Protein Data Bank

Quarterly Newsletter

Release #73 *July 1995*

July 1995 PDB Release

3689 full-release atomic coordinate entries (242 new additions)

3428	proteins, enzymes, and viruses
251	nucleic acids
10	carbohydrates

365 structure factor entries
31 NMR experimental entries

The total size of the atomic coordinate entry database is 1273 Mbytes uncompressed.

Table of Contents

The latest version of the Electronic Deposition Form should be obtained from the FTP /**pub** directory before depositing data.

What is New at the PDB

The Protein Data Bank (PDB) at Brookhaven National Laboratory (BNL) is an archive of experimentally determined three dimensional structures of biological macromolecules, serving a global community of researchers, educators, and students. The challenges presented by the enormous growth in data over the past several years have been met by the PDB staff and management, which now provide an up-to-date archive while simultaneously expanding network access and building strong worldwide collaborations with related databases and resources. Even larger challenges are at hand, as the deposition rate continues to rise along with the expectations of the consumers of these data. We are enhancing the capabilities of the PDB and transforming it into a new biological Three-Dimensional Macromolecular Structure Database while at the same time retaining many of the features of the PDB, for compatibility purposes.

Over the course of the next few years, we plan to:

- Ensure that the archive remains current, accurate, and relevant;
- Provide a rapid and *painless* automated deposition system, using internationally agreed-upon standards;
- Enhance data validation tools;
- Build a relational database to store and provide flexible access to the information;
- Integrate the PDB with complementary biological and chemical databases through semantic links and schema sharing;
- Provide easy access, through the Internet, to the PDB at BNL and other deposition and distribution sites worldwide.

– Joel L. Sussman

PDB Browser Linked to BioMagResBank

The PDB and BioMagResBank (BMRB), a data bank of NMR chemical shift information for over 400 proteins and peptides, have initiated a joint collaboration. Users can now retrieve BMRB entries containing chemical shift information — corresponding to PDB entries for the identical structures obtained using the PDB's WWW browser — by means of hyperlinks inserted automatically into the retrieved PDB entries by the browser. In the future, BMRB entries additionally will contain proton-exchange, order parameter, coupling constant, and other kinetic and thermodynamic data derived by NMR spectroscopy from proteins, peptides, and nucleic acids. As the BMRB and PDB databases expand, additional links will be established between complementary sets of data. BMRB is located at the University of Wisconsin-Madison and is supported by a grant from the National Library of Medicine.

IUCr WWW Server

The International Union of Crystallography (IUCr) has a home page on WWW that can be reached with the following URL: **http://www.iucr.ac.uk/welcome.html**. It includes information about the Union, journals, CIF, forthcoming meetings, and pointers to other sources of information.

Of particular utility is the World Database of Crystallographers that can now be accessed on-line either through a form if your WWW server (e.g. Mosaic, Netscape) supports forms or by gopher. You no longer need to have a printed copy on your desk. This database is being updated regularly.

There is information on the IUCr home page on how to get the most out of all these services. If you don't have direct Internet access, send an e-mail message to **cif@iucr.ac.uk** consisting of the following: send howto.asc. Be sure absolutely nothing else appears in the body of the e-mail, not even your signature or the period after the word ‘asc’. You will receive a complete description of information services supplied by the Union and available through e-mail.

A Protein Motions Database

This article was written by Mark Gerstein, Department of Structural Biology, Stanford University, Stanford, CA, USA.

We have created a database describing protein motions, arranged around a multi-level classification scheme including domain motions, loop motions, and subunit motions. In addition to listing relevant proteins and descriptions of their motion, the database contains pictures of the motion and hypertext links to journal articles and molecular coordinates. It is accessible *via* Internet using WWW with the following URL: **http://hyper.stanford.edu/~mbg/ProtMotDB/**. It is also accessible *via* anonymous FTP with the following URL: **ftp://hyper.stanford.edu/pub/mbg/ProtMotDB/ProtMotDB.raw.txt**.

