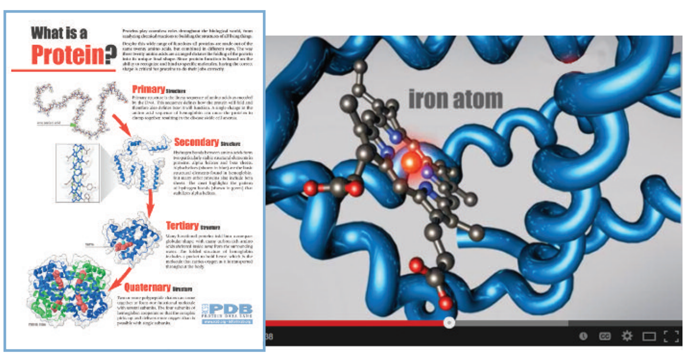
PDB-101 is an educational view of the RCSB PDB that packages together resources that promote exploration in the world of proteins and nucleic acids for teachers, students, and the general public.

Molecule of the Month: In this regular column by David S. Goodsell (The Scripps Research Institute), important biological molecules and how they function are illustrated through descriptive text, pictures, links to specific PDB entries, discussion questions, and 3D interactive molecular views.



The July 2013 Molecule of the Month *highlighted the HIV capsid structure using Jmol views and a downloadable template that can be used to create a paper model.*

Structural View of Biology offers a browser for top-down exploration of the PDB that starts with high-level functional categories and drills down to *Molecule of the Month* articles and related structural examples.



A new flyer and animation help answer the question What is a Protein?

Related **Educational Resources** available include posters, videos, downloadable paper models of proteins and nucleic acids, and other classroom materials.

Understanding PDB Data is an online reference to help explore and interpret individual PDB entries.

Author Profiles offer historical and educational timelines of the structures associated with a particular researcher.

New downloadable flyers include *What is a Protein?*, which gives an introduction to protein structure and function for beginners, and *The Structures of the Citric Acid Cycle*, which illustrates the PDB structures involved in this metabolic pathway. Two DNA-related resources have been translated into Spanish: the *Molecule of the Month* article and the paper model template that can be used to build a 3D model.

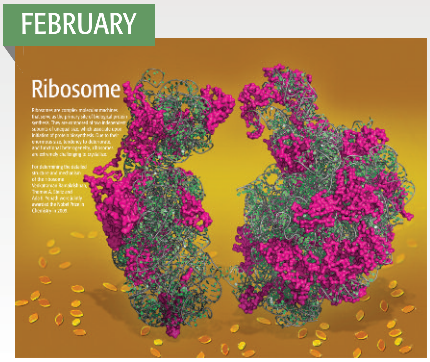
On campus at Rutgers and UCSD, RCSB PDB leaders teach graduate and undergraduate students how to understand and visualize PDB data in the context of biology. Other programs focus on working with students and teachers in middle and high school, such as the pilot program *Working Together to Visualize* that trained New Jersey high school teachers through workshops focusing on a structural view of biology. The teachers then incorporated the workshop materials in their classrooms. Students in these classes submitted short reports to the RCSB PDB. Top scoring students were invited to participate in the RCSB PDB's 2013 summer internship program.

RCSB PDB participates in local and national events such as Rutgers Day and the annual meeting of the National Science Teachers Association.

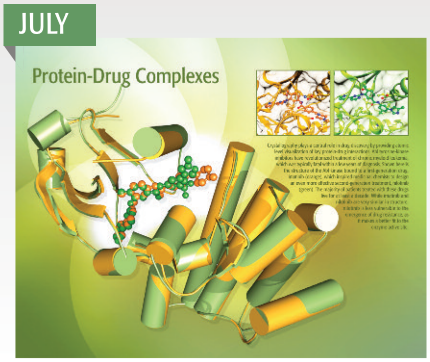
IYCr: 2014

For 2014: International Year of Crystallography, the wwPDB produced a calendar illustrating how X-ray crystallography enables our understanding of biology at the atomic level. The calendar, related images, and text can be downloaded from **wwpdb.org**.

