Protein Data Bank pdb™

Quarterly Newsletter published by Brookhaven National Laboratory Protein Data Bank ™

Release #82 **October 1997** **Brookhaven National Laboratory**

October 1997 CD-ROM Release

6491 Released Atomic Coordinate Entries

Molecule Type

5740	proteins, peptides, and viruses
253	protein/nucleic acid complexes
486	nucleic acids
12	carbohydrates

Experimental Technique

163	theoretical modeling
1017	NMR
5311	diffraction and other
1468	Structure Factor Files
355	NMR Restraint Files

The total size of the atomic coordinate entry database is 2.8 GB uncompressed.

Table of Contents

Internet Sites

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

What's New at the PDB

— Joel L. Sussman

There has been an incredibly rapid increase in the rate of determination of 3D structures of biomacromolecules, as reflected by a new structure being deposited at the PDB, on average, every five hours. Accompanying this growth in structure deposition is an increasing proportion of depositors who are taking advantage of the PDB policy to allow structures to be kept 'on hold' up to a year after coordinate deposition. We feel it is now time to reconsider whether such a policy is still appropriate.

PDB's release policy currently states that:

> "...coordinates may be held (before release) no longer than one (1) year and structure factors [and NMR restraint data] may be held no longer than four (4) years from the date of publication... Requests that the PDB delay release of your data (put it on hold) should be submitted at the time of the initial deposition."

This policy is based on rules drawn up by the International Union of Crystallography (IUCr) in the late 1980's, for papers published in IUCr journals, which is also the basis of the policies of most other scientific journals and of a number of government funding agencies for work undertaken on grants.

We at the PDB constantly have people contacting us to ask why a particular structure is not yet available. Many depositors themselves have been confronted by the inherent difficulties of describing 3D concepts in the limitation of words and pictures on a 2D page. Being able to interactively view the molecule in three dimensions greatly enhances a reader's comprehension and appreciation of the accomplishment of a newly determined structure.

Alexander Wlodower recently argued that reasons given for putting coordinates 'on hold' were largely outdated and irrelevant (*Nat. Struc. Biol.* **4**, 173-174, 1997), giving clear examples of how the availability of coordinates has been of benefit to all. Wlodower concluded:

> " ...early distribution of the coordinates, even to competing laboratories, has led to great acceleration of investigations and ultimately benefited the original team."

We are suggesting that the 'on hold' option be limited until the date of publication, bringing the structural biology community in line with almost every other discipline in having to give free and immediate access to the data which they are using to support their claims.

We would greatly appreciate comments from the scientific community on this suggested change in policy, and look forward to implementing it as soon as possible.

Archive Management

— Enrique Abola and Nancy Manning

— Proposed Mandatory Items

In August, the PDB released a list of proposed mandatory items **(http://www.pdb.bnl.gov)**, data items that will be required from depositors before a PDB ID code can be issued. The list was assembled to implement a fast track release of data using a layered approach. What we outline are those items that we believe are necessary to minimally describe an entry. The list specifies the data items and does not address issues related to implementation or data validation, and it also does not specify expected value (or value ranges) for these data items. These details are being worked out, and the PDB encourages your comments on this issue.

We are currently developing new protocols and policies to allow us to release and load entries on our server by an automatic procedure. This fast track release will depend on what we call a layered approach in which data are released within a day of submission and are subsequently updated after standardization (e.g., making data representation consistent with other entries) by PDB staff with assistance from the depositors.

In the layered approach, a PDB ID code will be issued only after a depositor gives approval to release immediately his/her entry. In the case of data on hold, depositors will receive a PDB ID code only after agreeing to the release of their entry on the hold expiration date. Thus there will be no backlog for the initial release of entries, as entries will be made available on the same day that they are given a PDB ID code. As authors will have the responsibility of giving the 'green light' to the release of their entries, it is likely that they will make sure that they are as accurate as possible.

The items of the list are a subset of those given in the Deposition Form. Presentation in AutoDep will be implemented once the list is finalized. We are now in the process of preparing a similar document which will outline validation criteria that will be applied to entries before they are accepted using the layered approach. There are a number of items to be pointed out:

- We expect that most depositors will make use of one of the refinement programs for which the PDB has specified recommended tokens to be submitted. The mandatory list is of a 'minimal subset' required for the PDB to process the data before issuing a PDB ID code.
- For MAD data, it would be best to get details for all wavelengths used, but since this mandatory list is the minimal set required, we are asking for details corresponding to the data set used for refinement. This may be modified at a later date.
- NMR entries are not yet well handled in this list. This is particularly true for describing the results of the refinement. The PDB is working with John Markley and Eldon Ulrich at the BMRB to improve the coverage for NMR experiments (see related article in this Newsletter).

Each token in the list of proposed mandatory items is described by six fields:

1. item No - Each item is assigned a number; item numbers with alphabetic extensions are for conditional_mandatory records.
2. status - This is either mandatory or conditional_mandatory. Conditional_mandatory items become mandatory depending on the answer to the previous mandatory item.
3. item - This is the name of the item to be requested.
4. Dep - Section in Electronic Deposition Form V. 3.30 in which the item is requested and described. Please note that the Deposition Form is divided into sections and subsections, each with a title. We include in this field the section title followed by the subsection title within parentheses.
5. Guide - Section in the PDB Contents Guide in which the field is described. This will normally be in the form of PDB record types. In some cases we include the field name in parentheses.
6. Descr - Description of the item whenever necessary.

Mandatory records are classified into four groups:

1. Administrative details - These include depositor information and special instructions.
2. Data for validation - Crystal parameters, space group information, scale matrix, coordinates, etc. are included in this group and are needed for validation of the entry.
3. Data for the experimental model - Compound, source, sequence, site, resolution range, R-value, free R-value, RMSD's, etc. will help the user in understanding and assessing the experimental model.
4. Experimental Data - Experiment type, resolution range, completeness of data, R-merge or R-sym, etc. will help the user in understanding the experiment on which the model is based.

Two examples of items in the list follow.

item No: 4
status: mandatory
item: e-mail (Y/N)
Dep: Depositor Information (Primary Contact)
Guide:
Descr: Question on whether the depositor has an e-mail address. We will require either a Y or N answer. N/A is not allowed. If yes we will require the e-mail address.

item No: 4a
status: conditional_mandatory
item: e-mail address
Dep: Depositor Information (Primary Contact)
Guide:
Descr:

item No: 55
status: **mandatory**
item: non-standard space group setting or origin choice (Y/N)
Dep: CRYSTAL AND COORDINATE SYSTEM
Guide:
Descr:

item No: 55a
status: conditional_mandatory
item: setting
Dep:
Guide: REMARK 285
Descr:

item No: 55b
status: conditional_mandatory
item: space group operators
Dep:
Guide: REMARK 285
Descr:

— AutoDep 2 Released

The PDB is pleased to announce the release of AutoDep version 2.0. This is an improved version of our Web-based submission procedure, with many new features and extended validation added. AutoDep 2.0 is designed to move toward the Layered Release, discussed above.

We wish to acknowledge the valuable contribution to this effort made by our beta testers, our collaborators at the EMBL Outstation: The European Bioinformatics Institute (EBI), Kim Henrick and Peter Keller, our collaborators at the BioMagResBank (BMRB), John Markley and Eldon Ulrich, and the many users of AutoDep 1 who have taken the time to send us their suggestions and bug reports.

