Protein Data Bank

Quarterly Newsletter

Release #76 *April 1996*

April 1996 CD-ROM Release

4432 full-release atomic coordinate entries

Molecule Type

4013	proteins, peptides, and viruses
100	protein/nucleic acid complexes
307	nucleic acids
12	carbohydrates

Experimental Technique

134	theoretical modeling
606	NMR
3692	diffraction and other

The total size of the atomic coordinate entry database is 1674 Mbytes uncompressed.

New Phone/Fax Numbers

PDB Main Telephone 516-344-3629
PDB Help Desk 516-344-6356
Fax 516-344-5751

Table of Contents

Internet Sites

WWW .... http://www.pdb.bnl.gov
FTP .... ftp.pdb.bnl.gov

What's New at the PDB

The PDB was established in 1971 under the leadership and direction of Dr. Walter Hamilton, Chemistry Department, Brookhaven National Laboratory, at the urging of members of the American Crystallographic Association (ACA) and based on discussions of probable future trends at the 1971 Cold Spring Harbor Symposium on Structure and Function of Proteins. From the beginning, the PDB has operated with continued support of the crystallographic community. The PDB has always been a truly international effort, initially with Affiliated Centers in Cambridge, United Kingdom; Melbourne, Australia; and Osaka, Japan. These centers, subsequently augmented by a number of on-line data providers, are all referred to as Affiliated Centers and are members of the PDB Service Association (PDBSA). There are currently twenty-four Affiliated Centers (see list inside back cover of this Quarterly Newsletter) as well as several full mirror sites of the PDB. Data acquisition and dissemination *via* tape media was on a global scale from the outset, with a small staff that handled approximately twenty-five structural depositions per year.

Rapid developments in both the growth of crystals of macromolecules and in experimental techniques for structure analysis have led to a revolution in Structural Biology. There has been an enormous increase in the number of laboratories performing structural studies to atomic resolution and at an ever-increasing speed. Advances include:

- Recombinant DNA techniques that permit almost any protein or nucleic acid to be produced in large amounts.
- More sensitive X-ray detectors.
- Real-time interactive computer graphics systems, together with automated methods for structure determination and refinement.
- Synchrotron radiation, allowing the use of smaller crystals, multiple wavelength anomalous dispersion (MAD) phasing, and time-resolved studies *via* Laue techniques.
- NMR methods permitting structure determination of macromolecules in solution.
- EM techniques for obtaining high-resolution structures of 2-D crystals.

These dramatic advances produced an abrupt transition from the linear increase of approximately twenty-five new structures deposited per year in the PDB before 1987 to a rapid exponential growth reaching the current rate of approximately twenty-five deposits per week. In the same period, the proliferation and increasing power of computers, the introduction of relatively inexpensive interactive graphics, and the growth of computer networks greatly increased the demand for access to PDB data — in many diverse ways. The requirements of molecular biologists, rational drug designers, and others in academia and industry were often fundamentally different from those of the crystallographers and computational chemists who had been the major PDB users since the 1970s.

As a result of this significant change in focus of those who deposit in and use the PDB, the eighteen-member staff of the PDB has just moved from Brookhaven's Chemistry Department to a newly-constructed wing of the Biology Department. With research programs in structural biology, molecular genetics, and DNA sequencing for the Human Genome Project, the Biology Department is an ideal location for the PDB.

This is the twenty-fifth year of the existence of the PDB — to celebrate, two special events have been planned:

- At the 17th International Union of Crystallography (IUCr) General Congress in Seattle this summer there will be a microsymposium entitled "Twenty-five Years of the PDB" (2 p.m., Sunday, August 11). For further information, see either URL **http://www.bmsc.washington.edu:80/iucr/** or URL **http://www.lmcp.jussieu.fr/sincris/iucr/**.

- An international meeting on "Bioinformatics<--->Structure" will be held in Jerusalem, Israel (November 17-21, 1996) in honor of the twenty-fifth anniversary of the PDB and the tenth anniversary of the Swiss-Prot Database. For additional information, see the article entitled "Joint Anniversary Celebration" within this Quarterly Newsletter, or either one of the following URLs: **http://www.pdb.bnl.gov/pdb25sw10** or **http://bioinformatics.weizmann.ac.il/conf/pdb25sw10**.

– *Joel L. Sussman*

Managing the Archives

– PDB Contents Guide, Version 2.0

The Protein Data Bank Contents Guide: Atomic Coordinate Entry Format Description, Version 2.0 was released on February 21, 1996. The purpose of the Contents Guide is to help several communities — depositors, software and information resource developers, and users of the PDB — to understand the contents of the coordinate entries. This format description is crucial in the effort to produce CIF-compliant data files from PDB entries.

PDB entries released after April 15, 1996 will be 2.0 compliant, and conversion of older entries to this new format will start in the fall of 1996. Beginning January 1997, the format of all PDB entries will be compliant with the current version of the Contents Guide. PDB will make sample entries and conversion tools available to the public. These will be accessible from URL **http://www.pdb.bnl.gov/Format.doc/Format_Home.html**.

Since publication of the Contents Guide, we have received many letters and phone calls related to both the document itself and the new format as described therein. Feedback falls into several categories:

- Corrections of typographical errors and misstatements.

- Requests for more information, such as further details to help clarify the semantics, examples, and diagrams.

- Recommendations for converting this to a fully HTML document.

- Suggestions for further changes to the format itself.

We are carefully considering each comment and will implement these suggestions starting in June 1996. Please watch our Web home page for news of the release of Contents Guide 2.1. Of course, any changes to the format itself will be implemented only after following the outlined Format Change Policy (please see related article in our October 1995 Quarterly Newsletter).

- **Changes to All PDB Records**

  Columns 71 - 80 now contain data. They previously contained the PDB ID code and record serial number. These items may be generated using scripts available from the PDB.

- **Changes to ATOM/HETATM Records**

  A segment identifier has been added to the coordinate records in columns 73 - 76. This allows unambiguous identification of regions of the chains and the relationship between them by specifying segments of molecules.

  The element symbol and charge now appear in columns 77 - 80 of the coordinate records.

  When temperature factors are provided, the tempFactor field (columns 61 - 66) always contains the isotropic B value, even when ANISOU records are provided.

  Insertion codes (column 27) are now defined as being alphabetic only.

- **Changes to Other Records**

  HELIX records now contain the length of the helix in columns 72 - 76.

  FTNOTE (Footnotes) has been dropped.

  SSBOND now includes the symmetry operator if a residue involved in the disulfide bond is in a symmetry-related unit.

In CRYST1 records:

- The full international Hermann-Mauguin symbol is used, *e.g.*, P 1 21 1 instead of P 21.
- For a rhombohedral space group in the hexagonal setting, the lattice-type symbol used is H.

A number of record types which previously contained free text have been restructured as follows:

- “Keyword: value” pairs have been introduced in certain records such as COMPND and SOURCE to allow easier parsing.
- EXPDTA has been expanded and now appears in every PDB coordinate entry.
- REMARK records have been restructured to allow easier parsing and to bring more organization to these records.