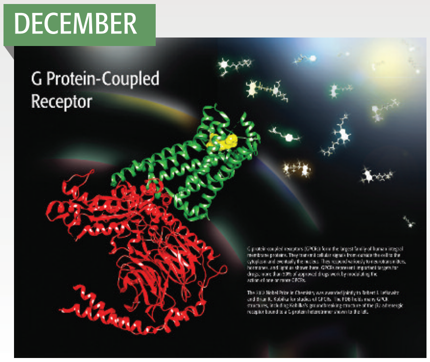
FEBRUARY



JULY



DECEMBER



High-throughput Structural Studies and the PSI Structural Biology Knowledgebase

PSI | nature StructuralBiologyKnowledgebase

Worldwide structural genomics centers have determined nearly 12,500 structures to date, which represents 13% of the PDB archive. RCSB PDB works closely with many of these centers and with the Protein Structure Initiative (PSI) Structural Biology Knowledgebase (SBKB).

The SBKB (**sbkb.org**) was established as a "marketplace of ideas and data" to facilitate research design and analysis for a wide variety of biological systems.[54-56] It serves as a single resource that integrates structure, sequence, and functional annotations, along with technical information regarding protein production and structure determination. Researchers can search the SBKB by sequence, PDB ID or UniProt accession code, and receive a current list of matching PDB entries, pre-built theoretical models from the Protein Model Portal,[44] annotations from 100+ open biological resources, structural genomics target histories and protocols from TargetTrack, and ready-to-use DNA clones from DNASU.[57] It is also possible to find structures according to functional relevance (KB-Rank tool),[58] or find related technologies and publications from the PSI Technology and Publications Portals, respectively. Interactive tools such as real-time theoretical modeling and biophysical parameter prediction also enhance understanding of proteins that are not yet well characterized.

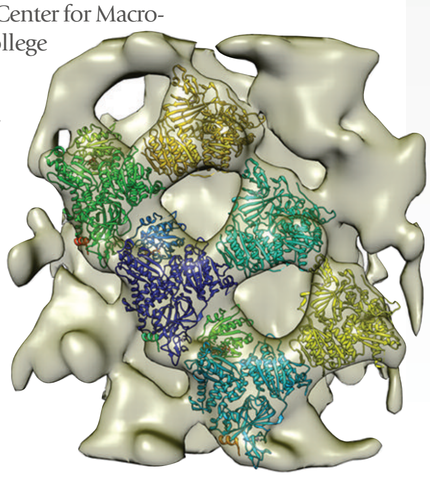
Monthly articles and editorials highlight the impact of structural biology on specific areas of biological research. Experimentally-focused website "hubs" offer up-to-date information about Structural Targets; Structure, Sequence and Function; Homology Models, Methods and Technologies, and Membrane Proteins. SBKB can be easily accessed from smartphones and tablets to enable on-the-go research.

EMDataBank: Unified Data Resource for 3-Dimensional Electron Microscopy

EMDataBank Unified Data Resource for 3DEM

The PDB archives large biological assemblies determined by 3DEM, a maturing methodology in structural biology that bridges the gap between cell biology and the experimental techniques of X-ray crystallography and NMR. 3DEM experiments produce 3D density maps, archived in the EM Data Bank, and often yield fitted coordinate models, which are archived in the PDB. **EMDataBank.org** is a deposition and retrieval network for 3DEM map, model, and associated metadata.[59]

The EMDB map archive is distributed as part of the PDB archive. This work has been carried out in collaboration with the RCSB PDB at Rutgers, PDBe, and the National Center for Macromolecular Imaging at Baylor College of Medicine.



The structure of the COPII transport-vesicle inner coat assembled on a membrane (colored ribbons; PDB ID 4bzi) along with the 3DEM map (grey; EMD-2428) that was determined in collaboration with 2013 Nobel Laureate Randy Schekman.

PDB ID 4bzi. G. Zanetti, S. Prinz, S. Daum, et al. *(2013) The structure of the COPII transport-vesicle coat assembled on membranes* eLife *2:e00951.*

