Quarterly Newsletter published by
Brookhaven National Laboratory
Protein Data Bank

Protein Data Bank
pdb™
Brookhaven National Laboratory

Release #85 **July 1998**

July 1998 CD-ROM Release

8035 Released Atomic Coordinate Entries

	Molecule Type
7109	proteins, peptides, and viruses
322	protein/nucleic acid complexes
592	nucleic acids
12	carbohydrates

	Experimental Technique
189	theoretical modeling
1272	NMR
6574	diffraction and other

2027	Structure Factor Files
429	NMR Restraint Files

The total size of the atomic coordinate entry database is 3.6 GB uncompressed.

Table of Contents

What's New at the PDB 1
Interoperability Between the PDB and mmCIF Data Formats ... 2
Release of AutoDep 2.1 3
Release Upon Publication 4
Staff Changes 4
PDB Lite Makes Its Debut 5
Sequence To and withIN Graphics PDB Viewer (STING – PDB viewer) 6
Searching for the Right PDB File with the 3DB Browser™ and Displaying Structure and Sequence Data with STING 7
Pattern Searching over the TOPS Protein Topology Database .. 7
BIOINFORMATICS: Creation of the Swiss Institute of Bioinformatics (SIB) 9
Development of a New Program for Structure Determination: Crystallography & NMR System 9
HIC-Up Updated 10
Notes of a Protein Crystallographer — Genomics, Proteomics, and the Secret of Life: A Faustian Dialog 10
Letters to the Editor 11
PDB Poetry: 1BAD story 12
Web Sites Referenced in this Newsletter 12
Related WWW Sites 13
Affiliated Centers and Mirror Sites 14
PDB™ Order Form 17
PDB Access, FTP Directory Structure, Consultants, Staff, Support, and Instructions to Authors 19