The new features of AutoDep 2.0 include:

- A Control Panel allows faster navigation through the process than was possible in the original AutoDep, and the user is led through the submission procedure as AutoDep "travels" to each new section to be completed.
- A function is included to look up the user's contact information in our database. If found, AutoDep completes this information automatically.
- Sending files is much easier through the use of a Web form with a file browser for selecting and uploading files to the PDB.
- The helps have been improved. Requesting "help" on one particular question will return information only on that subject.
- Non-standard space groups are handled properly.
- After the submission is complete, PDB validates the entry by running our internal validation programs as well as the WHAT IF program of R. Hooft and C. Sander, as implemented in collaboration with the PDB **(http://www.sander.embl-heidelberg.de/rob/checkhelp/)**. Checks include symmetry and geometry checks, and the results are compared with those obtained for other structures with similar resolution in the PDB. The author reviews these diagnostics before completing the submission. (See the April 1997 Newsletter for more on WHAT CHECK)
- AutoDep 2.0 was programmed to make it portable. In the near future EBI will begin accepting AutoDep submissions, making it even easier for our users in Europe to submit their structures. The data will be forwarded to PDB for processing and inclusion into the archives. This will be followed by satellite submission sites at the Weizmann Institute in Rehovot, Israel, and Osaka University in Osaka, Japan.

All new submissions will use AutoDep 2.0. However, submissions already begun with AutoDep 1 will be completed with the earlier interface. This will be handled automatically by the program just by "continuing" an AutoDep session of a BNL-number and password initiated in AutoDep 1. If you prefer to try AutoDep 2.0 instead, use the "Based on a previous submission" function to start a new session based on the uncompleted AutoDep 1 session. After two months, all submissions will be directed to AutoDep 2.0.

Please see the Release Notes **(http://www.pdb.bnl.gov/release-notes.html)** for more details. If you have any questions about either AutoDep 2.0 or AutoDep 1, please contact the PDB Help Desk at pdbhelp@bnl.gov or 516-344-6356.

PDB Computer Services

— John McCarthy

— PDB Internet Services Moved

The PDB's FTP/WWW/Gopher/Mailing List services were recently moved from a Silicon Graphics, Inc. Power Series 4D/480S with eight 40 MHz 32 bit processor computer system with 256 MB of main memory to a Silicon Graphics, Inc. Power Challenge L with four 90 MHz 64 bit processors and 512 MB of main memory. This new computer system is connected to the Internet *via* a 100 Mbit/second FDDI connection, ten times faster than the connection from the old computer system. It also has increased disk capacity to store all the entries being released by the PDB.

This move will allow for faster index searches using the PDB's Web browsing tools and faster file retrieval using our anonymous FTP service.

— PDB CD-ROM Files Compression

The current PDB CD-ROM set consists of six CDs. In the near future it will require seven or more CDs if an alternative to current storage methods is not found. Having seven or more CDs per set would require storing one set in multiple cases and requires more CD swapping to access the entire database.

After investigating alternatives, the PDB plans to compress files on its quarterly CD-ROM releases beginning with the October 1997 release. At first just the structure factor files will be compressed, with coordinate entries being compressed in the future.

The PDB is planning on using the Gnu gzip package to perform the compression and will distribute the Gnu Gunzip package in the CD-ROM set to allow CD-ROM users to perform the uncompression. The Gunzip package is free and the source code, as well as executable files ready to run on some of the more popular platforms, will be included in the CD-ROM set.

An effect of compression, since the CD-ROM set is written in ISO-9660 format requiring filenames in the DOS 8.3 format, will be that the filenames will be different. Files that had the PDB .ent suffix will have the .gz suffix when compressed. The PC-based browser PDB-Shell has been updated to be able to read compressed entry files.

A questionnaire has been sent to all users receiving the July 1997 CD-ROM requesting their views in this matter. A similar questionnaire was sent out in the summer of 1995 with most respondents very much in favor of compression.

If you would like to receive the questionnaire or wish to comment on this, please send e-mail to sysadmin@pdb.pdb.bnl.gov.

BMRB - PDB Collaboration Update

— Eldon Ulrich, BMRB, University of Wisconsin, Madison, WI, USA (elu@bmrb.wisc.edu) and Enrique Abola, PDB.

The BioMagResBank (BMRB, **http://www.bmrb.wisc.edu**) and the PDB are collaborating to develop efficient and standardized procedures for depositing and exchanging three-dimensional structural data derived by NMR spectroscopy. Our first step will be evident in the next version of the PDB AutoDep system (see related article). NMR spectroscopists will be given clear instructions concerning the types of data that can be deposited. These data include not only the calculated atomic coordinates, but also the constraints used in the calculations, statistics for the coordinates and constraints, and the assigned chemical shifts for the molecule(s) studied. The information requested will follow the recommendations of the IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy.

Processing and validation of PDB NMR entries will continue to be carried out at PDB at this time. PDB has, however, begun to transfer submitted NMR data to BMRB. BMRB staff will be contacting authors who have submitted structural data to PDB without a full list of chemical shift assignments and reference information to ask if they would be willing to deposit these data at BMRB. In addition, we are developing a standard format for NMR constraint data, procedures for validating the calculated atomic coordinates with the submitted constraint data, and software to convert constraints between commonly used input formats and the standard format. This work is being carried out with the help of Helen Berman and John Westbrook of the Nucleic Acid Database and Ton Rullmann, Robert Kaptein, and others involved in the European Structure Validation Project. Implementation of the complete system is expected to take several months.

BNL Contract Up for Bid; Announcement Expected in November

— Kara Villamil, Public Affairs Office, Brookhaven National Laboratory, Upton, NY, USA.

After a roller coaster ride of a year, the entire Brookhaven National Laboratory community is anxiously awaiting the announcement, expected in mid-November, of a new contractor to operate BNL for the U.S. Department of Energy.

The fate of BNL itself, and by extension the PDB, is not in question. But the issue of who will run the Laboratory, beginning in January 1998, is.

The current situation arose beginning in May 1997, when U.S. Secretary of Energy Federico Peña told the Lab's contractor of 50 years, Associated Universities Inc., that its current contract with DOE would be terminated, and a new bidding process begun, nearly two years earlier than planned.

Peña's announcement came on the heels of a DOE report detailing environmental management deficiencies at the Lab and within DOE itself. The report was commissioned following the discovery in mid-January of low-level radioactive groundwater contamination emanating from the spent fuel pool at BNL's main research reactor. The public and political uproar over that contamination, and the Lab's other environmental problems, reached a fever pitch in the spring and is still reverberating as the DOE weighs whether the reactor should be allowed to operate again.

— The Road to a New Contract

Peña's announcement of the contract's termination set in motion the creation of a Source Evaluation Board, based out of DOE's Chicago Area Office, and the solicitation of input from BNL employees and the local community on the desired attributes of the next contractor.

Among the main requirements set out in the DOE's resulting Request for Proposals was that the new contractor be non-profit, or a consortium led by a non-profit entity; that it maintain the mission and employment level of the Lab without substantial changes to employee benefits and union contracts; that it improve environmental management at BNL; and that it lay out plans for improved relations with the local community to repair the rift left by months of news about the groundwater contamination and other issues.