Using the database is very easy. You access its main page with one of the popular WWW browsers such as Mosaic or Netscape. You can type various keywords into this page and the database will display the entries for the relevant protein motions. Finally, you can click on various highlighted words in these entries, and this, in turn, will either link you directly to a relevant Medline abstract from the journal literature or bring up pictures and movies showing the protein motion.

[In order to view movies, pictures, postscript files, *etc.* you may have to build, and make your browser aware of, auxiliary applications. You can obtain information on how to do this for the two most popular browsers at **http://home.netscape.com/eng/mozilla/1.1/faq.html** and **http://www.ncsa.uiuc.edu/SDG/Software/XMosaic/mosaic-faq.html**, each of which can be easily accessed from the help menus of Netscape and Mosaic.]

The protein motions database should be of use to the community of biologists and biochemists because it collects information about many different kinds of motions in one place. This can greatly facilitate comparative analyses, for example. These data are particularly relevant at present because of the large and exponentially increasing number of solved protein structures (currently estimated at three thousand and increasing by about one per day). Furthermore, the graphical and interactive nature of the database is well-suited for the understanding of protein motions, which are often difficult to represent on a printed page. This is particularly important because many published papers about interesting protein motions do not precisely describe the relationship between the motion and specific publicly-accessible coordinate files and viewing orientations. For example, many papers do not tell you that the atomic coordinates for the open form are available as PDB entry 1ABC and those for the closed form are PDB entry 2ABC and that the motion is best viewed when looking down the crystallographic three-fold after fitting residues 5 to 90.

The database attempts to cover all instances of protein motion for which there is at least some structural information. The classification scheme is an expanded version of the one described previously; see M. Gerstein, A. Lesk, C. Chothia, Structural Mechanisms for Domain Movements in Proteins, *Biochemistry* **33**, 6739-6749 (1994).

- The basic division of items in this database is according to these three categories — loop and fragment motions, domain motions, and subunit motions. Within each category the motions are ranked according to the size of the entity that is moving.
- The next division is on the basis of available information such as whether structures of two conformations (e.g. open and closed) are known from crystallography or NMR, or is only one conformation known and the other suggested on the basis of other evidence?
- The final division is on the basis of mechanism. The motion is classified according to whether it involves rotation about distinct hinges or pivots, incremental sliding motions of different parts of the protein against each other (shear motions), partial refolding, or an allosteric transition.

Originally, the information associated with each protein was basic — its location in the classification described above, a brief text description of the motion, the PDB ident codes of the conformations (if available), and the references to the papers describing the motions. As the database project has progressed, more and more information has been added to each entry. Some entries, for instance, now have links to simple pictures and even movies (in MPEG format). The PDB ident codes associated with each entry have been made into hypertext links pointing directly to the structure entries at the Brookhaven PDB, to sequence and journal cross-references *via* the National Center for Biotechnology Information's (NCBI) Entrez database, and to a listing of related structures *via* the SCOP Database (see the January 1995 PDB Newsletter). Some of the references have also been directly linked to the corresponding Medline abstracts. The transformations and residue selections needed to optimally superimpose and orient each coordinate set to view the motion are provided with each entry. A WWW form linked to this database is provided for submitting entries.

For comments or additional information please contact Mark Gerstein at **mbg@hyper.stanford.edu.**

Announcement: Call for Targets for Protein Structure Prediction

IRBM (Istituto di Ricerche di Biologia Molecolare)
Practical Course: Frontiers of Protein Structure Prediction
October 8 - 17, 1995; Pomezia, Rome, Italy

Organizers: T. Hubbard (CPE, MRC), A. Tramontano (IRBM)

Instructors: G. Barton (Oxford), T. Hubbard (Cambridge), D. Jones (London), M. Sippl (Salzburg), A. Valencia (Madrid)

Lecturers: A. Lesk (Cambridge), J. Moult (Rockville), B. Rost (Heidelberg), C. Sander (Cambridge)

The aim of this workshop is to predict as much as possible about the structure of a number of proteins of biological interest, taking advantage of the most recent methodologies for fold recognition and *ab initio* prediction.