Internet Sites

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

What's New at the PDB

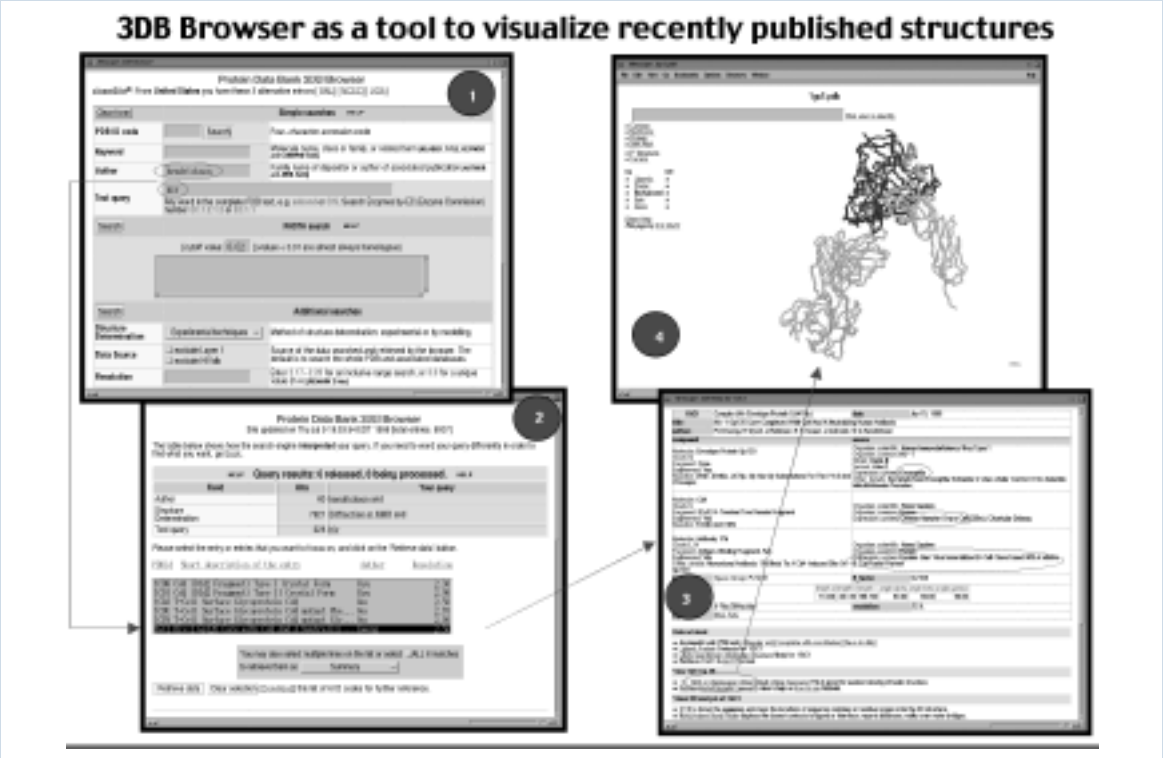
Joel L. Sussman

The PDB archive now has over 8,000 entries. Nineteen new structures were uploaded on the PDB server in the weekly Wednesday morning load of 22 July 1998, bringing the total number of entries past the 8,000 mark. The number of entries is still increasing very rapidly, with about one deposition every three hours! This is due in part to the inauguration of the first remote deposition site (January 1998) at the EMBL Outstation: European Bioinformatics Institute (EBI), UK, where approximately 30% of new entries are being deposited, making it faster and easier for European scientists to send their data to the PDB. We at the PDB are most interested in hearing your comments on tools for making structural information more easily accessible to the scientific community.

One of the easiest ways to search this large forest of entries in the PDB archives is to use the 3DB Browser™, which was developed by Jaime Prilusky of the Bioinformatics Unit of the Weizmann Institute of Science, Israel (see the PDB Home Page, http://www.pdb.bnl.gov). A detailed description of the 3DB Browser, as well as a number of helpful hints, can be seen in articles in the last two issues of the PDB Newsletter (Jan & Apr, 1998). Probably the best way to get a feeling for the 3DB Browser is just to try it. A simple example of its use is illustrated below in a search for a structure related to recent papers in *Nature* and *Science* (Kwong *et al.*, 1998; Rizzuto *et al.*, 1998). What one sees immediately upon retrieving the entry corresponding to these papers (PDB ID: 1GC1) is that the three proteins making up this complex were cloned in different expression systems and are derived from human cells and HIV Type 1. The tabular style of presenting the information is particularly useful for pinpointing the key information in the experiment, as well as for accessing various ways of visualizing the structure.

One of the easiest ways to view the basic structure is with the Chime™ software, a WWW plug-in. Chime is provided free by MDL Information Systems, Inc., and can be downloaded for PC's, Macintosh, and SGI computers from http://www.mdli.com/download/chimedown.html or http://www.mdli.co.uk/download chimedown.html.

Already one on-line journal, *Acta Crystallographica Section D* (*Acta Cryst. D*), has pointers from each article that refers to a PDB entry to the entry's 3DB Browser Atlas page. See, for example, the May 1998 issue at http://www.iucr.ac.uk/. This permits a reader of *Acta Cryst. D* to click on the entry and, within seconds, be able to see the structure in 3D. We are encouraging other journals to make these links. This has been facilitated with the inauguration of the PDB's Layered Release Protocol (on July 9, 1998) making it possible for depositors to explicitly indicate that their structural data should be 'Released on Publication' (see http://www.pdb.bnl.gov/pdb-docs/what_is_LR.html and the PDB Newsletters of Oct 97, Jan & Apr 98, http://www.pdb.bnl.gov/pdb-docs/newsletter.html.)



3DB Browser™ in action

1. *Search for Author: Hendrickson; Text query: HIV*
2. *6 hits obtained, 1GC1 highlighted*
3. *3DB Browser Atlas page. Red circles highlight the expression systems used for the different components in this multicomponent system.*
4. *Structure as visualized with MDL'S Chemscape Chime™ plug in.*

For novice users of the PDB an even simpler way to browse the archive is *via* PDB Lite, developed by Eric Martz at the University of Massachusetts, Amherst. PDB Lite is available directly from the PDB Home Page. It gives very clear instructions for searching and retrieving PDB entries, that are especially useful for those using the PDB for the first time. See the article on PDB Lite in this Newsletter. Other useful tools developed by Dr. Martz may be found at http://www.umass.edu/microbio/rasmol.

*See the article on PDB Lite in this Newsletter.

References

Kwong, P. D., Wyatt, R., Robinson, J., Sweet, R. W., Sodroski, J. & Hendrickson, W. A. (1998). Structure of an HIV gp120 Envelope Glycoprotein in Complex with the CD4 Receptor and a Neutralizing Human Antibody. *Nature* 393, 648-659.