The next step was for potential bidders to get to know BNL, a process that brought dozens of representatives from academic and industrial institutions to Long Island for tours and information sessions.

— A Tumultuous Summer

Speculation about the prospective bidding teams, and the chances that AUI would enter the race to try to reclaim its contract, ran wild until midsummer, when the field had narrowed to two competing teams that had made their bids public.

The State University of New York at Stony Brook, located less than 20 miles from BNL and with a long history of collaboration with the Lab, announced on July 28 that it had entered an agreement with the Battelle Memorial Institute, a non-profit research foundation based in Ohio that also runs the Pacific Northwest National Laboratory for DOE.

Then, on August 1, another New York institution, Rensselaer Polytechnic Institute (RPI), announced it would join with Westinghouse Corp. of Pittsburgh to bid. Just four days later, AUI announced it would not attempt a bid for the contract, with President Lyle Schwartz calling Peña's termination of the contract "a totally disproportionate response" to the problems at the Lab.

Once they had announced their intentions, the bidding teams began preparing their proposals, heading toward a September 1 deadline that changed to September 8 on RPI's request.

But the picture changed again on the 8th, a Monday, when news broke that RPI's president had decided to withdraw his institution's participation and that Westinghouse had lined up a new partner and been granted a two-week extension just 15 minutes before the deadline. While Stony Brook and Battelle successfully submitted their bid that day, the new team of Westinghouse and the Illinois Institute of Technology Research Institute finalized its bids and submitted it by the new due date of September 22. Both teams presented their bids to DOE in late September.

Despite the delays, DOE expects to evaluate the bids and announce the victor on November 17, kicking off a transition period to transfer control from AUI to the new contractor. By early January, the Lab will be on its way to a new future, a much different one than what it faced just a year ago.

Staff News

The PDB has two new staff members that we would like to introduce. We also were enriched by the close collaboration of several summer visitors and student interns this year.

The PDB welcomes Jiri Koutnik. Jiri is working toward his Ph.D. in computer science, and comes to us from the Czech Technical University, Prague, Czech Republic. As part of the PDB's Informatics division, he will be helping us develop better data handling tools for the "next generation PDB R&D".

Michael Miley, a recent graduate of the State University of New York at Stony Brook's Pharmacology School, joins the PDB's User Support Group. He is handling the Quarterly Releases, including the distribution of the CD-ROM and Newsletters, as well as some of the Help Desk queries. As part of the professional staff, Mike is also processing coordinate entries. Mike has been a student assistant with the PDB since January 1997, and we are pleased to welcome him as permanent staff.

Masami Kusunoki, Associate Professor of Crystallography at Osaka University in Japan, was here for the summer learning how the PDB functions. Upon returning home, he plans to set up a deposition and distribution node of the PDB at the University of Osaka. This important development will be a major asset to our users in Asia and the Pacific Rim, and the PDB is pleased to be collaborating with one of our long-standing Centers for this resource.

Manfred Hendlich, a computational chemist from Marburg University in Germany, worked at the PDB this summer as a visiting scientist. He devoted his time to the development of data handling and distribution tools for use in AutoDep and for the heterogen database, RELIBase. A brief description of this database is given in this Newsletter.

Petr Kocab of SoftDeC, Prague, and Martin Senger of the EBI, Hinxton Hall, UK were also here for several weeks this summer. Working closely with Otto Ritter, Head of Informatics, they are developing software components as part of the PDB information system reengineering effort.

Wen-Juan Xu comes to the PDB through the Suffolk County Community College Women's Internship Program. Wen-Juan is working part-time during the fall semester completing development of a database for the use of our User Support Group. We are also happy to report that Mariya Kobiashvili and Sabrina Hargrove are extending their time with us into the fall semester. They are assisting the Archive Management Group, Mariya with the Help Desk and Sabrina providing administrative and clerical support.

The PDB wishes a fond farewell to three students who were with us this summer, Keith Peters, Jie Zhang, and Steve Murtagh. Keith is a sophomore at the California Institute of Technology, majoring in computer science. He spent his third consecutive year with us aiding John McCarthy with computer system administration. Jie, who is a junior at William Floyd High School in Shirley, NY, spent his summer break with us for his second consecutive summer doing some computer programming and database work. Steve Murtagh, a sophomore at Shorham High School in Shorham, NY, worked with us this summer doing some computer programming. They have all gone back to school and we wish them well.

The PDB would also like to say good-bye to Brigitte Sylvain who has completed her term as a Professional Associate with the PDB. Brigitte is now pursuing graduate studies in molecular biology.

PDBObs: The Obsolete Structures Database has Been Upgraded

— Helge Weissig and Phil Bourne,
San Diego Supercomputer Center, San Diego, CA, USA
(helgew@sdsc.edu or bourne@scsc.edu; http://pdbobs.sdsc.edu).

PDBObs, the archive of obsolete PDB entries at the San Diego Supercomputer Center has moved to a new location: **http://pdbobs.sdsc.edu/**. Although requests to the former address will be automatically forwarded, all future queries should be directed to the new URL. PDBObs is now running on an eight processor Sun Enterprise Server 6000 as part of an initiative with Sun Microsystems Inc. to provide high quality services to the molecular biology community. The new system provides a significant reduction in the time taken to process requests.

Beyond improved hardware, several new features have been added to PDBObs since its introduction at the beginning of the year. The following additions further facilitate the comparison of all versions of a particular structure:

- a comprehensive table outlining differences and commonalities in experimental data (e.g., cell parameters, space group, number of atoms) and supporting information (e.g., authors, compound name, deposition, and release dates)
- comparison of dihedral angles, bond angles, and bond lengths for single residues and groups of residues (e.g., all proline or all alpha helical residues)
- introduction of a quantitative "fold deviation score" (FDS) as a means of comparison of structural parameters to "standard" small molecule data
- analysis of the whole structure for FDS scores as above and displayed as a colored C-alpha traces in Rasmol.

All new features are supported by the property object model (I.N. Shindyalov and P.E. Bourne (1997) *CABIOS* **13**: in press). For further information about the new features outlined above see **http://pdbobs.sdsc.edu/whatsnew.html.**

RELIBase: A Database of Receptor/Ligand Complexes

— Manfred Hendlich, Department of Pharmaceutical Chemistry,
University of Marburg, Germany
(hendlich@pharmazie.uni-marburg.de).

RELIBase (**http://pdb.pdb.bnl.gov:8081/home.html**) is the first free Web-based service that can access receptor/ligand structures deposited in the PDB by true 3D search queries.

The main purpose of RELIBase is to provide a selective and efficient access to receptor/ligand complexes currently deposited in the PDB and to make the enormous wealth of information contained in the receptor/ligand structures available for structure-based drug design studies.

The WWW-based interface to RELIBase can be used to input a sub-structure search object either by text, a smiles string, or by an interactive Java-based molecule editor, and the system can perform the following functions:

- 2D substructure searches for fast identification of all ligands which contain a specific functional group.

- 2D similarity searches.
- Protein sequence similarity searches.
- 3D substructure searches permit the identification and analysis of specific interactions between ligands and proteins.
- Analysis of interaction preferences. RELIBase permits a statistical analysis of the atomic environment seen by functional groups in protein/ligand complexes.

More on RELIBase will appear in the next PDB Quarterly Newsletter. Please send your comments on RELIBase to: hendlich@pharmazie.uni-marburg.de.