- **New Record Types Added**

TITLE	HETNAM
CAVEAT	HETSYN
KEYWDS	LINK
MODRES	HYDBND
DBREF	SLTBRG
SEQADV	CISPEP

For details on each of these changes, see the section of the associated record type in the Contents Guide.

Postdoctoral Fellowship with the Protein Data Bank

The Alfred P. Sloan Foundation and the U.S. Department of Energy are jointly sponsoring a program that awards postdoctoral research programs for scientists interested in computational biology. Anyone interested in developing an application in conjunction with the Protein Data Bank should contact us immediately. The deadline for submission of proposals is August 1, 1996. Please direct inquiries to Dr. Enrique Abola.

Joint Anniversary Celebration

In celebration of the Protein Data Bank’s twenty-fifty anniversary and Swiss-Prot’s tenth anniversary, the 24th Aharon Katzir-Katchalsky Conference on Bioinformatics <--> Structure will be held in Jerusalem, Israel, November 17-21, 1996. The meeting will focus on the relationship of sequence, structure, function, and databases. Sessions will include talks, poster presentations and computer demonstrations with access to the Internet.

The scientific program will cover:

- **3-D Structures**

  Proteins, Nucleic Acids, Protein-Nucleic Acid Interactions Folding, Classification, and 3-D Patterns.

- **Biological Sequences**

  Sequence Analysis, Profiles and Patterns, Molecular Evolution, Algorithm Development, Sequence to Function, and Gene Organization.

- **Modeling & Prediction**

  Knowledge-based Modeling, Threading, 3-D Profiles, Protein Engineering, Molecular Recognition, Structure-Based Drug Design, Molecular Graphics, and Validation.

- **Databases**

  3-D Structure Databases, Protein/Nucleotide Sequence Databases, Genomic Databases, Federation and Integration, and Database Queries.

Invited speakers include:

Enrique Abola (USA)	Michael Levitt (USA)
Amos Bairoch (Switzerland)	Victor Markowitz (USA)
Helen Berman (USA)	Manuel Peitsch (Switzerland)
Tom Blundell (UK)	Jaime Prilusky (Israel)
Peer Bork (Germany)	Jane Richardson (USA)
Graham Cameron (UK)	Chris Sander (Germany/UK)
Cyrus Chothia (UK)	Manfred Sippl (Austria)
David Eisenberg (USA)	William Studier (USA)
Ken Fasman (USA)	Joel L. Sussman (Israel/USA)
Leroy Hood (USA)	Tomitake Tsukihara (Japan)
Rod Hubbard (UK)	Keith D. Watenpaugh (USA)
Ephraim Katchalski-Katzir (Israel)	Keith Wilson (UK/Germany)
John Kendrew (UK)	Ada Yonath (Israel)
Doron Lancet (Israel)	

The meeting will be held under the auspices of the Protein Data Bank, Brookhaven National Laboratory, University of Geneva Israel Academy of Science and Humanities, and The Weizmann Institute of Science.

The Organizing Committee consists of Enrique E. Abola, Amos Bairoch, Manuel Peitsch, and Joel L. Sussman. The Local Organizing Committee consists of Marvin Edelman, Michal Harel, Michael Levitt, Uriel Z. Littauer, Jaime Prilusky, and Joel L. Sussman.

The deadline for registration and poster abstract submission is June 15, 1996. Please access the conference information and the registration form at either **http://www.pdb.bnl.gov/pdb25sw10** or **http://bioinformatics.weizmann.ac.il/conf/pdb25sw10**.

Updated World Wide Web Home Page

PDB is pleased to announce a new look for our Web home page, URL **http://www.pdb.bnl.gov**. Those who have accessed this site since March 27, 1996 may have noticed a new design. The home page serves as a roadmap to our WWW site, providing links to several sub-pages as well as statistics on the contents of the archives and a mail link for sending mail to the PDB Help Desk. The sub-pages are organized to quickly lead you to what you need.

PDB plans to continue to upgrade our Web site to provide the full range of information, tools, and links that our user community needs. We invite your suggestions and comments on our effort.

A brief summary of each page follows.

- **Searching and Browsing the PDB**

  This is where to access the PDB Browser for searching the coordinate entries and the Swiss-3DIMAGE Browser for searching the annotated 3D images created by Manuel Peitsch of GLAXO Institute for Molecular Biology, Geneva, Switzerland. Lists of pending and on-hold entries and molecule types and sequences are found here, as well as indices to the entries and links to the prepared biological molecule coordinate files.

- **Submitting Data to the PDB**

  Anyone wishing to deposit new data should jump to this page for instructions, guidelines, and to access the latest Deposition Form and Contents Guide.

- **PDB FTP Server**

  This link takes you to the PDB anonymous FTP server, at **ftp.pdb.bnl.gov**. From here you may access any of the entries, programs, PDB Quarterly Newsletters, Full Tables, and other documents that are publicly available.

- **PDB Documentation**

  The PDB Documentation page allows quick and convenient access to our documents. The PDB Contents Guide can be viewed and retrieved in several formats. A link to its Web page gives an introduction to the document as well as links to related files. The PDB Quarterly Newsletters are found here as well as definitive copies of our deposition and format change policies.

- **PDB Mirror Sites**

  Here are links to official PDB Mirror Sites. For more information on mirror sites, please see the following article. New mirror sites will be announced and added to this page as they come on-line.

- **PDB User Group**

  Information on the PDB User Group, headed by Jane Richardson of Duke University, is found here. Links to the Pending and Waiting List, the directory of biological molecule coordinates, the user_group directory on our FTP server, and a mail link to the User Group are provided.

- **mmCIF - Macromolecular Crystallographic Information File**

  This page contains links to several resources related to the macromolecular Crystallographic Information File (mmCIF). For further information, please see the article entitled “mmCIF Dictionary” within this Quarterly Newsletter.

- **Molecular Biology Servers/Databases/Web Sites of Interest**

  Here are links to sites of interest to our user community, including related databases, browsers, institutions, and associations. This page is continually growing and can serve as a springboard to your Web surfing.

- **Software and Related Information**

  PDB provides several software programs to the public, some written in-house and others written elsewhere and distributed by us. This is where you may access these programs as well as obtain information on several other program suites.

Mirror Sites

The PDB is pleased to announce the establishment of three international mirror sites. These sites will enable users from all parts of the world to access all of PDB's archives, search tools, documentation, programs, *etc*. We encourage you to try out each of these sites in order to determine which may be better for your own access needs. Our appreciation is extended to those responsible for setting up the mirror at each location. We will continue to offer our complete support and encouragement to them as well as to other sites as they are established.

- **Weizmann Institute of Science, Rehovot, Israel**

  http://pdb.weizmann.ac.il
  Jaime Prilusky and Sasha Faibusovich

  A complete PDB mirror site including all of the FTP directories was set up in the Bioinformatics Unit, at the Weizmann Institute of Science, Rehovot, Israel. It also cross-links to the local SCOP mirror site.