Rizzuto, C. D., Wyatt, R., Hernandez-Ramos, N., Sun, Y., Kwong, P. D., Hendrickson, W. A. & Sodroski, J. (1998). A Conserved HIV gp120 Glycoprotein Structure Involved in Chemokine Receptor Binding. *Science* 280, 1949-1953.

Interoperability Between the PDB and mmCIF Data Formats

Otto Ritter

The PDB archive will in the near future support the mmCIF format as its primary canonical format for file-based data exchange, and the support for the traditional PDB format will be gradually phased out in the long perspective.

So far the canonical format has been the traditional PDB one, and mmCIF has been supported as an optional format for retrieval *via* the 3DB Browser. The standard format for structure factor files (http://www.pdb.bnl.gov/pdb-bin/ftp_index.pl?dir=ftp/structure_factors) has, however, long been the mmCIF.

A majority of software tools world-wide still depend on the PDB format. This is the main reason why, during a transition period, both the mmCIF and PDB format should be supported for both input and output operations on the PDB archive. Since the existing tools for translating between PDB and mmCIF do not do so without loss of information and/or without compromising the granularity of data representation, the PDB staff has undertaken the work of constructing a fine granularity syntactic and semantic mapping between the two formats.

This mapping, together with the software implementation of it, constitutes a means of high interoperability between the mmCIF and PDB formats, *i.e.*, a way of translating entry files in both directions without losing information and/or distorting the intended meaning of the data.

Specification

For automated translation between two different formats, three pieces of information are required: unambiguous definition of the two formats, so that files in each one could be unambiguously parsed, and an algorithmic description of how to transform a parsed file into the other format.

We have transformed the informal definition of the PDB format as in the PDB Contents Guide (versions 2.1 and 2.2), into a formal context-free grammar (one in the Backus-Naur Form) with additional algebraic and logical constraints. This grammar has over 1300 rules.

The mmCIF dictionary, together with its definition language (DDL), is already a formal grammatical representation. We have taken 1620 definitions from the official mmCIF Dictionary version 1.0.0, and added over 315 definitions for a local PDB extension dictionary.

Since the order of categories within an mmCIF data block and the order of items within a category instance is arbitrary, and so is the style of formatting with respect to white space, comments, etc., we have also defined a preferred PDB order and style for the mmCIF format. This part of the specification has almost 2000 additional rules.

Implementation

The GLASS (Grammatical, Logical and Algebraic System Specification) toolkit is used to interpret the grammatical and transformation rules. It is a formal model of structured data, where the structure and allowable values are defined by a declarative context-free grammar with an optional set of logical and algebraic constraints. GLASS grammar can also include structured annotation of itself, and rewrite rules for transforming documents into another form or language.

In addition to the 6000+ mapping and formatting rules, we have implemented 14 procedural functions (external methods called from for GLASS) for cases where pure GLASS transforms would end up being too slow, cumbersome to express, or impossible at all. Overall, it is much more flexible, cheaper, and less error-prone to maintain the declarations of uniform rules and the procedural code of only 14 functions, as opposed to a fully procedural code in some general-purpose programming language.

Status as of July 1998

The PDB → mmCIF direction is fully specified and fully operational. We are in the process of final alpha testing and fine-tuning of the rules, and we will shortly make available a large batch of translated mmCIF entries for external beta testing.

The mmCIF → PDB direction is fully specified in an mmCIF-style dictionary, and a large part of it is already translated into GLASS rules. The completion of this direction is scheduled after successful testing of the PDB → mmCIF direction. The local mmCIF dictionary and all the mapping and formatting specifications will shortly become publicly available from the PDB Web/FTP site.

Acknowledgments

The work on mappings and conversions has been a large team effort.

The semantic mappings were specified by the PDB Archive Management group members Enrique Abola, Jiansheng Jiang (who also programmed one major external function), Nancy Manning, and Regina Shea. Format for these specifications was extended mmCIF.

The conversion of mapping rules into the GLASS algebraic form, other syntactic transformation rules and GLASS-external methods were developed by Jiri Koutnik, who also did most of the work of formalizing the PDB format grammar.

The GLASS toolkit has been implemented by Ivo Marvan, Jiri Hovorka, and Petr Kocab of SoftDeC Ltd., Prague.