HIC-Up: Dealing with Hetero-Compounds

— Gerard J. Kleywegt, Department of Molecular Biology, Uppsala University, Uppsala, Sweden (gerard@xray.bmc.uu.se).

As more and more complexes of biomacromolecules and small molecules (inhibitors, ligands, substrate analogues, etc.) are solved, there is a need for resources to deal with these compounds. In particular, during (crystallographic) refinement and rebuilding operations, users need "dictionaries" describing the geometry and stereochemistry of such hetero-entities. Generating such dictionary files is a cumbersome, time-consuming and error-prone undertaking. In order to simplify the process, we have collected a large set of hetero-entities from the Protein Data Bank (PDB), and implemented some tools that automate much of the dictionary-generation process.

— Collection of Hetero-Compounds

The best place to start looking for coordinates for a hetero-compound is the CSD, the Cambridge small-molecule crystallographic database. If such a search yields no useful clues (either the exact compound you are looking for, or a related compound, or a set of fragments), or if one does not have access to the CSD, the next-best thing may be to check if anyone else has previously used the compound in a macromolecular refinement. In that case, the PDB is the most suitable place to look. In order to make the search as simple as possible, we have written a script which automatically scans the PDB and finds all unique hetero-compound names. This list is fed into a program that scans all PDB entries again, in order of decreasing resolution (and NMR structures after X-ray structures). As soon as one of the hetero-compounds from the list is encountered, the coordinates are stored. Once the scan is complete, every hetero-compound is translated to put its centre-of-gravity at the origin, all occupancies are set to 1.0 and all temperature factors to 20.0 Å^2. The coordinates and some additional information (PDB entry from which it was taken, resolution, a list of other PDB files which contain the same compound, etc.) are then written to one large file. Our present collection, generated using the December 1996 version of the PDB, contains more than 1,100 (mostly unique) hetero-compounds - the file is updated about twice a year. This file (called "hetero.pdb") is also used by the BioMagResBank (**http://www.bmrb.wisc.edu/**) in its electronic deposition procedure. As an example, the entry for retinol looks as follows:

```
COMPND RTL RETINOL
REMARK RTL Extracted from PDB file 1hbp.pdb
REMARK RTL Formula C20 H30 O1
REMARK RTL Nr of non-hydrogen atoms 21
REMARK RTL Residue type RTL
REMARK RTL Residue name 951
REMARK RTL Original residue name (for O) $176
REMARK   2 RESOLUTION. 1.9  ANGSTROMS.  1HBP  29 RTL
REMARK RTL Compound also present in : 1RBP 1CRB 1BRP
HETATM    1  C1  RTL   951      -2.428   1.801  -3.065  1.00 20.00 1HBP
HETATM    2  C2  RTL   951      -2.950   1.847  -4.505  1.00 20.00 1HBP
HETATM    3  C3  RTL   951      -2.239   0.962  -5.522  1.00 20.00 1HBP
[...]
HETATM   20  C20 RTL   951       1.882  -0.696   5.344  1.00 20.00 1HBP
HETATM   21  OR  RTL   951       5.637  -1.079   5.994  1.00 20.00 1HBP
REMARK RTL ENDHET
```

Using an editor, or a unix tool as simple as "grep", one can quickly find out if the compound one is looking for occurs in the file. Since compound names are not always clear and unambiguous (suspiciously many compounds are called "SEE REMARK 7", for instance), one may also use a jiffy program (CT2HET) to search the collection based on a connectivity file. This program is also useful if one wants to find all six-membered sugar compounds, etc.

— X-PLOR/CNS and TNT Dictionaries

XPLO2D is a utility program that contains an option to generate appropriate dictionaries for X-PLOR/CNS. Given a PDB file containing the coordinates of a hetero-compound, it generates four new files:

- a topology file (defining atom types, masses, etc., bonds, impropers [chiral carbons, flat groups and bonds], possible dihedrals, hydrogen-bond acceptors and possible donors). This file may need to be edited, for instance to add charges and the masses of implicit hydrogen atoms (although both are usually ignored by crystallographers nowadays).
- a parameter file (defining target values and force constants for bonds, etc.). The target values are simply the averages of the observed values. The force constants are set to the same value for all bonds, angles and impropers (the defaults being in the same ball-park as those of the Engh and Huber force field).
- an X-PLOR input file which, when executed, will energy-minimise the structure of the compound and print a list of violations. This should always be done prior to inclusion of the compound into the refinement process, since the resulting structure reveals what X-PLOR will try to make the compound look like once it is included. If, for instance, a dihedral angle was given a target value of 180 degrees whereas it should have been 0 degrees, this will show up immediately after the energy minimisation. Hence, this is a quick and easy way to prevent the frustration of finding that X-PLOR has "ruined" your compound after a 4,000 K slow-cool which took two days to execute ...
- a "clean" PDB file suitable for use by X-PLOR in the energy-minimisation procedure.

Recently, this program has been modified so that it can also produce rudimentary dictionaries for use with the crystallographic refinement program TNT. The program can be fed multiple copies of a hetero-entity, in which case all bond lengths, etc. are averaged to produce target values.

— O Dictionaries

Another utility program (MOLEMAN or MOLEMAN2) can be used to generate four of the five types of dictionary file that may be needed for the display and manipulation of a hetero-compound inside O. The only dictionary that cannot be generated in this fashion is that required for regularisation. On the other hand, regularisation can be done rapidly in X-PLOR, and if one uses sensible manipulation commands in O it will rarely be necessary to regularise a hetero-compound. The four types of dictionary that can be generated automatically involve:

- Connectivity. In order for O to draw the correct bonds, e.g., no bonds between hydrogen atoms, a connectivity entry is sometimes needed (although in most cases the defaults in O will do the job).
- Real-space fit. To include a compound in real-space fit calculations, a list of all its atoms can be provided as an O datablock.
- Real-space refinement. A similar datablock is needed to include the compound in some real-space refinement calculations (RSR_zone).
- Torsions. Some torsion-angle manipulations in O require that these angles and the affected atoms be defined.

— Availability

The Uppsala resources related to hetero-compounds have been collected into a new, freely-accessible Web site called HIC-Up (Hetero-compound Information Centre-Uppsala), which can be accessed at URL: **http://alpha2.bmc.uu.se/hicup/.**
The site contains a subset of the following items for each of over 1,100 hetero-compounds:

- a PDB file with some REMARKs and with HETATM records
- a "clean" PDB file, ready to use in your favourite model building and refinement programs
- the output of the program HETZE which attempts to check the quality of hetero-compounds
- the "official" PDB dictionary file (containing CONECT records, etc.)
- a VRML file (requires a VRML browser or plug-in)
- a ChemScape Chime page (requires the Chime plug-in)
- a connection table
- a list of other hetero-compounds which are either identical to the present compound, or a superstructure of it (this search is only carried out if the compound contains 10 or fewer non-hydrogen atoms)
- X-PLOR/CNS dictionary files
- O dictionary files
- a TNT dictionary file
- a pointer to the PDBSUM site at UCL, London, showing a list of all PDB files which contain the present hetero-compound
- a summary of the HETZE report
- a GIF image of the compound (generated with LIGPLOT)
- an "ASCII picture" showing atom names and connectivity, taken from the PDB dictionary file

The compounds can be accessed through several sorted index lists (residue name, trivial name, chemical formula, number of non-hydrogen atoms).