If you are interested in a structure prediction being made for a protein which does not appear to be homologous to any known structure and for which there is no sign of an experimental structure being determined, please consider submitting it as a target for this course. The predictions will be made public as a technical document and also available *via* WWW.

For further information and on-line target submission forms see **http://www.mrc-cpe.cam.ac.uk/predict/.**

Suggestions for Filling Out the Electronic Deposition Form

Since June 1994, the PDB has been accepting depositions on our Electronic Deposition Form (Dep Form). We are pleased with the high level of compliance and have been able to save time in processing those entries received on the new Dep Form. Here are a few suggestions that depositors should be aware of when completing the Form. These examples were chosen because they represent some of the most-often made mistakes or most frequently misunderstood instructions. If you have any questions while preparing your deposition, we will be happy to assist you.

PDB programs read the Electronic Dep Form and automatically convert the information into a PDB entry. By following the procedures described here, you will be helping to reduce processing time as the programs will produce a cleaner entry with the first pass. Please note that the entire process will be changing in the coming months as PDB institutes automatic submission procedures.

– *Features of Deposition Form*

• Dashed Lines

Dashes, as shown here,

```
--------------------------------------------------------
```

mark the start and end of each section. If you need to insert your own dashed lines, please insert spaces every 5 scores, as shown:

```
----- ----- ----- ----- ----- ----- ----- ----- ----- -----
```

• Instructions

Instructional text appears throughout the Dep Form. Please do not delete or edit any of the explanatory lines — they will be stripped out automatically when the entry is processed. The process utility, expecting these lines to be there, will get confused and may ruin the entry if any of these lines are deleted by the depositor.

• Keywords

Changes to the keyword (including case and spacing) will create problems for the parsing program. Please leave them in the submitted Dep Form, leaving the value blank if appropriate.

• Colon

The colon indicates the presence of a keyword: value pair (the information preceding the colon is the keyword; the information following the colon is the value). If you remove the colon, the value may be misinterpreted.

Correct

```
residue name     : ASN
chain identifier : A
```

Incorrect

```
residue name     : ASN
chain identifier A
```

The correct interpretation is residue ASN of chain A. The incorrect example is interpreted as residue "ASN CHAIN IDENTIFIER A" with NO chain identifier.

• **Repeat Block Indicators,** // \\ **and** \\ //

These slashes are used to indicate repeating blocks. If you put information on the line within the slashes, the information may be lost. Copy the entire block, including the slashes, when you need to provide repetitive data for the items contained within the block. For example:

```
//                    \\
   keyword_A : texta
   keyword_B : textb
 \\                   //


//                    \\
   keyword_A : textc
   keyword_B : textd
\\                    //
```

– NULL Values

When data are unavailable, not supplied, or irrelevant to your entry, leave the value field blank. Do not put comment text in these value fields. If you wish to present some explanation for null fields, you may do so as a REMARK record.

Correct

free R value : 0.021
free R value :

Incorrect

free R value : NO FREE R WAS CALCULATED
free R value : -----

– Author Names

Our programs expect last names to follow initials. Do not split a name across two lines.

Correct

author(s): J.L.Sussman,S.R.Holbrook,R.W.Warrant,
G.M.Church, S.-H.Kim

Incorrect

author(s): JL Sussman,SR Holbrook,RW Warrant,GM
Church,S-H Kim.

– Residue Name and Chain Identifier

One-character, three-character, or full residue names may be given in either upper or lower case. Multiple residue names are acceptable in the SEQUENCE section. Please do not use symbols or words which are difficult for programs to interpret.