- **European Bioinformatics Institute, EMBL Outstation, Hinxton Hall, United Kingdom**

  http://www.ebi.ac.uk/PDB
  Philip McNeil

  A complete PDB mirror site including all of the FTP directories was set up at EBI, Hinxton Hall, United Kingdom.

- **Institute of Physical Chemistry, Peking University, Beijing, China**

  http://www.ipc.pku.edu.cn/npdb/
  Luhua Lai and Dawei Lin

  A complete PDB mirror site including all FTP directories was set up at the Institute of Physical Chemistry, Peking University, Beijing, China. It also cross-links to the local SCOP mirror site and the first China Bioinformatics server. Users in China and nearby locations may find this site helpful.

Three additional European mirrors will be operational in the very near future. These are the Cambridge Crystallographic Data Centre, Cambridge, United Kingdom; IGBMC Laboratory of Structural Biology, Strasbourg, France; and Uppsala University Department of Molecular Biology, Uppsala, Sweden. Each will be announced as soon as it comes on-line.

Function Discussion

A lively discussion on the PDB Listserver pertaining to the adding of a description of biological function to PDB entries was initiated by Sandeep Kumar from the Indian Institute of Science, Bangalore, India with significant input from Eleanor Dodson, York University, United Kingdom, and Sherry Mowbray, Uppsala University, Sweden. The relevant letters are available for perusal on our Web server at **http://www.pdb.bnl.gov/listserv.html**.

Currently, PDB entries allow the inclusion of functional information in PDB HEADER, KEYWDS, and REMARK records. In addition, the PDB DBREF record points to related entries in other databases. Some, for example SWISS-PROT, contain a description of biological function. PDB is also developing a relational database, 3DBase, that has facilities to include user-supplied annotations. One use of this facility is envisioned to be linking PDB entries to specialized databases that describe protein activity or function. If such a database were built we would be prepared to provide active links. Alternatively, anyone wishing to annotate PDB entries with descriptions of biological function can do so using this mechanism. A description of 3Dbase is available from URL **http://www.pdb.bnl.gov**.

San Diego Supercomputer Center Macromolecular Structure Database

*This article was written by Philip E. Bourne, San Diego Supercomputer Center, San Diego, CA, USA **(bourne@sdsc.edu)**.*

A macromolecular structure database is available to WWW users from multiple servers located at the San Diego Supercomputer Center (SDSC). This database may be accessed from URL **http://www.sdsc.edu/moose**.

The database, which contains a variety of native and derived data, is incremented daily from the FTP archives of the PDB. A summary of the current contents of the database is available from the WWW page. Three classes of query are supported: text, entity, and pattern. Each returns a subset of structures which may be examined graphically (3-D rendering, property profiles, contact maps, and snapshot summaries, *etc.*) or used as input to a more detailed query.

Entity queries search for structures based on features of each polypeptide chain, *e.g.*, arrangement of secondary structure elements, percent alpha or beta, and molecular weight.

Pattern queries search for property patterns based on the amino acid sequence which may be environmental (exposure, polarity, and B values), static (volume, isoelectric point, and hydrophobicity), or a combination of both. Query patterns can be derived from known structures in known sequences (within some threshold) or in relation to absolute values.

Hashing techniques are used to provide real-time query and as an aside provide current tables of property value distributions across the whole database.

Current developments include making the database, query language, and Web interface into an export package for users to construct their own databases and to support 3-D motif searches.

The SDSC macromolecular structure database was developed by Ilya Shindyalov, Rich Toscano, and Phil Bourne. Comments are welcome and may be made *via* the Web page.

mmCIF Dictionary

*This article was written by Paula M. D. Fitzgerald, Merck Sharp and Dohme Research Laboratories, Rahway, NJ, USA **(paula_fitzgerald@Merck.Com)**.*

Version 0.8.03 of the macromolecular Crystallographic Information File (mmCIF) dictionary is now available. For those not already familiar with this project, mmCIF provides a data structure, dictionary, and associated software tools for the archiving and interchange of a variety of data pertaining to the macromolecular crystallographic experiment and its results. The mmCIF project is an initiative of the IUCr.

Rather than providing full details of mmCIF here — its history, its structure, and its many uses, we invite you to visit the mmCIF Web page **(http://ndbserver.rutgers.edu/mmcif)** which contains all of this information and more. Once there, you will find pointers to other CIF materials and other WWW sites containing software and related information. If more convenient, there is also a full mirror of the mmCIF home page in England **(http://www.ebi.ac.uk/NDB/mmcif)**.

The mmCIF dictionary was released for public review in August 1995 and that review process has now been completed. You are invited to participate in the final stage of the review process which will happen between now and June 15, 1996. Then the dictionary will be finalized and published. Although changes are still possible in this final review stage, we do not anticipate any of them affecting the basic structure of the data model, only the details of that model. We welcome comments and suggestions.

We have used the mmCIF mailing list as the mechanism for dialog during the review process.

To subscribe to the mmCIF List Server, send a one line message containing the text:

subscribe mmciflist Your Name

to the e-mail address:

requests@ndbserver.rutgers.edu

To post to the list, send messages to the e-mail address:

mmciflist@ndbserver.rutgers.edu.

Your input is welcome. Please contact one of the members of the mmCIF Working Group:

Paula Fitzgerald.................paula_fitzgerald@Merck.Com
Helen Berman....................berman@dnarna.rutgers.edu
Phil Bourne........................bourne@sdsc.edu
Brian McMahon ....b.mcmahon@iucr.ac.uk
Keith Watenpaugh ....kdwatenp@radon.upj.com
John Westbrook.................westbrook@rutchm

Announcement — O Rebuilding Tutorial Available

This article was written by Gerard J. Kleywegt, Department of Molecular Biology, Biomedical Centre, Uppsala University, Uppsala, Sweden ***(gerard@xray.bmc.uu.se)***.

The tutorial "O for the Structurally Challenged," designed to introduce protein-model rebuilding and quality control using O and refinement using X-PLOR, has recently been released. It may be obtained *via* anonymous FTP from **rigel.bmc.uu.se** (130.238.37.29) in the **/pub/tutorial** directory as file **rebtut.dirtar.Z**. When uncompressed and tar-ed, it will create a directory named rebtut which contains a README file; the text of the tutorial in plain ASCII, HTML, and PostScript format; and a directory tree named gmrp.

The tutorial contains the following ten chapters:

- Introduction
- Preparations
- Quality Checks
- Oops
- Rebuilding
- Preparing For Refinement
- Refinement
- Post-refinement
- Another Cycle
- Miscellaneous Refinement Issues

Working with "O for the Structurally Challenged" requires:

- Basic familiarity with O, *e.g.*, through the "O for Morons" tutorial.
- O version 5.10 x and an O manual.
- Access to the O public domain directory, generically called OMAC. Check the correct name on your local system. If not available, you may obtain it from the O FTP server **(/pub/gerard/omac.dirtar.Z)**.
- The list of Frequently-Asked Questions **(OMAC/software.faq)**. This document is also available from the Web server at URL **http://kaktus.kemi.aau.dk/.**
- Several Uppsala utility programs (also available from the O FTP server — you will need the manuals for these programs).
- X-PLOR version 3.1 or later and the X-PLOR manual.
- CCP4 programs (optional; they are used for map calculations, but they may also be carried out with X-PLOR).