References

1. H. M. Berman, J. D. Westbrook, Z. Feng, *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res* **28**: 235-242.
2. H. M. Berman, K. Henrick, H. Nakamura. (2003) Announcing the worldwide Protein Data Bank. *Nat Struct Biol* **10**: 980.
3. P. W. Rose, C. Bi, W. F. Bluhm, *et al.* (2013) The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res* **41**: D475-D482.
4. H. M. Berman, G. J. Kleywegt, H. Nakamura, *et al.* (2013) How community has shaped the Protein Data Bank. *Structure* **21**: 1485-1491.
5. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, *et al.* (1977) Protein Data Bank: a computer-based archival file for macromolecular structures. *J Mol Biol* **112**: 535-542.
6. H. M. Berman, K. Henrick, G. Kleywegt, *et al.* (2012). The Worldwide Protein Data Bank. In *International Tables for X-Ray Crystallography* (E. Arnold, D. M. Himmel & M. G. Rossmann, eds.), Vol. F: Crystallography of biological macromolecules, pp. 827-832 Springer, Dordrecht, The Netherlands.
7. S. Velankar, Y. Alhroub, C. Best, *et al.* (2012) PDBe: Protein Data Bank in Europe. *Nucleic Acids Res* **40**: D445-452.
8. A. R. Kinjo, H. Suzuki, R. Yamashita, *et al.* (2012) Protein Data Bank Japan (PDBj): maintaining a structural data archive and resource description framework format. *Nucleic Acids Res* **40**: D453-460.
9. E. L. Ulrich, H. Akutsu, J. F. Doreleijers, *et al.* (2008) BioMagResBank. *Nucleic Acids Res* **36**: D402-408.
10. J. D. Westbrook, P. M. D. Fitzgerald. (2009). Chapter 10 The PDB format, mmCIF formats, and other data formats. In *Structural Bioinformatics*, Second Edition (P. E. Bourne & J. Gu, eds.), pp. 271-291. John Wiley & Sons, Inc., Hoboken, NJ.
11. J. D. Westbrook, N. Ito, H. Nakamura, *et al.* (2005) PDBML: The representation of archival macromolecular structure data in XML. *Bioinformatics* **21**: 988-992.
12. P. M. D. Fitzgerald, J. D. Westbrook, P. E. Bourne, *et al.* (2005). 4.5 Macromolecular dictionary (mmCIF). In *International Tables for Crystallography* G. Definition and exchange of crystallographic data (S. R. Hall & B. McMahon, eds.), pp. 295-443. Springer, Dordrecht, The Netherlands.
13. J. Westbrook, K. Henrick, E. L. Ulrich, *et al.* (2005). 3.6.2 The Protein Data Bank exchange data dictionary. In *International Tables for Crystallography* (S. R. Hall & B. McMahon, eds.), Vol. G. Definition and exchange of crystallographic data, pp. 195-198. Springer, Dordrecht, The Netherlands.
14. M. D. Winn, C. C. Ballard, K. D. Cowtan, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Cryst* **D67**: 235-242.
15. R. A. Nicholls, F. Long, G. N. Murshudov. (2012) Low-resolution refinement tools in REFMAC5. *Acta Cryst* **D68**: 404-417.
16. P. D. Adams, P. V. Afonine, G. Bunkoczi, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst* **D66**: 213-221.
17. K. Henrick, Z. Feng, W. F. Bluhm, *et al.* (2008) Remediation of the Protein Data Bank Archive. *Nucleic Acids Res* **36**: D426-D433.
18. S. Dutta, D. Dimitropoulos, Z. Feng, *et al.* (2013) Improving the representation of peptide-like inhibitor and antibiotic molecules in the Protein Data Bank. *Biopolymers doi: 10.1002/bip.22434.*
19. C. L. Lawson, S. Dutta, J. D. Westbrook, *et al.* (2008) Representation of viruses in the remediated PDB archive. *Acta Cryst.* **D64**: 874-882.
20. H. M. Berman, K. Henrick, H. Nakamura, *et al.* (2007) The Worldwide Protein Data Bank (wwPDB): Ensuring a single, uniform archive of PDB data. *Nucleic Acids Res* **35:** D301-303.
21. M. Quesada, J. Westbrook, T. Oldfield, *et al.* (2011) The wwPDB common tool for deposition and annotation. *Acta Cryst* **A67**: C403-C404.
22. J. Y. Young, Z. Feng, D. Dimitropoulos, *et al.* (2013) Chemical annotation of small and peptide-like molecules at the Protein Data Bank. *Database* **2013**: bat079.
23. UniProt Consortium. (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* **40**: D71-75.
24. D. A. Benson, I. Karsch-Mizrachi, K. Clark, *et al.* (2012) GenBank. *Nucleic Acids Res* **40**: D48-53.
25. S. Gore, S. Velankar, G. J. Kleywegt. (2012) Implementing an X-ray validation pipeline for the Protein Data Bank. *Acta Cryst* **D68**: 478-483.
26. R. J. Read, P. D. Adams, W. B. Arendall, III, *et al.* (2011) A new generation of crystallographic validation tools for the Protein Data Bank. *Structure* **19**: 1395-1412.
27. G. T. Montelione, M. Nilges, A. Bax, *et al.* (2013) Recommendations of the wwPDB NMR Structure Validation Task Force. *Structure* **21**: 1563-1570.
28. R. Henderson, A. Sali, M. L. Baker, *et al.* (2012) Outcome of the first electron microscopy validation task force meeting. *Structure* **20**: 205-214.
29. J. Trewhella, W. A. Hendrickson, M. Sato, *et al.* (2013) Meeting report of the wwPDB small-angle scattering task force: Data requirements for biomolecular modeling and the PDB *Structure* **21**: 875-881.
30. The Gene Ontology Consortium. (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**: 25-29.
31. Enzyme Nomenclature, Enzyme Classification. **www.chem.qmw.ac.uk/iubmb/enzyme**.
32. E. W. Sayers, T. Barrett, D. A. Benson, *et al.* (2012) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **40**: D13-25.
33. A. G. Murzin, S. E. Brenner, T. Hubbard, *et al.* (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**: 536-540.