Examples of how to provide residue name and chain identifier:

Correct

residue name : N
chain identifier :

residue name : ASN
chain identifier : A

residue name : adenosine
chain identifier : A

residue name : GLY SER ASN GLY MET
chain identifier : A

residue name : GSNGM
chain identifier : A

Incorrect

residue name : ASN A
chain identifier :

residue name : GLY 27 A--> MET 31 A
chain identifier :

residue name : from GLY 27 A TO MET 31 A
chain identifier :

– Sequence Number and Insertion Code

Sequence numbers are numeric. Insertion codes are usually letters, although, they are currently defined as alphanumeric. More than one sequence number per line is allowed, however, do not include the residue name, chain identifier, or insertion code on the same line as the sequence number.

Examples of how to provide sequence numbers and insertion codes:

Correct

sequence number : 84
insertion code : A

sequence number : 84,85,86
insertion code : A

sequence number : 84 - 120
insertion code :

sequence number : 84 - 120, 130 - 150
insertion code :

Incorrect

sequence number : thr 84
insertion code :

sequence number : 84A
insertion code :

– Continued Lines

There is no limitation on the length of a line. Text following a carriage return continues onto the next line with no indentation required. Do not repeat the keyword or the colon.

Correct

remark: This model is one of two independently refined models of the reduced SOD structure. Comparison of the two models was used to analyze coordinate errors and water structure. The other model, as well as the final model, obtained after combining the two corresponding structure factor data sets, can also be found in this database.

Incorrect

remark: This model is one of two independently refined models
remark: of the reduced SOD structure. Comparison of the two
remark: models was used to analyze coordinate errors and
remark: water structure. The other model, as well as the final
remark: model, obtained after combining the two correspond-
remark: ing structure factor data sets, can also be found in this
remark: database.

Incorrect

remark: This model is one of two independently refined models
of the reduced SOD structure. Comparison of the two
models was used to analyze coordinate errors and
water structure. The other model, as well as the final
model, obtained after combining the two correspond-
ing structure factor data sets, can also be found in this
database.

– Repeating Blocks

Please edit the repeating block to include the indicators // \\ and all keywords, as in this example. Do not repeat only a portion of the repeating block.

Example of how to repeat blocks:

```
TERMINAL RESIDUE (TER)

//                          \\
residue name        : LYS
chain identifier    : P
sequence number     : 18
insertion code      :
\\                          //

//                          \\
residue name        : CYS
chain identifier    : L
sequence number     : 219
insertion code      :
\\                          //

//                          \\
residue name        : ARG
chain identifier    : H
sequence number     : 220
insertion code      :
\\                          //
```

– PDB Entries

Correctly-formatted records from a PDB entry may be inserted into any of the relevant sections of the Dep Form instead of providing the requested values. These must remain intact as PDB records, and appear at the bottom of the section, as in this example. Often this will speed up the process of filling out the Dep Form.

Example of how to insert correctly-formatted PDB records:

```
HELIX   (optional, see box above)
        Give inclusive pairs of residues to define each helical
        substructure and indicate the kind of helix found in
        each case. Place numbers/letters within brackets.
        Repeat this block for each helix.

        (text omitted)

  determination method :

  //                                    \\
  helix identifier :
  helix class :
  remark :

  initial     chain id [ ] seq number [ ] insert code [ ]
  terminal    chain id [ ] seq number [ ] insert code [ ]
  \\                                    //

  HELIX  1  1  ALA  L   80  ASP  L   82  5
  HELIX  2  2  SER  L  122  THR  L  126  1
  HELIX  3  3  LYS  L  183  GLU  L  187  1
  HELIX  4  4  PHE  H   29  ASN  H   31  5
  HELIX  5  5  LYS  H   73  SER  H   75  5
  HELIX  6  6  SER  H   84  ASP  H   86  5
  HELIX  7  7  ASN  H  162  GLY  H  164  5
  HELIX  8  8  PRO  H  213  SER  H  215  5
```

Good-bye to John Skora

We would like to sincerely thank John Skora for his five years of dedication and hard work with the PDB as our Systems Manager. John has decided to leave the PDB for another position. Our very best wishes to him in the future.