John Westbrook (Rutgers University), Frances and Herbert Bernstein (Bernstein + Sons), and Peter Murray-Rust (Nottingham University) provided us with several mmCIF-related tools, and gave helpful advice.

We acknowledge the hardware and software support of Digital Equipment Corporation and Silicon Graphics.

Funding for a large part of the work was provided by Brookhaven National Laboratory.

Release of AutoDep 2.1

Dawei Lin and Nancy Manning

The PDB released a new version of it Web-based submission program, AutoDep, on July 9, 1998, simultaneously at Brookhaven and the EMBL Outstation: The European Bioinformatics Institute (EBI). AutoDep 2.1 is compliant with the new Layered Release Protocol.

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

The major new features of AutoDep 2.1 include:

• **AutoDep 2.1 complies with the new Layered Released Protocol** adopted by the PDB. A complete submission must contain all mandatory tokens, and must pass the validation criteria as described in the document Validation for Layered Released and in the October 1997 and January 1998 PDB Newsletters, all available from our Web page (http://www.pdb.bnl.gov/). The coordinate entry and all generated diagnostics are reviewed by the author before completing the submission process. Following submission, the Layer 1 entry and its associated report file are loaded on the PDB servers.

• **AutoDep 2.1 complies with the new Layered Release Protocol** adapted by the PDB. Complete submissions must contain all mandatory tokens, and they must pass the stated validation criteria as described in Validation for Layered Release (http://www.pdb.bnl.gov/pdb-docs/validation.html). The entry and the associated report are reviewed by the author before submission. Following submission, the Layer 1 entry and the report file are loaded on the server.

- **The option "Release on Publication"** is added.
- **Templates** are added for **CNS and REFMAC**.
- **If several general remarks are merged into AutoDep 2.1**, they are treated as separate remarks.
- **AutoDep 2.1 preserves any special formats** in the following remarks:
    - Heterogen Description
    - Crystallization Description
    - Sequence Description
    - Site Description

- **The PDB validation summary output** is now being reported to the depositor. After reviewing the output, the depositor may decide to either complete the submission process or to revisit his/her structure and make changes before completing submission.
- **Cross checking of data items** has been added. For example, the relationship of the values given in R and R free, high and low resolution and the completeness of site records, such as site id, site description, and site residue(s), will now be cross checked.
- **AutoDep is now more customized** for experiment type. If experiment type is given as NMR, the crystal and crystal system items will be automatically marked N/A.
- **Warnings are provided** - if the molecule name or source organism cannot be found within lists of compounds and organisms in the Swiss-Prot database; if a journal does not have a match in the PDB journal list; or if a free text box has a line more than 59 characters long. To override the warning and get the green check mark, merely hit "Save Answers" again.
- **New report pages** - The reports are driven by a report menu page. There are up to four reports generated, depending on the data. These are: 1) Serious error report, 2) Caveat information report, 3) PDB summary report, and the 4) WHAT_CHECK(Hooft *et. al., 1996*) diagnostic report. A short WHAT_CHECK report is presented which highlights possible problems in the entry. This short report links to the complete WHAT_CHECK output. (Also see the April 1998 PDB Quarterly Newsletter.)
- **An option has been added to "Instructions to PDB"** to request that the PDB not re-label atoms to comply with IUPAC conventions.

All new submissions will use AutoDep 2.1 at either BNL or EBI. However, submissions previously begun with AutoDep 2.0 can be completed with the earlier interface. This will be handled automatically by the program just by "continuing" an AutoDep session of a BNL-number or EBI-number and password initiated in AutoDep 2.0. If you prefer to try AutoDep 2.1, use the "Based on a previous submission" function to start a new session based on the uncompleted AutoDep 2.0 session. After September 9, 1998, all submissions will be directed to AutoDep 2.1.

Please see the Release Notes (http://www.pdb.bnl.gov/release-notes.html) for more information about AutoDep 2.1. For more information about Layered Release Protocol go to http://www.pdb.bnl.gov/what_is_LR.html.

If you have any questions, please contact the PDB Help Desk at pdbhelp@bnl.gov or +1-516-344-6356.

Reference:

Hooft, R.W.W., Vriend, G., Sander, C., & Abola, E.E. (1996) *Nature* **381**, 27.