This tutorial may be used in several different ways:

- As a quick introduction to quality control and rebuilding (Chapters 2, 3, 4, 5, and optionally, 8 and 9).
- As a quick introduction to refinement, only for absolute X-PLOR beginners (Chapters 6 and 7).
- As a complete course in rebuilding and refinement from the first to the final model (all chapters, including several iterations of Chapter 9).

Typically, Chapters 5, 7, and 9 will be the most time-consuming, depending on the level of experience the student has with O and rebuilding, and refinement in general. Chapters 2 through 8 contain a few questions in the section "The Swedish Inquisition". Generally, these will aid or extend the understanding of the subject matter covered in each chapter.

The tutorial does not include *ab initio* model-building in an MIR map. For this purpose, there are some macros on the O FTP server **(/pub/p2_course)**. Also, the structure used as an example has no NCS, so averaging is not covered either. There is, however, a separate RAVE tutorial available, also from the O FTP server **(/pub/gerard/rave/exam.dirtar.Z)**.

The structure being worked with is that of cellular retinoic-acid binding protein type II (CRABP II) in complex with all-trans-retinoic acid. This structure was solved at 1.8 Å, but we use only data to 2.8 Å initially to show the effects that limited resolution may have. The starting model is derived from a related protein, CRABPI, which was solved at 2.9 Å resolution. This structure was changed as follows:

- The N and C terminal residues were removed.
- The region near the insertion in CRABPII was removed (residues 114 - 117).
- A loop with poor density and high temperature factors was removed (residues 100 - 106).
- Residues which differ between CRABPI and II were cut back to alanines (unless they are glycines in CRABPI).

The temperature factors of this model were reset, and it was subjected to mild simulated annealing refinement. The structure was then "rebuilt" to introduce some deliberate errors; some energy minimization and individual temperature-factor refinements were carried out; and 2Fo-Fc and Fo-Fc maps were calculated.

This structure has the advantage that it is fairly small, yet contains most of the common errors and problems (main chain, side chain, poor loop, insertion site, and low resolution) associated with protein model refinement and rebuilding.

Before starting, copy your local version of the gmrp directory tree to your area. This contains all the files needed. Begin working in the O directory **/gmrp/o**. Additionally, ask your system manager to install the "run", "ono", and "oplot" scripts from the FTP server **(/pub/gerard/scripts)** to make your O-life simpler. It may be useful to have a copy of the O manual and tutorial at hand. Answers to most questions may be found in the literature references, program manuals, or by inspection of the model, the density, or one of the files.

It may be handy to use a Web browser to consult the O manual and FAQ **(http://kaktus.kemi.aau.dk/)**. Manuals for the utility programs are also available in HTML format — in Uppsala, from our home page and elsewhere from the O FTP server **(/pub/gerard/html_manuals.dirtar.Z)**. Documentation for X-PLOR and CCP4 programs is also available on the Web. While working through the tutorial, you may want to use the Good Model-Building and Refinement Forms for keeping notes **(OMAC/gsp_forms.ps)**.

Rebuilding is a lot more fun if you play loud music on your personal stereo!

Comments and suggestions pertaining to this tutorial should be sent to **gerard@xray.bmc.uu.se**.

Naomi — Analysis, Refinement, Modelling, Prediction, and Design

*This article was written by Simon M. Brocklehurst, Oxford Centre for Molecular Sciences, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK **(smb@bioch.ox.ac.uk)**.*

NAOMI is an integrated computer program system for studying the three-dimensional structures of proteins at the atomic level. Developed by Simon M. Brocklehurst of the Oxford Centre for Molecular Sciences at the University of Oxford, UK, the program allows the user to perform manipulations of and analyses on protein structures deposited in the PDB. The program is aimed at both the specialist and non-specialist user and has been under continual development since 1991. Since its release in mid-1995, the program has been licensed by several hundred laboratories worldwide.

Almost all of the algorithms used in the program are novel. For example, the secondary structure recognition algorithm employs fuzzy logic to define elements of structure, and the hydrogen bond energy calculations take into account both electrostatic effects (using explicit electron lone-pairs and hydrogen atoms) and the quality of wave-function overlap. Many of the features in NAOMI are unique and thus the program may be regarded as being complementary to other protein structure-related software. Considerable effort has also been spent in optimizing the speed at which the program runs. For example, solvent accessibility calculations typically take about ten seconds for one thousand atoms on a mid-range workstation. The program is installed as a single executable, but different features may be accessed by obtaining license keys which switch on various modules.

One of the key points about NAOMI which distinguishes it from many other program systems is that it was designed to be easy to use for both the structure-specialist (*e.g.*, NMR spectroscopist or X-ray crystallographer) and non-specialist (*e.g.*, molecular biologist or protein engineer). The program includes interactive clickable documentation and is operated by using concise, intuitive English-like commands. For example, to identify potential ion-pairs in a protein structure, the command "use salt-bridges" followed by the command "table salt-bridges" would instruct NAOMI to first compute the positions of ion-pairs in the protein and then output the results in a table that may be easily read by the user.

Typical users may be interested in the following:

- **General Structure Analysis:** ion-pairs, hydrogen bonds, key residues, secondary structure, super-secondary structure, disulphide bridges, hydrophobic packing, solvent accessibility calculations, or amino acid hydrophobicity scales.

- **Bioinformatics:** incorporation of three-dimensional structural information into multiple amino-acid sequence alignments.
- **Structure Modelling:** mutation of residues, manipulation of main-chain and side-chain conformations, minimization of bad contacts, or building of missing atoms into a structure.
- **NMR Structure Refinement:** analysis of ensembles to identify hydrogen bond restraints (including homo- and hetero-oligomeric structures), rapid generation of truly random starting structures prior to beginning structure calculation, or prediction of expected NOEs from structure.
- **Engineering and Design:** design of stabilizing disulphide bonds and design of stabilizing PRO->non-PRO mutations.
- **Prediction of Protein Folding Pathways:** production of three-dimensional models of folding intermediates.
- **Characterization of Protein-Protein Complexes:** characterization of protein binding sites into component hydrophobic, hydrogen bonding, and electrostatic components.
- **Prediction of Protein-Protein Recognition Sites:** prediction of which residues in a protein may be important for protein-protein recognition from a knowledge of only a single structure.

By using commands in loops, analyses can easily be performed on the entire PDB. Commands are provided to automatically locate structures from their PDB ID codes alone if the user's local copy of the database maintains the standard PDB directory hierarchy. Graphical output is provided in the form of interfaces to several computer programs (primarily to MOLSCRIPT and RASMOL) in order to facilitate time-efficient preparation of high quality figures. For example, information in The Cytokines Web (**http://www.ocms.ox.ac.uk/~smb/cyt_web/**) was largely prepared automatically using NAOMI to produce schematic illustrations of structures and to characterize and illustrate binding sites on cytokine and receptors. NAOMI may be downloaded from the Web **(http://www.ocms.ox.ac.uk/~smb/Software/N_details/naomi.html)** or *via* anonymous FTP **(ftp://nmrz.ocms.ox.ac.uk/pub/smb/naomi/)**.