The HIC-Up site also contains:

- a service to generate dictionaries for O, X-PLOR/CNS and TNT from a structure you provide;
- a service to run HETZE, a program which does some basic quality checks on hetero-compound structures;
- a list of hyperlinks to other sites containing useful information related to hetero-compounds in structural biology.

SBNet: The Swedish Structural Biology Network

— Gerard J. Kleywegt, Department of Molecular Biology, Uppsala University, Uppsala, Sweden (gerard@xray.bmc.uu.se).

Shortly before it lost the 1994 elections, the previous Swedish government allocated a whopping six billion Swedish crowns (almost one billion US$) to stimulate scientific research and development efforts. A major recipient of this money was the Swedish Foundation for Strategic Research (**http://www.stratresearch.se/**). The "mission" of this Foundation is to support scientific, technical and medical research. It does so by promoting the development of important research environments of the highest international standing with a view to enhancing Sweden's long-term competitiveness.

The Foundation issued a call for proposals, and the Swedish structural biologists were among the first to comply. In the Autumn of 1994, a programme committee consisting of Alwyn Jones (Uppsala University), Hans Eklund (Swedish University of Agricultural Sciences, Uppsala), Björn Nilsson (Pharmacia & Upjohn, Stockholm), Torleif Härd (Karolinska Institute, Stockholm) and Sture Forsén (Lund University) compiled the first detailed proposal, outlining the contours of a Structural Biology Network. A sum of roughly 65 million Swedish crowns (approximately ten million US$) for the period 1996-2000 was eventually allocated by the Foundation, and in 1996 the Network (dubbed SBNet) took off. The Network is run by the Programme Director (Alwyn Jones) and the Chairman of the Board (Björn Nilsson), assisted by the Network Coordinator (Gerard Kleywegt). The Network Board is responsible for the allocation of finances and for steering the Network into areas of strategic importance. Besides Jones and Nilsson, its members are Jan Hoflack (Astra, Gothenburg), Guy Dodson (University of York) and Iain Campbell (Oxford University).

The Network is officially hosted by Uppsala University, and it unites academic and industrial structural biologists from the following disciplines:

- biomacromolecular X-ray crystallography
- high-resolution NMR spectroscopy of biomacromolecules
- near-atomic resolution electron microscopy of biomacromolecules
- computer modeling of biomacromolecules

The Network aims to strengthen the strategic value of structural biology in Sweden by using a four-tiered approach:

1. To reinforce excellence
2. To remedy current weaknesses
3. To elevate graduate training to a world-class level

In order to achieve these aims, the Network:

- funds the building of a dedicated expression laboratory in Uppsala
- funds a number of Ph.D. student and researcher positions throughout Sweden
- organises a mentor system for the students who are funded through the Network

- organises an Annual Conference open to all interested structural biologists in Scandinavia
- organises workshops and advanced graduate courses on specialised topics in any of the disciplines of structural biology
- awards travel grants to stimulate lab rotation visits by students and post-docs
- maintains a mailing list through which news and information is disseminated to Swedish structural biologists
- maintains a Web site to foster the exchange of news and information relevant to the Network or Swedish structural biology in general

The First Annual Conference of the Network was organised in June. It was set up in a format not unlike that of the Gordon Research Conferences, with lectures by invited speakers from abroad and by young people from within Sweden, and with the afternoons dedicated to informal discussions. There were five sessions, dealing with X-ray techniques and applications, drug design, membrane proteins, hot structures, and NMR techniques and applications. In addition there was a poster session (and a football match between the crystallographers and the NMR spectroscopists that ended with a score of 4-1). Pharmacia & Upjohn sponsored an award for the best poster, which went to Mathias Eriksson ("Ribonucleotide reductase"). Astra's Structural Chemistry Laboratory sponsored an award for the best presentation by a young person. This award was won by Helena Berglund ("Structural studies of the ribosomal protein S15").

The Conference was attended by 120 structural biologists (the maximum number that could be accommodated) from all four disciplines, and with a strong representation from the two industrial labs at Pharmacia & Upjohn and at Astra. The atmosphere was very friendly and the discussions lively. Almost 60 young people presented posters of a generally high quality, indicating that structural biology is a rapidly expanding and dynamic field in Sweden. In his closing remarks, the "eminence grise" of Swedish structural biology, Carl-Ivar Brändén, noted that the fact that so many young people were attracted to this field was a good sign, since students often have a better impression of which scientific areas are important than the people who supervise them. After the meeting, one of the invited speakers said that he wished that there were such a Network in his own country.

For more information about the Network, see the SBNet web site at URL: **http://alpha2.bmc.uu.se/~gerard/srf/** or send e-mail to: sbnet@xray.bmc.uu.se. For a (biased) report of the X-ray versus NMR football match see: **http://alpha2.bmc.uu.se/~gerard/srf/conference_1997.html.**

Two New Programs for Accessing the PDB

— Chaitali Mukhopadhyay and Sanjib Kumar Gupta,
Distributed Information Centre, Bose Institute, Calcutta, India
(dicroot@giascl01.vsnl.net.in, dic@boseinst.ernet.in,
root@boseinst.ernet.in).

— PDBVIEW

PDBVIEW is a PC - based software program that can search PDB entries by compound, source, and header. The Boolean operators "and, or, not" can be used for searching. The program then asks for the display option. The display of an individual entry contains the header, compound, and source fields. It also asks whether the sequence is to be displayed or not. The sequence display appears at the bottom of the screen as a line with horizontal boxes of different types for helix and sheet components. The length of the boxes is proportional to the length of the secondary structural elements.

A repeated display of a homologous family of proteins helps one to identify the common secondary structural elements in them. This program is also helpful for browsing the PDB files and for a two dimensional representation of the secondary elements of a protein.

The program uses several databases (generated from the PDB files) containing various fields and secondary structural information. These databases have to be generated prior to the searching. The program that can read the CD-ROMs containing PDB entries and generates the databases are available on request from dicdbt@boseinst.ernet.in or root@boseinst.ernet.in.

All the programs have been developed in C in collaboration with students of Regional Computer Centre, Jadavpur, Calcutta, India.

— LOOP SEARCH

The program LOOP SEARCH has been developed in C and can be compiled on unix workstations having a CD-ROM drive. The program will read the PDB entries from the CD-ROM drive. The program takes the secondary structural information from the file. The structural elements are then arranged sequentially and the lengths of the loop residues connecting two such elements are calculated. The loops are categorized based on their connectivity, e.g., a loop connecting two helical segments are assigned as category 1, connecting one helix and one strand - 2, connecting two beta strands - 3, and connecting beta strand and helix - 4. The output file contains the entry name, loop length, loop category and the sequence in one letter code. The output file is in the FASTA format.

The output file is very useful for homology modeling purposes, where finding a loop of similar length and same category might aid in the modeling process. Also, creation of a general database of loops will be possible.

This program was developed in collaboration with Mr. Debanu Das, Chemistry Dept, IIT Kanpur, India and students of Regional Computer Centre, Jadavpur, Calcutta, India.

Notes of a Protein Crystallographer

— The Lunar Element and our Collective MADness

— Cele Abad-Zapatero, Dept. of Structural Biology,
Abbott Laboratories, Abbott Park, IL, USA (abad@abbott.com).