34. C. A. Orengo, A. D. Michie, S. Jones, *et al.* (1997) CATH–a hierarchic classification of protein domain structures. *Structure* **5**: 1093-1108.

35. WHO Collaborating Centre for Drug Statistics Methodology. (2013). Guidelines for ATC classification and DDD assignment, Oslo, 2012.

36. M. H. Saier, Jr., M. R. Yen, K. Noto, *et al.* (2009) The Transporter Classification Database: recent advances. *Nucleic Acids Res* **37**: D274-278.

37. P. W. Rose, B. Beran, C. Bi, *et al.* (2011) The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Res* **39**: D392-D401

38. Jmol: an open-source Java viewer for chemical structures in 3D. **www.jmol.org**

39. J. L. Moreland, A. Gramada, O. V. Buzko, *et al.* (2005) The Molecular Biology Toolkit (MBT): a modular platform for developing molecular visualization applications. *BMC Bioinformatics* **6**: 21.

40. C. Knox, V. Law, T. Jewison, *et al.* (2011) DrugBank 3.0: a comprehensive resource for 'omics' research on drugs. *Nucleic Acids Res.* **39**: D1035-1041.

41. D. Weininger. (1988) SMILES 1. Introduction and encoding rules. *J Chem Inf Comput Sci* **28**: 31.

42. S. Velankar, J. M. Dana, J. Jacobsen, *et al.* (2013) SIFTS: Structure Integration with Function, Taxonomy and Sequences resource. *Nucleic Acids Res* **41**: D483-489.

43. M. Punta, P. C. Coggill, R. Y. Eberhardt, *et al.* (2012) The Pfam protein families database. *Nucleic Acids Res* **40**: D290-301.

44. J. Haas, S. Roth, K. Arnold, *et al.* (2013) The Protein Model Portal--a comprehensive resource for protein structure and model information. *Database* **2013**: bat031.

45. G. Fermi, M. F. Perutz, B. Shaanan, *et al.* (1984) The crystal structure of human deoxyhaemoglobin at 1.74 A resolution. *J Mol Biol* **175**: 159-174.

46. H. Robinson, Y. G. Gao, X. Yang, *et al.* (2001) Crystallographic analysis of a novel complex of actinomycin D bound to the DNA decamer CGATCGATCG. *Biochemistry* **40**: 5587-5592.

47. C. Zardecki. (2008) Interesting structures: Education and outreach at the RCSB Protein Data Bank. *PLoS Biol 6*: **e117**.

48. S. Dutta, C. Zardecki, D. Goodsell, *et al.* (2010) Promoting a structural view of biology for varied audiences: an overview of RCSB PDB resources and experiences. *J Appl Cryst* **43**: 1224-1229.

49. H. M. Berman. (2012) Creating a community resource for protein science. *Protein Sci* **21**: 1587-1596.

50. H. M. Berman, B. Coimbatore Narayanan, L. D. Costanzo, *et al.* (2013) Trendspotting in the protein data bank. *FEBS Lett* **587**: 1036-1045.

51. D. S. Goodsell. (2012) Illustrating the machinery of life: viruses. *Biochem Mol Biol Educ* **40**: 291-296.

52. D. S. Goodsell. (2012) Putting proteins in context: scientific illustrations bring together information from diverse sources to provide an integrative view of the molecular biology of cells. *Bioessays* **34**: 718-720.