NAOMI runs only on Silicon Graphics workstations running IRIX 5.0 or greater. For further information, please contact Simon Brocklehurst, Department of Biochemistry, University of Oxford **(smb@bioch.ox.ac.uk)**.

The PDB Pipeline

*This article was written by David J. Schuller, Department of Molecular Biology, University of California, Irvine, CA, USA **(schuller@uci.edu)**.*

The PDB Pipeline, a collection of simple utilities for manipulating PDB coordinates, is freely available *via* anonymous FTP from **indigo4.biomol.uci.edu** in the **pub/schullersoft/pdb.tar.Z** directory. This is a compressed Unix tar file, therefore, use binary mode for retrieval. See the README file for information on other available software.

Users working with PDB coordinates on Unix systems may find these utilities helpful, particularly if it is necessary to move coordinates between two or more programs with different implementations of the PDB format.

The two basic concepts behind the PDB Pipeline are:

- Each manipulation is performed by a short, simple program.
- The programs read from standard input and write to standard output so that they work well with Unix command piping.

The current version contains approximately seventy-five separate utilities, with the longest source file containing four hundred lines. This makes it easy to understand program function by dissecting the source code. YES! source code is included! If you need a specific functionality not provided, it may be possible to hijack one of the existing utilities for your own purposes with minimal coding effort.

The use of standard input and output (FORTRAN units 5 and 6) allows several utilities to be combined by piping output from one into another. The support of command line piping also allows PDB Pipeline utilities to be easily combined with Unix utilities like grep, sed, and awk or any pipe-ready homegrown utilities you may have.

These utilities will do the following: select coordinates by residue zone, chain ID, segment ID, or NMR model; set temperature factor, occupancy, residue number, segment ID, or chain ID; apply rotations and translations to coordinates; orthogonalize or fractionalize coordinates; truncate sidechains to poly-alanine; add OXT terminal oxygen atoms; derive sequence, volume, or mean temperature factor from coordinates; and more.

As an example; you have just refined your model in program X and need to fix up the coordinates in order to view them in graphics program F (we won't name names here). You want to remove the hydrogens, restore the chain IDs which were lost by program X, change your water residue names from "WAT" to "HOH", and since you refined with strict NCS, you want to generate coordinates for viewing the second molecule. Your script file may look similar to this:

```
pdb_remh X.pdb | pdb_chain -A | sed -e “s^WAT^HOH^g” |tee F.pdb |\
pdb_rotm -f ncs.matrix |pdb_chain -B >>F.pdb
```

where ncs.matrix is a file containing the NCS rotation/translation operator. Note that this work was performed using Unix utilities sed and tee. The functions of the PDB Pipeline utilities are, hopefully, obvious from their names.

The code is mainly FORTRAN77, with some extensions. All of the programs compile and run on an SGI Indigo with Irix 5.3

and SGI's FORTRAN compiler — they should perform well with any modern Unix system with some tweaking of the makefile.

Comments are welcomed at **schuller@uci.edu**.

Notes of a Protein Crystallographer

– A Brilliant Light in the Midwest Firmament

This article was written and submitted by Cele Abad-Zapatero, Abbott Laboratories, Abbott Park, IL, USA ***(abad@abbott.com)****. It has previously appeared in the Spring 1996 American Crystallographic Association Newsletter.*

On Friday night, January 12, 1996, the Advanced Photon Source (APS) achieved a stored beam current of 100 mA inside the storage ring of the first third-generation synchrotron source in the hard X-ray range in America. The APS will be officially dedicated on May 1, 1996. These milestones signal the beginning of a bright future for the community of macromolecular crystallographers in the U.S. and for the structural community at large. Chicago, the Midwest city symbolized by buildings such as The Sears Tower and The Art Institute, will now have a different icon to represent it: the unique ring of the APS, located at Argonne National Laboratory.

The APS was conceived in 1984 by a scientific committee assembled by the Office of Basic Energy Sciences of the U.S. Department of Energy (DOE). The committee recognized the tremendous impact that high-brilliance x-ray beams, generated by a specially designed machine, could have in research programs in materials science, physics, biology, chemistry, geosciences, and many other fields. Following the recommendations of the committee and years of design and planning, ground was broken in June 1990 and on the evening of August 9, 1995, the first X-rays were observed on the APS experimental floor.

It is probably not appropriate to list here the technical specifications of the APS and compare them with the corresponding values at other synchrotrons in the U.S., which already offer beamlines for several diffraction applications including macromolecular crystallography. Those details may be found at the following Web sites:

APS: http://www.aps.anl.gov/welcome.html
NSLS: http://www.nsls.bnl.gov
CHESS: http://www.chess.cornell.edu/
SSRL: http://ssrl01.slac.stanford.edu/ssrl.html
ALS: http://beanie.lbl.gov:8001/als/als_homepage.html

Prospective users should carefully study the technical information at these sites when planning experiments at a synchrotron radiation facility. Most of these sites offer cross references to other sites of interest for the synchrotron and x-ray community.

Nonetheless, there are two general considerations that will attest to the future importance of third-generation synchrotron X-ray sources like ALS (Berkeley), ESRF (Grenoble), APS (Chicago), and SPring8 (Japan). On the technical side, these facilities have been designed specifically to generate X-ray beams of unprecedented brilliance, and the impact that these sources will have on our understanding of the material world surrounding us can only be a matter of conjecture. On the sociological side, these installations have been planned and executed with the user community in mind. Although this may appear trivial, this second factor should not be overlooked. One has to consider for instance the amount of space allocated to the estimated two thousand researchers that will visit the APS per year, with an approximate floating population of three to four hundred investigators at any one time. The planning team of the APS has provided additional space on the periphery of the ring for offices and laboratories (laboratory office modules or LOMs) adjacent to the experimental hall. These clusters of five triangular extensions give the APS its unique serrated appearance. In addition, a user residence facility is located within half a kilometer from the storage ring; this residence will provide accommodations for up to 240 users and will provide computer links to the experimental hall.

Future access to the APS experimental facilities by qualified researchers will be as members of Collaborative Access Teams (CATs) or as Independent Investigators. A CAT consists of investigators with a common research focus; each will be responsible for its own planning, construction, funding, and operation of beamlines specifically designed with a certain experimental focus. CATs are required to allocate a certain percentage of beamtime to Independent Investigators.

When completed, the APS will offer its user community a total of approximately seventy beamlines, one half bending-magnet beamlines and the other half using insertion devices as the radiation source. Fifteen CATs are now actively pursuing the planning, design, and construction of beamlines at the APS with applications covering the three domains in the study of condensed matter: diffraction, spectroscopy, and imaging.