John McCarthy will be taking over John Skora's responsibilities, and we are confident that he will continue to provide reliable service. Mr. McCarthy has been with the PDB for five years and looks forward to his new responsibilities. Questions previously directed to John Skora should now go to John McCarthy at **mccarthy@bnl.gov.**

The *Portable* ***WWW Browser***

The WWW version of the PDB browser has met with success in many areas of the networked community, with roughly 1,800 accesses per week from around the world. When it was designed, a modular approach was used that allowed for the separation of the user interface, search engine, and file access functions in order for us to adapt quickly to advances in each of those three areas. As a side benefit, this also allowed the browser code to be portable, so that those who have poor network access to the computers at the PDB may have the search engine and file service parts of the browser installed locally and thereby improve access to the PDB archive. This article describes the portable browser. Detailed installation instructions may be found on-line from the PDB.

Since this version of the browser uses the WWW, Dave Stampf, in collaboration with Jaime Prilusky of the Bioinformatics Unit of the Weizmann Institute, set it up to have hyperlinks to the E. C. Database and Entrez Reference Database added automatically to the PDB html file as it is being transmitted to the user's viewer program. This valuable information is thus a mouse click away, permitting an expansion of the research possibilities afforded the user.

In order to run the WWW browser locally, you have to be running an http server. (This is the communications protocol underlying the WWW.) For information on how to do this, please see **http://hoohoo.ncsa.uiuc.edu/docs/Overview.html** for one possible http server. You must also have Perl installed on your system. Once you have a system with a functional http server and Perl, you can install the WWW browser in layers.

– *The Browser Code*

The first layer is a Perl script and its associated help file. This script creates the form with which the user works and constructs, but does not execute, the queries. This is easy to install and uses few system resources.

In order to perform the installation, you need to edit between one and five lines in the Perl script called browse.pl. The lines specify the name and location of the Perl script, the name and location of the help file, the machine and port of the query server (second layer) and the name of the script that retrieves and marks up the files (third layer).

– *The Search Engine*

The second layer consists of a second Perl script and a number of index files. Together, these actually execute the user's query, returning the ident code and perhaps a line of text describing the entry, but not the entry itself. This layer requires much more disk space and some CPU cycles as well, but it is primarily I/O bound. The current version of this search engine uses UNIX dbm files, but we are currently testing the use of SYBASE to perform the same (and more complex!) searches.

This Perl script accepts incoming requests on an unprivileged port and then forks a copy of itself to handle the request and awaits another request. In order to install this part you must have an up-to-date set of the browser's index files installed (and, preferably, also mirrored) on your local machine. You must also tell the Perl script the location of these indices in your directory tree. The index files require less than 18 Mbytes of disk space as of April 1, 1995. These files are available from the PDB in compressed format, may be automatically mirrored, and are identical to the index files used in the tk/tcl version of the browser.

– *File Service*

The third and final layer consists of a small number of Perl scripts, together with a few more index files, and requires all of the PDB archive to be on-line locally. Many locations will find this requirement unworkable — those sites can simply install layer one, or layers one and two. Our experience indicates that not all searches culminate in the transfer of a PDB file, so installation of the first two layers may suffice.

To install the third layer, you simply need to identify the location of the files on your system within the Perl scripts and make sure that the index files are installed.

All of the files necessary to install the portable WWW browser, together with more detailed instructions, are contained in **ftp://ftp.pdb.bnl.gov/pub/pdbbrowse/WWWBrowse/INSTALL**. We welcome bug reports and suggestions. Please send these to Dave Stampf **(drs@bnl.gov)**.