In the old part of Grenoble (Department of Isère and capital of the Dauphiné province) in France, there is a café named "Café de la Table Ronde", right across from the old Palais de Justice in the Place de St. André. According to some of the newspaper clippings hanging from its rustic walls, this coffee house was the second in France, founded in 1739, seventy years after the first coffee house in France the "Café d'Europe" in Paris. Among its clientele this café boasted actors such as Fernandel and Gerard Depardieu; folk singers and balladeers like Jacques Brel and George Brassens, and writers of the caliber of Pierre Choderlos de Laclos (1741-1803; *Les Liaisons Dangereuses*, object of an American film by the same title) and Stendhal (1783-1842; pseudonym of Marie-Henri Beyle,

Le Rouge et le Noir), who was a native of Grenoble. I am certain that in most recent years it has also been the place of gathering for many of the most illustrious scientists that visit the ESRF in Grenoble every year. However, they were not listed yet.

I was sitting at this cafe late in the evening on June 16, 1996 at the conclusion of the workshop on Multiple Anomalous Diffraction (MAD, Grenoble, June 10-16, 1996) marking time until the departure of my TGV train back to Paris. Somehow, an anonymous photographer managed to capture the very same spot I was sitting down in one of the Grenoble postcards. The square where the outdoor tables were set was full of life, and a full Moon was shining up on one of the corners of the quadrilateral defined by the roofs of the buildings around me, including the parish of Saint André. As I remember that night, and the subsequent developments in our laboratory, I feel as if, in that particular moment, I was "moonstruck" in a scientific rather than romantic manner. At that time, our laboratory had just solved, in a matter of only a few weeks, its first structure using the MAD methodology. From then on, I have felt captivated by the "power and sheer elegance of the method (1)" derived from the substitution of Sulfur by Selenium in the protein.

Selenium (from Greek selene, "moon") derives its name from its silvery appearance and was recognized as an element in 1818 by Jöns Jacob Berzelius. Its location in *The Periodic Kingdom* (2) is in the Eastern Rectangle (the p-block), right after the isthmus that connects the strong metals (Western rectangle) with the nonmetallic elements and has the yellow Sulfur as its northern neighbor. Its gray metallic, lunar-like, appearance was probably the main reason behind its name but its ability to exist in several different colored forms, akin to the phases of the Moon established for me another connection with our night planet. As an amorphous (noncrystalline) powder it is red but transforms easily into a black vitreous glass. As a crystal, it can be red or gray with the latter being the most stable under ordinary conditions. By itself it is not poisonous but many of its compounds are very toxic. However, it is also an essential mineral that has to be provided in the diet of certain animals.

After that night, it seems as if I am under the spell of this gray metalloid element and I think that I am not alone in view of the many structures that are solved nowadays using this methodology. The spell of this lunar element on protein crystallographers is undoubtedly due to the unique properties of its electron cloud with an absorption K-edge at 12.658 KeV (0.9795 Å). Its strong attractive influence is also due to the fact that it can be easily incorporated into our protein samples and that, most of the time, our crystals do not even notice it (3,4).

As judged by the number of protein structures solved by this method in the last year (5), my lunar hallucinations have spread to a large number of members of the protein crystallography community and may soon reach the level of a collective MADness. The origins of this hysteria can be traced back for a number of years and I will not elaborate on this (6,7). However, I would argue that there must be an instigator, a culprit, for such a large group of "intelligent" people to become engulfed in such a collective state of agitation. I have not taken a poll but I think that, for once, the community is unanimous and unambiguous as to who that person is, and I will not even have to name the author of such witchcraft.

May these few lines be an expression of respect and admiration for the leader of this unusual and benign cult. May these words pay homage to one of the unsung heroes of the structural revolution in Biology, who was recently justly honored with the Gregori Aminoff Prize in Sweden. May these letters express my personal gratitude, that of my colleagues, and the one from the community at large, to such an "anomalous" and visionary colleague. He has opened for us a path to solve new structures, illuminated by the pale light and color of the lunar element combined with the brightness and tunability of our most brilliant X-ray sources. When I go to the *Café de la Table Ronde* next time, I would like to see the names of some scientists among its illustrious clientele, and our MADman among them.

References

1. Words used by Dr. Steve Cusack at the opening of the Workshop at the EMBL Grenoble Outstation.
2. *The Periodic Kingdom: A Journey into the Land of the Chemical Elements* by P.W. Atkins. Basic Books, a Division of Harper Collins Publishers, 1995.
3. Cowie, D. B. and Cohen, G. N. (1957). *Biochim. Biophys. Acta*, **26**, 252-261.
4. Hendrickson, W. A., J. R. Horton and David M. LeMaster (1990) *EMBO J.*, **9**, 1665-1672.
5. Abstracts of the Annual ACA meeting in St. Louis, Missouri, July 19-15, 1997.
6. Hendrickson, W. A. (1985). *Transactions of the American Crystallographic Society* **21**, 11-21.
7. *Macromolecular Crystallography with Synchrotron Radiation* by J.R. Helliwell, Cambridge University Press, Cambridge, England, 1992.

WEB SITES

Referenced in the October 1997 PDB Newsletter

Archive of Obsolete PDB Entries
http://pdbobs.sdsc.edu/
Recent Changes and Additions
http://pdbobs.sdsc.edu/whatsnew.html

AutoDep Release Notes
http://www.pdb.bnl.gov/release-notes.html

BioMagResBank
http://www.bmrb.wisc.edu/

HIC-Up (Hetero-compound Information Centre Uppsala)
http://alpha2.bmc.uu.se/hicup/

Protein Data Bank
http://www.pdb.bnl.gov

RELIBase
http://pdb.pdb.bnl.gov:8081/home.html

SBNet (Swedish Structural Biology Network)
http://alpha2.bmc.uu.se/~gerard/srf/
Report of the X-ray versus NMR football
http://alpha2.bmc.uu.se/~gerardsrfconference_1997.html

Swedish Foundation for Strategic Research
http://www.stratresearch.se/

What Check
http://www.sander.embl-heidelberg.de/rob/checkhelp/

AFFILIATED CENTERS AND MIRROR SITES

Thirty-six affiliated centers offer the Protein Data Bank database archives for distribution. These centers are members of the Protein Data Bank Service Association (PDBSA). Centers designated with an asterisk(*) may distribute the archives both on-line and on magnetic or optical media; those without an asterisk are on-line distributors only. Information is given for those Centers that are now also official PDB Mirror Sites.