For experiments in the realm of macromolecular structure, four CATs are currently building experimental stations with different but complementary focus. Bio-CARS, within the Consortium for Advanced Radiation Sources, will construct three experimental stations: two on the bending-magnet line and the other directed toward collecting data from small and/or weakly diffracting crystals. Beamline optics are designed for crystals with very large unit cells such as crystalline samples of viruses and ribosome-like aggregates. The stations will allow data to be collected in both monochromatic and polychromatic (Laue and time-resolved) mode including MAD (Multiwavelength Anomalous Diffraction) phasing experiments. They will have biohazard facilities at Bio-safety level 3 and will offer twenty-four hour per day user support.

The Structure Biology Center (SBC-CAT) is designed as a user facility for macromolecular crystallographers. The technical facilities have been designed specifically to support the following:

- MAD phasing experiments.
- Data collection from microcrystals.
- Data collection from crystals of large macromolecular aggregates.
- Routine rapid data collection from similar or related crystal structures.

In addition, the SBC-CAT will provide users with a full complement of instrumentation, ancillary facilities, software, and support staff to enhance productivity.

The Industrial Macromolecular Crystallography Association (IMCA), a consortium of twelve chemical and pharmaceutical companies, together with the Illinois Institute of Technology, form the IMCA-CAT. This group is pursuing the construction of two beamlines (one sector) with the objective of collecting diffraction data from macromolecular crystals of biomedical interest to aid in their drug design and protein engineering programs.

Bio-CAT is planning experimental stations to study the structure and dynamics of partially ordered samples such as membranes and fibers. The main research techniques will be resonant (anomalous) and non-resonant X-ray diffraction as well as X-ray absorption fine structure (EXAFS) spectroscopy, with emphasis on time-resolved studies, polarized XAFS, and combined diffraction/spectroscopic techniques. The equipment, software, and laboratories of the Bio-CAT station will be available to a wide variety of users.

A truly brilliant star has been born in the Midwest Sky which, together with the ones already existing in the world and the ones currently under construction, will illuminate our path toward a better understanding of the atomic landscapes, islands, continents, and universes that surround us.

Second Meeting on the Critical Assessment of Techniques for Protein Structure Prediction

This article was written by Tim Hubbard, Centre for Protein Engineering, Cambridge, UK ***(th@mrc-cpe.cam.ac.uk)****.*

Methods for obtaining information about protein structure from the amino acid sequence have been advancing rapidly. But just what can these methods currently deliver? The first large-scale experiment aimed at beginning to answer these questions was conducted in 1994 and culminated at a meeting in Asilomar, California at the end of that year. Some 135 predictions were made by thirty-five different groups. The results are published in a special issue of *Proteins: Structure, Function and Genetics*, **23,** No. 3, (Nov. 1995). Meeting abstracts and predictions are available at URL **http://iris4.carb.nist.gov/** or at the European mirror URL **http://www.mrc-cpe.cam.ac.uk/mirrors/carb/**.

We are now announcing the second experiment. As with the first experiment, the goal is to obtain an in-depth and objective assessment of our current abilities and inabilities in this area. For this to occur, participants will predict as much as possible about a set of soon-to-be-known structures in advance of the meeting. Sessions will be devoted to presentation of the results, comparison with the experiment, and a description of the methods used.

As previously, for the experiment to succeed, it is essential that we obtain help from the experimental community. Therefore, we would like to invite protein crystallographers and NMR spectroscopists who expect to solve a structure before October 1, 1996 to submit the sequence so that attempts may be made to predict it before the structure is publicly announced. Each prediction will be given a deadline prior to the initial date information about the structure will be made public.

Targets of all sizes and types are required. Small structures (less than one hundred residues) are needed to test some of the *ab initio* structure prediction methods. Proteins with folds related to those of known structures are needed to test fold identification methods. Proteins with sequences homologous to one or more known structures are needed to test comparative modeling methods. Protein-protein and protein-ligand complexes are required to test docking methods.

Participants are asked to provide the following:

- The sequence or a sequence database accession number of the protein.
- An estimate of the expected date of public release and subsequent updates as work proceeds.
- A commitment to make the coordinates available to the independent assessors no later than October 1, 1996, should the structure be solved by then. If necessary, the coordinates will be for limited distribution until the meeting.

Coordinates provided will be treated with strict confidentiality and used only to evaluate the accuracy of predictions. For further information and on-line forms and documents, please see the Web sites **http://iris4.carb.nist.gov/casp2/** or **http://www.mrc-cpe.cam.ac.uk/casp2/**

To be useful to the predictors, a period of at least one month is required before any details of the structure will be released. Please notify us immediately when the details are going to be made public, so that we may ask the predictors to stop work in a timely manner. This can be done by sending e-mail to **casp2@mrc-lmb.cam.ac.uk**. In order for the predictions to be assessed in time for the meeting in December 1996, sets of coordinates will be needed by October 1, 1996 at the very latest.

The meeting organizers are:

Tim Hubbard, Co-chair, Centre for Protein Engineering, Cambridge, UK.

Steve Bryant, Co-chair, NCBI, National Library of Medicine, USA.

John Moult, President, CARB, University of Maryland, USA.

Jan T. Pedersen, CARB, University of Maryland, USA.

Krzysztof Fidelis, Lawrence Livermore National Laboratory, USA.

Richard Judson, Sandia National Laboratory, USA.

Understanding Protein Structure Determination

*This article was written by Lennart Nilsson, Karolinska Institutet, Department of Biosciences at NOVUM, Center for Structural Biochemistry, S-141 57 HUDDINGE, Sweden **(Lennart.Nilsson@csb.ki.se)**.*

The Karolinska Institute's Center for Structural Biochemistry will hold its sixth summer school, entitled "Understanding Protein Structure Determination" on September 1-6, 1996. This is one in a series of short graduate courses organized each summer by the Summer University of Southern Stockholm at NOVUM Research Park, located about 15 km south of Stockholm. As usual we are aiming for an informal format with students on the graduate and postdoctoral levels. A fair amount of time will be set aside for discussions and social activities.

The topics of the 1996 summer school are intended to cover the two major protein structure determination techniques: X-ray crystallography and NMR spectroscopy. Topics such as the physical basis, experimental aspects, computational and software aspects, quality assessment of determined structures, and effects of structural heterogeneity and dynamics will be addressed. Emphasis will be put on the complementarity of the two techniques.

The intended speakers are: Kurt Wüthrich, Marius Clore, David Neuhaus, Tom James, Walter Chazin, Oleg Jardetzky, Arthur Palmer, Randy Read, Gert Vriend, Alwyn Jones, Gerard Kleywegt, Jan Drenth, Hans Bartunik, Jack Johnson, Stephen Cusack, Keith Moffat, Richard Henderson, Jorge Navaza, and Peter Timmins.

The course fee is SEK 2500 for academic participants and SEK 5000 for nonacademic participants. The deadline for application is June 14, 1996, and the deadline for the poster abstract is August 2, 1996.

The Organizing Committee consists of Hans Hebert, Torleif Hard, Rudolf Ladenstein, and Lennart Nilsson.