Accessing Individual PDB Entries via the WWW

The PDB is providing easy access to all entries using the URL mechanism of the WWW for developers creating new hypertext links to the PDB archive, looking for a faster way to access individual entries, or experimenting with programs written in Fortran, C, or Perl who would like access to entries without having the entire archive locally mounted.

Every entry contained in the PDB is now available *via* the URL

http://www.pdb.bnl.gov/cgi-bin/get-pdb-entry?id=xxxx

where xxxx is the usual PDB ident code (*e.g.*, 1abc). In addition, if your network links are slow, you may wish to try the variation

http://www.pdb.bnl.gov/cgi-bin/get-pdb-entry?id=xxxx&encoding=compress

which will send the file to you in UNIX compressed format, typically saving two-thirds of the network bandwidth.

The first URL should work from any "reasonable" WWW browser. The second is successful on most recent versions of WWW browsers, but if you have a simpler browser or an earlier version of the more complete browsers, you may have to save and uncompress the data in a separate command. Of course if you use either of the above URL's from a UNIX command line, please be sure to quote the string since '?' and '&' are usually special characters in the shell.

Finally, one more parameter is available for either of the lines. If you append the string "&type=view," then the file (either compressed or uncompressed) is transmitted with the MIME Content-type = "chemical/x-pdb." In a manner similar to the action of the PDB WWW browser, this will transmit the file and start the external viewer (*e.g.* RasMol) to display the structure.

In order to distribute the network load evenly around the globe, we will make these scripts available to all PDB redistributors running an http server. We will announce availability of the scripts on the PDB's listserver **(pdb-l@pdb.pdb.bnl.gov)**.

Thanks are due to Steve Brenner and Jaime Prilusky who originally asked that we make this public and provided some of the ideas for options. We would be happy to consider any other extensions to this scheme. Please contact Dave Stampf (**drs@bnl.gov**) with your ideas.

3DBbase – A Relational Database

The PDB is in the process of building a relational database to manage the archive and to give its users a more powerful engine to browse its contents. The schema is presented on PDB's WWW server, and at this time the database that it represents is being made available to the community for evaluation and testing. Access to the data stored in the database will be through PDB's WWW browser.

This database development work is a collaboration that includes the following groups:

The Protein Data Bank, Brookhaven National Laboratory
Bioinformatics Unit, Weizmann Institute of Science
Data Management Tools Group, Lawrence Berkeley Laboratory
The Genome Data Base, Johns Hopkins University

The schema was generated with OPM Schema Editor 3.0, developed by the Data Management Tools Group, Lawrence Berkeley Laboratory.

The database is being implemented on a SYBASE engine. In addition to all coordinate entries found in the PDB, the PDB database includes semantic links to entries found in other biological databases. The first steps in the formation of a federation of biological databases are also being made possible by this work. All PDB database bibliographic citations are stored and maintained on GDB's citation database (CitDB), also built using OPM.

We plan to have a fully operational database system by October 1995. In the interim, all are welcome to test our system, study our schema, and suggest changes and extensions. Further documentation providing details and explanations of the objects presented in the schema, as well as our development plans, will be released in the coming weeks.

The schema can be accessed through the PDB WWW home page **(http://www.pdb.bnl.gov)**.

Access to PDB

– World Wide Web (WWW)

PDB has a WWW server on the computer system **www.pdb.bnl.gov** (Internet address 130.199.144.1). This server is accessible using the URL **http://www.pdb.bnl.gov/**.

Besides including links to the PDB FTP and Gopher servers, the WWW server includes links to many other useful databases and information servers.

– Gopher

PDB has a Gopher server on the system **gopher.pdb.bnl.gov** (Internet address 130.199.144.1). This server is accessible using a Gopher client connecting to the following link:

Name	=	Protein Data Bank FTP server
Type	=	1
Host	=	gopher.pdb.bnl.gov
Port	=	70
Path	=	1/

As a Gopher client, you may navigate through a hierarchy of directories and documents or ask an index server to return a list of all documents that contain one or more specified words. For instance, you can choose "The PDB Anonymous FTP" after reaching PDB's Gopher server in order to search and download the same information and coordinate files as through FTP. Alternatively, you can select "An (almost) full-text search of the PDB Bibliographic Headers" in order to search PDB using any keyword.