53. A. Prlić, A. Yates, S. E. Bliven, *et al.* (2012) BioJava: an open-source framework for bioinformatics in 2012. *Bioinformatics* **28**: 2693-2695.

54. M. J. Gabanyi, P. D. Adams, K. Arnold, *et al.* (2011) The Structural Biology Knowledgebase: a portal to protein structures, sequences, functions, and methods. *J Struct Funct Genomics* **12**: 45-54.

55. L. K. Gifford, L. G. Carter, M. J. Gabanyi, *et al.* (2012) The Protein Structure Initiative Structural Biology Knowledgebase technology portal: A structural biology web resource. *J Struct Funct Genomics* **13**: 57-62.

56. H. M. Berman, J. D. Westbrook, M. J. Gabanyi, *et al.* (2009) The protein structure initiative structural genomics knowledgebase. *Nucleic Acids Res* **37**: D365-368.

57. C. Y. Cormier, J. G. Park, M. Fiacco, *et al.* (2011) PSI:Biology-materials repository: a biologist's resource for protein expression plasmids. *J Struct Funct Genomics* **12**: 55-62.

58. E. S. Julfayev, R. J. McLaughlin, Y. P. Tao, *et al.* (2012) KB-Rank: efficient protein structure and functional annotation identification via text query. *J Struct Funct Genomics* **13**: 101-110.

59. C. L. Lawson, M. L. Baker, C. Best, *et al.* (2011) EMDataBank.org: unified data resource for CryoEM. *Nucleic Acids Res* **39**: D456-D464.

60. D. G. Kuroda, J. D. Bauman, J. R. Challa, *et al.* (2013) Snapshot of the equilibrium dynamics of a drug bound to HIV-1 reverse transcriptase. *Nat Chem* **5**: 174-181.

61. S. Hare, S. J. Smith, M. Metifiot, *et al.* (2011) Structural and functional analyses of the second-generation integrase strand transfer inhibitor dolutegravir (S/GSK1349572). *Mol Pharmacol* **80**: 565-572.

62. R. S. Yedidi, K. Maeda, W. S. Fyvie, *et al.* (2013) P2' benzene carboxylic acid moiety is associated with decrease in cellular uptake: Evaluation of novel nonpeptidic HIV-1 protease inhibitors containing P2 bis-tetrahydrofuran moiety. *Antimicrob Agents Chemother* **57**: 4920-4927.

63. M. Soundararajan, A. K. Roos, P. Savitsky, *at al.* (2013) Structures of Down syndrome kinases, DYRKs, reveal mechanisms of kinase activation and substrate recognition. *Structure* **21**: 986-996.

Molecular images were created using Chimera (E.F. Pettersen, T.D. Goddard, C.C. Huang, *et al.* (2004) UCSF Chimera–a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612.), and mMaya (MOLECULAR MAYA™, Version 1.3, Digizyme Inc., open-source at **www.moleculardmovies.com/toolkit/**).

The cover image of HIV envelope was created using autoPACK (Graham Johnson, Ludovic Autin, Mostafa Al-Alusi, David Goodsell, Michel Sanner and Art Olson; open-source at **autopack.org**).

wwPDB Members

RCSB PDB: **rcsb.org**
Protein Data Bank Europe: **pdbe.org**
Protein Data Bank Japan: **pdbj.org**
BioMagResBank: **www.bmrb.wisc.edu**

RCSB PDB Partners

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

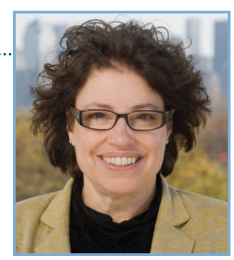
San Diego Supercomputer Center and the Skaggs School of Pharmacy and Pharmaceutical Sciences
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0743

RCSB PDB Management

DR. HELEN M. BERMAN

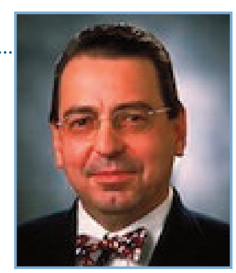
Director
Board of Governors Professor of Chemistry & Chemical Biology
Rutgers, The State University of New Jersey
berman@rcsb.rutgers.edu

DR. MARTHA QUESADA



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

DR. STEPHEN K. BURLEY



Associate Director
Rutgers, The State University of New Jersey
sburley@proteomics.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

RCSB PDB
PROTEIN DATA BANK

RCSB.ORG • INFO@RCSB.ORG