For further information or application forms please contact Ms. Aila Holappa *via* e-mail **(Aila.Holappa@cbt.ki.se)** or fax (+46-8-608 9290), or you may obtain additional information from URL **http://www.csb.ki.se/events/summer96.html**.

PDB Moves within Brookhaven

During April, the PDB moved within Brookhaven National Laboratory from the Chemistry Department to the Biology Department. Our new mailing address is:

Protein Data Bank
Biology Department, Bldg. 463
Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973-5000
USA

In addition, our telephone exchange was recently changed from 282 to 344. Our main phone number is 516-344-3629 and the new help desk phone number is 516-344-6356. Our e-mail addresses, FTP address **(ftp.pdb.bnl.gov)**, and WWW home page **(http://www.pdb.bnl.gov)** have remained the same.

Letter to the Editor on Crystallographic Data Deposition

Reprinted from the January 1996 Quarterly Newsletter. The following "Letter to the Editor" has been sent to journals in which X-ray crystallographic structure determinations of macromolecules are published.

A formal discussion of the archival journal requirements for data deposition was held at the November 1995 International Seminar-cum-School on Macromolecular Crystallographic Data in Calcutta, India.

The current policy of the International Union of Crystallography (IUCr) is that upon publication of a crystal structure determination of a macromolecule, the atomic parameters used or represented in the publication must be deposited in the Brookhaven Protein Data Bank. The deposition of structure amplitudes is recommended but not required. The policy provides crystallographers with the option of delaying the release of atomic parameters for one year and of structure amplitudes for up to four years from the date of publication. Participants strongly supported this policy and felt it should be strictly applied by the journals (referees).

Recent developments in X-ray crystallographic experimental and refinement techniques and the huge expansion in computing power and networking, however, necessitate the review of deposition arrangements.

It was noted that the new validation procedures are much more effective but require the experimental structure amplitudes as well as the atomic parameters. In addition, the technical arrangements for deposition, analysis, and validation of macromolecular crystal structures are now much easier.

The undersigned consider it vital for the macromolecular crystallographers to respond to these developments in their deposition practices. We recommend, therefore, that publication of macromolecular crystal structures should be accompanied by deposition of atomic parameters and also structure amplitudes. Amongst the many reasons identified for this practice, the following two are critical:

- Rigorous validation of the structure determination results can only be carried out using both atomic parameters and experimental structure amplitudes. It is important that journals ensure that referees have sufficient information to prevent incorrect structures being published.

- Archiving of this data will ensure they are not lost. There were numerous reports at this Meeting of data being lost. This most probably reflects a general problem in the crystallographic community.

Edward N. Baker - Member of IUCr Executive Committee and Member of the IUCr Commission on Biological Molecules
Department of Chemistry and Biochemistry
Massey University
Palmerston North
New Zealand

Tom L. Blundell
ICRF Unit of Structural Molecular Biology
Department of Crystallography
Birkbeck College
Malet Street
London, WC1E 7HX
England

Mamannamana Vijayan - Chairman of IUCr Commission on Biological Molecules
Molecular Biophysics Unit
Indian Institute of Science
Bangalore 560012
India

Eleanor Dodson - Member of IUCr Electronic Publishing Committee
Department of Chemistry
University of York
York, YO1 5DD
England

Guy Dodson - Previous Chairman of IUCr Commission on Biological Molecules
Department of Chemistry
University of York
York, YO1 5DD
England

Gary L. Gilliland, Associate Director
Center for Advanced Research in Biotechnology
9600 Gudelsky Drive
Rockville, MD 20850
USA

Joel L. Sussman - Head, Protein Data Bank
Departments of Biology and Chemistry
Brookhaven National Laboratory
Upton, NY 11973
USA
and
Department of Structural Biology
Weizmann Institute of Science
Rehovot 76100
Israel

Current PDB Submission Procedures and Requirements

Reprinted from the January 1996 Quarterly Newsletter.

There are three essential elements of a complete PDB deposition:

- PDB-formatted coordinate data.

- A completed up-to-date version of our Electronic Deposition Form which you can pick up off of the PDB WWW home page **(http://www.pdb.bnl.gov)**, download from our FTP server **(ftp.pdb.bnl.gov)**, or request *via* e-mail from the PDB. Please fill out this Deposition Form using an on-line editor rather than by hand, as we run a program on the form to generate a preliminary header for your final PDB entry. As our older, non-electronic Deposition Forms are more difficult to fill out and because those forms are not as complete, we request that you no longer send the old forms to us. Our new Form has been available for about one and a half years, and we prefer that you use it.

- Copies of all relevant preprints and reprints OR a copy (print-out) of your submitted manuscript with the PDB Tracking Number it relates to prominently noted. As questions frequently arise about this, we would like to stress that IF YOUR PAPER IS NOT AT THE PREPRINT STAGE YET, a regular xerox or print-out of the submitted manuscript fulfills this requirement. If there is no manuscript in progress, please indicate so in the Journal (JRNL) section of the Deposition Form and the print requirement will be waived.

  Referenced papers can be sent electronically using FTP or e-mail, by fax to 516-344-5751, or by postal mail to:

  Protein Data Bank Depositions
  Biology Department, Bldg. 463
  Brookhaven National Laboratory
  P.O. Box 5000
  Upton, NY 11973-5000
  USA

  For depositors needing to obtain an ID code as quickly as possible, we suggest you e-mail or fax your manuscript.

  References are kept completely confidential and are used only to aid us with the processing of your entries and to ensure that we reference related papers correctly in the entries themselves.

The PDB now issues ID codes to depositors as soon as we receive the above three items. For more information on PDB submissions, please contact Minette Cummings at **pdb@bnl.gov**.

When sending replacement data to the PDB, please send an accompanying e-mail to **pdb@bnl.gov** indicating the related Tracking Number or ID code as well as a listing of all files sent.■

BROOKHAVEN ORDER FORM

Name of User	______________________	Date	______________
Organization	______________________	Phone	______________
Address	______________________	Fax	______________
	______________________	E-mail	______________
	______________________		

- ***Price is valid through September 30, 1996***
- ***Price is per CD-ROM set released --- releases occur four times per year***
- ***Facsimile and phone orders are not acceptable***

The Protein Data Bank ***MUST*** receive all three of the following items before shipment can be completed (please send all required items together *via* postal mail — facsimile and phone orders are ***NOT*** acceptable):

1. Completed order form;
2. Mailing label indicating exact shipping address;
3. Payment (using one of the two options below):

--> Check payable to Brookhaven National Laboratory in US dollars and drawn on a US bank. Foreign checks cannot be accepted and will be returned.

--> Original purchase order payable to Brookhaven National Laboratory. After your order is processed, you will be invoiced by Brookhaven National Laboratory. Please indicate exact address invoice should be sent to:

____

____

____

____

A wire transfer is acceptable only ***AFTER*** we have received an original purchase order from your organization and you have been invoiced by Brookhaven. After receiving Brookhaven's invoice, your bank may send a wire transfer to:

Bank name	:	Morgan Guaranty Trust Company of New York
Account name	:	Brookhaven National Laboratory
Account number	:	076-51-912

Please send all required items together *via* postal mail to:

Protein Data Bank Orders
Biology Department, Building 463
Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973-5000
USA

1 Protein Data Bank CD-ROM Set — ISO 9660 Format *$332.26*

(tax and shipping charges not applicable)

This page intentionally left blank.