***APPLIED THERMODYNAMICS, LLC**
Hunt Valley, Maryland, USA
George Privalov (410-771-1626)
George_Privalov@classic.msn.comhttp://www.mole3d.com

BIRKBECK
Crystallography Department
Birkbeck College, University of London
London, United Kingdom
Ian Tickle (44-171-6316854)
tickle@cryst.bbk.ac.uk
http://www.cryst.bbk.ac.uk

BMRB
BioMagResBank
University of Wisconsin - Madison
Madison, Wisconsin, USA
Eldon L. Ulrich (608-265-5741)
elu@bmrb.wisc.edu
http://www.bmrb.wisc.edu

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

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@ccdc.cam.ac.uk
http://www.ccdc.cam.ac.uk

PDB Mirror Site:
http://pdb.ccdc.cam.ac.uk/
Ian Bruno (mirror@ccdc.cam.ac.uk)

CMU
Carnegie Mellon/Pittsburgh Supercomputing Center
Pittsburgh, Pennsylvania, USA
Hugh Nicholas (412-268-4960)
nicholas@psc.edu
http://pscinfo.psc.edu/biomed/biomed.html

CSC
CSC Scientific Computing Ltd.
Espoo, Finland
Erja Heikkinen (358-9-457-2433)
erja.heikkinen@csc.fi
http://www.csc.fi

DKFZ
German Cancer Research Center
Heidelberg, Germany
Otto Ritter (49-6221-42-2372)
o.ritter@dkfz-heidelberg.de
http://www.dkfz-heidelberg.de

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

PDB Mirror Site:
http://www2.ebi.ac.uk/pdb
Philip McNeil (pdbhelp@ebi.ac.uk)

FUJITSU KYUSHU SYSTEM ENGINEERING LTD.
Computer Chemistry Systems
Fukuoka, Japan
Masato Kitajima (81-92-852-3131)
ccs@fqs.fujitsu.co.jp
http://www.fqs.co.jp/CCS

FMI
Friedrich Miescher Institute
Basel, Switzerland
Carl David Nager (41-61-697-5678)
carl.david.nager@fmi.ch
http://www.fmi.ch

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

:
IGBMC
Laboratory of Structural Biology
Strasbourg (Illkirch), France
Frederic Plewniak (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)

LIGM
Laboratorie d'ImmunoGenetique Moleculaire
Montpellier, France
Marie-Paule LeFranc (33-04-67-61-36-34)
Lefranc@ligm.crbm.cnrs-mop.fr
http://imgt.cnusc.fr:8104

MAG
Molecular Applications Group
Palo Alto, California, USA
Margaret Radebold (415-846-3575)
bold@mag.com
http://www.mag.com

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

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

NATIONAL RESEARCH COUNCIL OF CANADA
Institute for Marine Biosciences
Halifax, N.S., Canada
Christoph W. Sensen (902-426-7310)
sensencw@niji.imb.nrc.ca
http://cbrmain.cbr.nrc.ca

NATIONAL TSING HUA UNIVERSITY
Department of Life Science
HsinChu City, Taiwan
J.-K. Hwang (+886 3-5715131, extension 3481) or lshjk@life.nthu.edu.tw
P.C. Lyu (+886 3-5715131 extension 3490) lslpc@life.nthu.edu.tw http://life.nthu.edu.tw

PDB Mirror Site:
http://pdb.life.nthu.edu.tw/
Tony Wu (mirror@life.nthu.edu.tw)

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

NCSA
National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign
Champaign, Illinois, USA
Allison Clark (217-244-0768)
aclark@ncsa.uiuc.edu
http://www.ncsa.uiuc.edu/Apps/CB

NCSC
North Carolina Supercomputing Center
Research Triangle Park, North Carolina, USA
Linda Spampinato (919-248-1133)
linda@ncsc.org
http://www.mcnc.org

****OML***
Oxford Molecular Ltd.
Oxford, United Kingdom
Kevin Woods (44-1865-784600)
kwoods@oxmol.co.uk
http://www.oxmol.co.uk or http://www.oxmol.com

****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, Institute of Physical Chemistry
Beijing 100871, China
Luhua Lai (86-10-62751490)
lai@ipc.pku.edu.cn
http://www.ipc.pku.edu.cn

PDB Mirror Site:
http://www.ipc.pku.edu.cn/pdb
Li Weizhong (liwz@csb0.ipc.pku.edu.cn)

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

SEQNET
Daresbury Laboratory
Warrington, United Kingdom
User Interface Group (44-1925-603351)
uig@daresbury.ac.uk
http://www.seqnet.dl.ac.uk

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

TURKU CENTRE FOR BIOTECHNOLOGY
University of Turku and Abo Akademi University
Turku, Finland
Adrian Goldman (358-2-3338029)
goldman@btk.utu.fi
http://www.btk.utu.fi

UNIVERSIDAD NACIONAL DE SAN LUIS
Facultad de Ciencias Fisico Matematicas y Naturales
Universidad Nacional de San Luis
San Luis, Argentina
Jorge A. Vila (54-652-22803)
vila@unsl.edu.ar
http://linux0.unsl.edu.ar/fmn

PDB Mirror Site:
http://pdb.unsl.edu.ar
Fernando Aversa (aversa@unsl.edu.ar)

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciencias Biologicas
Belo Horizonte, MG - Brazil
Marcelo M. Santoro (55-31-441-5611) santoro@icb.ufmg.br
Ari M. Siqueira (55-31-952-7470) siqueira@icb.ufmg.br
http://www.icb.ufmg.br

PDB Mirror Site:
http://www.pdb.ufmg.br
Ari M. Siqueira (siqueira@cenapad.ufmg.br)

UNIVERSITY OF GEORGIA
BioCrystallography Laboratory
Department of Biochemistry and Molecular Biology
University of Georgia
Athens, Georgia, USA
John Rose or B.C. Wang (706-542-1750)
rose@BCL4.biochem.uga.edu
http://www.uga.edu/~biocryst

PDB Mirror Site:
http://BCL10.bmb.uga.edu
John Rose (rose@BCL4.biochem.uga.edu)

UPPSALA UNIVERSITY
Department of Molecular Biology
Uppsala University
Uppsala, Sweden
Alwyn Jones (46-18-174982)
alwyn@xray.bmc.uu.se
http://pdb.bmc.uu.se or http://alpha2.bmc.uu.se

WARSAW UNIVERSITY
Interdisciplinary Centre for Modelling
Warszawa, Poland
Wojtek Sylwestrzak (48-22-658-4901)
W.Sylwestrzak@icm.edu.pl
http://www.icm.edu.pl

WEHI
The Walter and Eliza Hall Institute
Melbourne, Australia
Tony Kyne (61-3-9345-2586)
tony@wehi.edu.au
http://www.wehi.edu.au

PDB Mirror Site:
http://pdb.wehi.edu.au/pdb
Tony Kyne (tony@wehi.edu.au)

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

PDB Mirror Site:
http://pdb.weizmann.ac.il
Alexander Faibusovich (pdbhelp@pdb.weizmann.ac.il)