Release Upon Publication

Regina K. Shea

The PDB's new Layered Release protocol allows the depositor several options for releasing data submitted: release immediately, after a 6 or 12-month hold period, or upon publication of the article describing the coordinates. This last option, release upon publication, has always been available by request. However it has now taken on new meaning. Recently several journals announced a policy of requiring not only that the coordinates be submitted to the PDB before an article is accepted, but also that these coordinates be released before or at the time of publication of the article. Journals with this policy include Nature, Nucleic Acids Research, Proceedings of the National Academy of Sciences, Proteins, and Science.

In order to coordinate the publication of articles and the release of their associated coordinates, the PDB is setting up communication channels with these journals so that the publication information can be included in entries when they are released. This will also provide a timely trigger for the release of the entry.

The PDB is interested in expanding the list of journals that will send publication information on articles with deposited coordinates, even if release of the entry on publication is not required. The Journal of Biological Chemistry is sending references to the PDB. This is needed since automated searches of MEDLINE only pick up a fraction of the references that are not yet published, either because the journal is not indexed there or because the final version of the title has changed. PDB will also periodically contact depositors about the publication status of their articles.

Although an entry can be placed on hold until publication, in no case will this period exceed the PDB's stated maximum time of eighteen (18) months from date of deposition. In addition, PDB will continue to follow the IUCr guidelines that state that coordinates may be held (before release) no longer than one (1) year from the date of publication. These are maximum hold periods. If the article describing the coordinates is published in a journal with a shorter time limit, the PDB will follow the policy of that journal.

Staff Changes

Departures:

Frances Bernstein left the PDB following 24 years of service, most recently in the Archive Management Group. During her years at the PDB, Frances helped create PDB format, develop procedures and programs, and ensure the quality of the data. By her tireless devotion to the PDB and our users, Frances contributed to the success and growth of the PDB and we wish her well in the future.

Michael Miley has been a productive member of the Archive Management Group since January 1997. He has left in order to pursue a Ph.D. in bioorganic chemistry at Washington University, St. Louis, MO. We are confident that he will do very well in graduate school.

Arrivals:

The PDB is pleased to welcome Lu Sun to the staff. Lu has just received her BS in Biochemistry from the State University of NY at Stony Brook. She is responsible for answering the Help Desk telephone and e-mail while working on the coordinate entries as part of the Archive Management Group.

Janice Picarelli has been helping the PDB with various projects since April, providing specialized secretarial assistance. She comes to us through Manpower Temporary Services. Janice's expertise in word processing has been and will continue to be put to good use in the next few months.

Five students have been summer interns with the PDB this year. Keith Peters will be a junior Computer Science major at the Califor-

nia Institute of Technology. This is Keith's 4th summer with the PDB. He assists with computer operations, both hardware and software. Jeremy Praissman will be a sophomore majoring in Computer Science at Carnegie Mellon University this fall. Jeremy is writing programs for validation of NMR coordinate entries. Wenjuan Xu is with us for her third summer. She just graduated from Suffolk Community College. Wenjuan designed a new database for tracking orders and provides support for the PDB Web site. Two high school students are also working with us.

Steve Murtagh is writing database programs and Latrice Turpin is providing clerical assistance for the PDB.

PDB Lite Makes Its Debut

Eric Martz

University of Massachusetts, Amherst, MA, USA

(http://www.umass.edu/microbio/rasmol/emartz.htm)

PDB Lite, which first appeared on the PDB Web site in June, offers a simplified path for finding, viewing, and downloading entries from the PDB as an alternative to the more powerful 3DB Browser. At the invitation and with the support and encouragement of Joel Sussman (Head of the PDB), I designed PDB Lite in collaboration with Jaime Prilusky[1], who earlier had developed the 3DB Browser and its database searching infrastructure. With the exception of the Chime-based viewing interface, the implementation of PDB Lite was done by Jaime Prilusky. Some of the results of this collaboration, such as the Chime-based viewing interface, were incorporated into the 3DB Browser as well.

I have had an opportunity to observe many college professors use the PDB on a first-time or occasional basis, which helped me identify some of the issues addressed in PDB Lite. With NSF support, I am conducting 3-day workshops to enable college faculty to incorporate macromolecular visualization (with RasMol or Chime) into their teaching (http://www.umass.edu/microbio/rasmol/workshop.htm). About 100 faculty from over 80 colleges, largely in the northeastern US, are participants.