AFFILIATED CENTERS

Twenty-four affiliated centers offer DATAPRTP information for distribution. These centers are members of the Protein Data Bank Service Association (PDBSA). Centers designated with an asterisk(*) may distribute DATAPRTP information both on-line and on magnetic or optical media; those without an asterisk are on-line distributors only.

BMERC
BioMolecular Engineering Research Center
College of Engineering, Boston University
Boston, Massachusetts
Nancy Sands (617-353-7123)
sands@darwin.bu.edu
http://bmerc-www.bu.edu/

BIRKBECK
Crystallography Department
Birkbeck College, University of London
London, United Kingdom
Alan Mills (44-171-6316810)
a.mills@cryst.bbk.ac.uk
http://www.cryst.bbk.ac.uk/PDB/pdb.html/

CAOS/CAMM
Dutch National Facility for Computer Assisted Chemistry
Nijmegen, The Netherlands
Jan Noordik (31-80-653386)
noordik@caos.caos.kun.nl
http://www.caos.kun.nl/

****CCDC***
Cambridge Crystallographic Data Centre
Cambridge, United Kingdom
David Watson (44-1223-336394)
watson@chemcrys.cam.ac.uk

CSC
CSC Scientific Computing Ltd.
Espoo, Finland
Heikki Lehvaslaiho (358-0-457-2076)
heikki.lehvaslaiho@csc.fi
http://www.csc.fi/

EMBL
European Molecular Biology Laboratory
Heidelberg, Germany
Hans Doebbeling (49-6221-387-247)
hans.doebbeling@embl-heidelberg.de
http://www.EMBL-Heidelberg.DE/

EMBL OUTSTATION: THE EUROPEAN BIOINFORMATICS INSTITUTE
Wellcome Trust Genome Campus
Hinxton, Cambridge, United Kingdom
Philip McNeil (44-1223-494-401)
mcneil@ebi.ac.uk
http://www.ebi.ac.uk

ICGEB
International Centre for Genetic Engineering and Biotechnology
Trieste, Italy
Sandor Pongor (39-40-3757300)
pongor@icgeb.trieste.it

IGBMC
Laboratory of Structural Biology
Strasbourg (Illkirch), France
Frederic Pewniak (33-8865-3273)
plewniak@igbmc.u-strasbg.fr
http://www-igbmc.u-strasbg.fr

****JAICI***
Japan Association for International Chemical Information
Tokyo, Japan
Hideaki Chihara (81-3-5978-3608)

****MAG***
Molecular Applications Group
Palo Alto, California
Margaret Radebold (415-846-3570)
bold@mag.com
http://www.mag.com

****MSI***
Molecular Simulations Inc.
San Diego, California
Mark Forster (619-458-9990)
mjf@msi.com
http://www.msi.com/

NCHC
National Center for High-Performance Computing
Hsinchu, Taiwan, ROC
Jyh-Shyong Ho (886-35-776085; ex: 342)
c00jsh00@nchc.gov.tw

NCSA
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
Champaign, Illinois
Patricia Carlson (217-244-0768)
pcarlson@ncsa.uiuc.edu

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION
National Library of Medicine
National Institutes of Health
Bethesda, Maryland
Stephen Bryant (301-496-2475)
bryant@ncbi.nlm.nih.gov
http://www.ncbi.nlm.nih.gov/

****OML***
Oxford Molecular Ltd.
Oxford, United Kingdom
Steve Gardner (44-1865-784600)
sgardner@oxmol.co.uk
http://www.oxmol.co.uk/

****OSAKA UNIVERSITY***
Institute for Protein Research
Osaka, Japan
Masami Kusunoki (81-6-879-8634)
kusunoki@protein.osaka-u.ac.jp

PEKING UNIVERSITY
Molecular Design Laboratory
Insititute of Physical Chemistry
Beijing 100871, China
Luhua Lai (86-10-2751490)
lai@ipc.pku.edu.cn
http://www.ipc.pku.edu.cn

PITTSBURGH SUPERCOMPUTING CENTER
Pittsburgh, Pennsylvania
Hugh Nicholas (412-268-4960)
nicholas@psc.edu
http://pscinfo.psc.edu/biomed/biomed.html/

SEQNET
Daresbury Laboratory
Warrington, United Kingdom
User Interface Group (44-1925-603351)
uig@daresbury.ac.uk

SAN DIEGO SUPERCOMPUTER CENTER
San Diego, California
Philip E. Bourne (619-534-8301)
bourne@sdsc.edu
http://www.sdsc.edu/SDSC/Staff/bourne/pb.html

****TRIPOS***
Tripos, Inc.
St. Louis, Missouri
Akbar Nayeem (314-647-1099; ex: 3224)
akbar@tripos.com

UPPSALA UNIVERSITY
Department of Molecular Biology
Uppsala University
Uppsala, Sweden
Alwyn Jones (46-18-174982)
alwyn@xray.bmc.uu.se

WEIZMANN INSTITUTE OF SCIENCE
Rehovot, Israel
Jaime Prilusky (972-8-9343456)
lsprilus@weizmann.weizmann.ac.il

Protein Data Bank
Biology Department, Bldg. 463
Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973-5000
USA

Statement of Support

PDB is supported by a combination of Federal Government Agency funds (work supported by the U.S. National Science Foundation; the U.S. Public Health Service, National Institutes of Health, National Center for Research Resources, National Institute of General Medical Sciences, and National Library of Medicine; and the U.S. Department of Energy under contract DE-AC02-76CH00016) and user fees.

Contacts

PDB Main Telephone 516-344-3629
PDB Help Desk 516-344-6356
Fax 516-344-5751

Internet:
help desk .... pdbhelp@bnl.gov
general correspondence ..... pdb@bnl.gov
depositions .... pdb@bnl.gov
order information .... orders@pdb.pdb.bnl.gov
network services .... sysadmin@pdb.pdb.bnl.gov
Listserver subscriptions ...... listserv@pdb.pdb.bnl.gov
Listserver postings .... pdb-l@pdb.pdb.bnl.gov
entry error reporting .... errata@pdb.pdb.bnl.gov

Staff

Joel L. Sussman, Head
Enrique E. Abola, Deputy Head and Sci. Coord.
Jaime Prilusky, Interim Head Database Dev.
David R. Stampf, Sr. Computer Analyst

Frances C. Bernstein
Judith A. Callaway
Minette Cummings
Betty R. Deroski
Pamela A. Esposito
Arthur Forman
Patricia A. Langdon
Michael D. Libeson
Nancy O. Manning
John E. McCarthy
Christine Metz
Regina K. Shea
Janet L. Sikora
Dejun Xue

Internet Sites

WWW .... http://www.pdb.bnl.gov
FTP .... ftp.pdb.bnl.gov

First Class