Related WWW Sites

Databases

Archive of Obsolete PDB Entries	http://pdbobs.sdsc.edu/
BMRB (BioMagResBank)	http://www.bmrb.wisc.edu
CCDC (Cambridge Crystallographic Data Centre)	http://www.ccdc.cam.ac.uk
EBI (European Bioinformatics Institute)	http://www.ebi.ac.uk
EMBL (European Molecular Biology Laboratory)	http://www.embl-heidelberg.de
ExPASy Molecular Biology Server	http://expasy.hcuge.ch
GDB (Genome Data Base)	http://gdbwww.gdb.org
GenBank (NIH Genetic Sequence Database)	http://www.ncbi.nlm.nih.gov/Web/Genbank/index.html
HIC-Up (Hetero-compound Information Centre Uppsala)	http://alpha2.bmc.uu.se/hicup/
HIV Protease Database	http://www-fbsc.ncifcrf.gov/HIVdb/
Klotho: Biochemical Compounds Declarative Database	http://www.ibc.wustl.edu/klotho/
Library of Protein Family Core	http://WWW-SMI.Stanford.EDU/projects/helix/LPFC/
Crystal MacroMolecule Files at EBI	http://www2.ebi.ac.uk/msd/macmol_doc.shtml
NCBI (National Center for Biotechnology Information)	http://www.ncbi.nlm.nih.gov
NDB (Nucleic Acid Database)	http://ndbserver.rutgers.edu
PDB (Protein Data Bank)	http://www.pdb.bnl.gov
PIR (Protein Information Resource)	http://www-nbrf.georgetown.edu/pir
Prolysis: A Protease and Protease Inhibitor Web Server	http://delphi.phys.univ-tours.fr/Prolysis/
Protein Kinase Database Project	http://www.sdsc.edu/kinases/
Protein Motions Database	http://hyper.stanford.edu/~mbg/ProtMotDB/
RELIBase	http://pdb.pdb.bnl.gov:8081/home.html
SCOP: Structural Classification of Proteins	http://scop.mrc-lmb.cam.ac.uk/scop/
Mirrored at Protein Data Bank	http://www.pdb.bnl.gov/scop/
Swiss-Prot Sequence Database	http://expasy.hcuge.ch/sprot/sprot-top.html
CATH Protein Structure Classification	http://www.biochem.ucl.ac.uk/bsm/cath
Enzyme Structures Database	http://www.biochem.ucl.ac.uk/bsm/enzymes/
PDBsum	http://www.biochem.ucl.ac.uk/bsm/pdbsum

Software-Related Sites

CCP4	http://www.dl.ac.uk/CCP/CCP4/main.html ftp://ccp4a.dl.ac.uk/pub/ccp4
mmCIF	http://ndbserver.rutgers.edu/NDB/mmcif
O Home Page	http://imsb.au.dk/~mok/o/
OPM (Object-Protocol Model) Data Management Tools	http://gizmo.lbl.gov/DM_TOOLS/OPM/OPM.html
RasMol Home Page	http://www.umass.edu/microbio/rasmol/
SHELX Home Page	http://linux.uni-ac.gwdg.de/SHELX
Squid: Analysis and Display of Data from Crystallography and Molecular Dynamics	http://www.yorvic.york.ac.uk/~oldfield/squid/
VMD - Visual Molecular Dynamics	http://www.ks.uiuc.edu/Research/vmd/
X-PLOR Home Page	http://xplor.csb.yale.edu/

Other Resources

Crystallography Worldwide	http://www.unige.ch/crystal/w3vlc/crystal.index.html
BioMoo	http://www.cco.caltech.edu/~mercer/htmls/BioMOOHomePage.html
DALI - Comparison of Protein Structures in 3D	http://www.embl-heidelberg.de/dali/dali.html
NCSA Biology Workbench	http://biology.ncsa.uiuc.edu/
MOOSE (Macromolecular Structure Database at San Diego Supercomputer Center)	http://db2.sdsc.edu/moose
PDB_select: Representative PDBStructures	ftp://ftp.embl-heidelberg.de/pub/databases/protein_extras/pdb_select/recent.pdb_select
PROCHECK - To Submit a PDB File for Analysis	http://www.cryst.bbk.ac.uk/PPS/procheck/test.html
Protein Structure Verification-Biotech Server	http://biotech.embl-heidelberg.de:8400/
Mirrored at Protein Data Bank	http://biotech.pdb.bnl.gov:8400/
Resources for Macromolecular Structure Information	http://www.ucmb.ulb.ac.be/StructResources.html
The Virtual School of Molecular Sciences	http://www.vsms.nottingham.ac.uk/vsms/
Weizmann Institute, Genome and Bioinformatics	http://bioinfo.weizmann.ac.il/

Brookhaven Order Form

Name of User _____ Date _____

Organization _____ Phone _____

Address _____ Fax _____

_____ E-mail _____

- Price is valid through September 30, 1998
- 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; and
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. Please indicate exact address to which invoice should be sent:

_____

_____

_____

_____

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 Co. 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
Biology Department, Building 463
Brookhaven National Laboratory
P.O. Box 5000
Upton, NY 11973-5000

1 Protein Data Bank CD-ROM — ISO 9660 Format *$362.45*
Total for four (4) releases *$1449.80*
(tax and shipping charges not applicable)

This page intentionally left blank.

Access to the PDB

Main Telephone 1 516-344-3629

Help Desk Telephone 1 516-344-6356

Fax 1 516-344-5751

Help Desk ..... pdbhelp@bnl.gov

General Correspondence ..... pdb@bnl.gov

WWW Home Page ..... http://www.pdb.bnl.gov

FTP Server ..... ftp.pdb.bnl.gov

Network Services ..... sysadmin@pdb.pdb.bnl.gov

Entry Error Reports ..... errata@pdb.pdb.bnl.gov

Order Information ..... orders@pdb.pdb.bnl.gov

User Group ..... PDBusrgrp@suna.biochem.duke.edu

Listserver Postings ..... pdb-l@pdb.pdb.bnl.gov

Listserver Subscriptions ..... listserv@pdb.pdb.bnl.gov
to subscribe, the text of
your message should be .... subscribe PDB-L Your Name

FTP Directory Structure for Entries

The PDB has updated the FTP server (ftp.pdb.bnl.gov) in order to have a more standardized directory structure. This will facilitate use of mirror software to keep local copies of the database current.

Entry files are now found under the directory **pub/pdb/**

all_entries/
coordinate entry files in compressed and uncompressed format

biological_units/
generated coordinates for the biomolecules

current_release/
current database, with entries removed or added since the last CD-ROM

fullrelease/
static copy of the database as found on the last CD-ROM

latest_update/
entries added or removed in the most recent FTP update

newly_released/
entries released since the last CD-ROM

nmr_restraints/
compressed NMR restraint files

obsolete_entries/
withdrawn and/or replaced entries

structure_factors/
compressed structure factor files

fullrelease, newly_released, and current_release are divided into multiple subdirectories.

Scientific Consultants

John P. Rose, *University of Georgia, Athens, Georgia, USA*

Sasha Faibusovich
Clifford Felder
Kurt Giles
Jaime Prilusky
Mia Raves
Vladimir Sobolev
Yehudit Weisinger
Weizmann Institute of Science
Rehovot, Israel

PDB Staff

Joel L. Sussman, Head
Enrique E. Abola, Deputy Head and Head of Scientific Content/Archive Management
Otto Ritter, Head of Informatics

Frances C. Bernstein
Betty R. Deroski
Arthur Forman
Sabrina Hargrove
Jiansheng Jiang
Mariya Kobiashvili
Jiri Koutnik
Patricia A. Langdon
Michael D. Libeson
Dawei Lin
Nancy O. Manning
John E. McCarthy
Christine Metz
Michael J. Miley
Regina K. Shea
Janet L. Sikora
John Spiletic
S. Swaminathan
Wen-Juan Xu
Dejun Xue

Statement of Support

The 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.

Instructions to Authors

Contributions to the PDB Newsletter may be sent by e-mail or diskette to:
Nancy O. Manning, Editor
oeder@bnl.gov
Deadlines for contributions are:
March 1, June 1,September 1, and December 1.
Tel. 516-344-3629 Fax 516-344-5751
WWW http://www.pdb.bnl.gov

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

Telephone (516) 344-3629 Fax (516) 344-5752

Protein Data Bank
pdb™
Brookhaven National Laboratory

First Class