– FTP

PDB has an anonymous FTP account on the computer system **ftp.pdb.bnl.gov** (Internet address 130.199.144.1). Files may be transferred to and from this system using anonymous as the FTP user name and your e-mail address as the password. Besides downloading entries, data files, and documentation, it is possible to upload any files that you may wish to send to PDB, into the special directory /**new_uploads**. Those using VMS may need to place quotes around file names.

– Listserv

PDB has a mailing list devoted to discussions concerning its operation, contents, and access procedures.

To subscribe, send e-mail to **listserv@pdb.pdb.bnl.gov** with the one-line message: **subscribe PDB-L Firstname Lastname**.

To find out what can be done with this mailing list, send e-mail to the same address (**listserv@pdb.pdb.bnl.gov**) with the one-line message: **help**.

To send a message to all PDB-L subscribers, e-mail the message to: **PDB-L@pdb.pdb.bnl.gov**.

This page intentionally left blank.

BROOKHAVEN ORDER FORM

Name of User ______________________________ Date ______________

Organization ______________________________ Phone ______________

Address ______________________________ Fax ______________

______________________________ E-mail ______________

____

- ***Price is valid through September 30, 1995***
- ***Price shown is per release - there are four releases 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 U.S. dollars and drawn on a U.S. 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.

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 three required items together *via* postal mail to:

Protein Data Bank Orders
Chemistry Department, Building 555
Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973-5000
USA

Protein Data Bank CD-ROM — ISO 9660 Format *$300.00*

(tax and shipping charges not applicable)

This page intentionally left blank.

AFFILIATED CENTERS

Twenty-two 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/

*BIOSYM
BIOSYM Technologies, Inc.
San Diego, California
Rick Lee (619-546-5536)
rickl@biosym.com
http://www.biosym.com/

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/

CAN/SND
Canadian Scientific Numeric Data Base Service
Ottawa, Ontario, Canada
Roger Gough (613-993-3294)
cansnd@vm.nrc.ca

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)
dgw1@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/

CINECA
NE Italy Interuniversity Computing Center
Casalecchio di Reno (BO), Italy
Laura Setti (39-51-6599478)
asltc0@icineca.cineca.it

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

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

INN
Israeli National Node
Weizmann Institute of Science
Rehovot, Israel
Leon Esterman (972-8-343934)
lsestern@weizmann.weizmann.ac.il

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

*MAG
Molecular Applications Group
Palo Alto, California
Hilary Jensen (415-473-3039)
hilary@suerte.mag.com
http://hyper.stanford.edu/~Mag/

*MSI
Molecular Simulations Inc.
Burlington, Massachusetts
Lance J. Ransom Wright (617-229-9800)
lance@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
Yoshiki Matsuura (81-6-879-8605)
matsuura@protein.osaka-u.ac.jp

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

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

Protein Data Bank
Chemistry Department, Bldg. 555
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

Telephone 516-282-3629
Facsimile 516-282-5751

Internet:

pdb@bnl.gov .... general correspondence
orders@pdb.pdb.bnl.gov .... order information
sysadmin@pdb.pdb.bnl.gov .... network services
listserv@pdb.pdb.bnl.gov .... Listserver subscriptions
pdb-l@pdb.pdb.bnl.gov .... Listserver postings
errata@pdb.pdb.bnl.gov .... entry error reporting

Please include your name, postal mailing address, e-mail address, facsimile

PDB Staff

Joel L. Sussman, Head
David R. Stampf, Sr. Project Mgr.
Enrique E. Abola, Science Coordinator
Jaime Prilusky, Interim Head Database Dev.

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
Regina K. Shea
Karen E. Smith
Dejun Xue

First Class