There are several goals for PDB Lite. Most importantly, we want to make it easy for nonspecialists, occasional users, and students from the high school level on up to find molecules in the PDB. Then, when molecules are found, we want the links which offer information about them to be understandable to this group of users, which may be called the general scientific public. We want to make it easy to view a molecule, with a few generally useful renderings and color schemes, for those not familiar with RasMol or other visualization software. Hence we provided a viewing interface with MDL's Chemscape Chime freeware plugin, an interface kept simple so as to work on all platforms supported by Chime 1 (SGI, Macintosh, Windows 3.1, and Windows 95). Finally, we wanted to enable people not thoroughly familiar with file and folder management to succeed at downloading atomic coordinate files from the PDB. For this latter goal, we provided "click by click" instructions for the three platforms likely to be used by this group (Macintosh, Windows 3.1, and Windows 95/NT). Throughout, we have striven to avoid jargon and unnecessary technical terminology and to make everything as clear and explicit as possible.

PDB Lite is organized into a sequence of pages titled as follows (where xxxx is a PDB ID code). If a single entry matches the search terms, pages 2, 3, and 4 are skipped. If a single entry is selected on page 3, page 4 is skipped.

PDB Lite Page Sequence

1. Find Macromolecules
2. Count of Macromolecules Found
3. Select Macromolecules for Retrieval
4. Short Descriptions
5. View/Analyze/Save xxxx

New pages available from page 5:

How to save file xxxx.pdb to your disk
xxxx.pdb (shown in Chime with image control buttons)

To keep it simple, PDB Lite has a single slot for entering search terms. Under the search slot are three checkboxes for "Crystal diffraction, NMR, Theory" and a fourth, "Non-redundant structures only". Searching for "lysozyme" finds more than 450 entries; doing the same with "Non-redundant structures only" checked finds five. Here is also a button labeled with a question mark which brings up a page titled "Nature of 3D Structural Data". This page has a brief introduction to the methods used for obtaining 3D structures, and the limitations of the resulting data. Under the search slot is a brief explanation of how to phrase a search. Considerable effort was devoted to making this description accurately reflect the behavior of the search engine. Examples given on that page illustrate that placing an asterisk after a term signifies leading as well as trailing ambibuity, *e.g.,* "lacto*" finds not only "lactoferrin" but "muconolactone". Two situations were found which do not consistently give correct results in the present implementation: use of parentheses, or a "loose" term following "not" (such as "milk not lacto*"). When the user enters such a search, a message appears stating that this search phrase is not supported, preventing the user from obtaining incorrect results.

The "Count of Macromolecules Found" page highlights the interpretation of the query, showing the word "and" between search terms and explicitly stating "Experimental Results or Theoretical Models" even when no category was excluded. The display of pending/on-hold entries has been moved to a separate page to avoid confusion, and this category is explained. Terms such as "PDBid" or "Resolution" in the headings for the select-list on the "Select Macromolecules" page are explained. Advanced crystallographic information such as symmetry, R factor, and crystal cell are omitted from the "Short Descriptions" page and the "View/Analyze/Download" page. On the latter page, which lists many links to further resources, an effort has been made to state briefly, in non-technical language, the unique or important type of information offered by each link. The download link now forces a file-save dialog, instead of displaying the PDB file as text in the browser, and for Windows, the default filename for a download is now the PDBid followed by .pdb, *e.g.,* 1ace.pdb, instead of "send-pdb". (Unfortunately you still get "send-pdb" on the Macintosh.)

A number of enhancements are already being planned for PDB Lite and its Chime viewing interface, but while these are being implemented, it is hoped that the existing version of PDB Lite will improve accessibility of the PDB to the general scientific public.

[1] Head of BioInformatics, Weizmann Institute of Science, Rehovot, Israel

Sequence To and withIN Graphics PDB Viewer (STING —— PDB viewer)

Goran Neshich[1], Roberto C. Togawa[1], Wellington Vilella[1], and Barry Honig[2]

[1] *Embrapa/Cenargen, Biotechnology/Bioinformatics, BrazilianBioNet (BBNet), Brazilian Bioinformatics Resource Center, Brasilia, Brazil*

[2] *Columbia University, Department of Biochemistry and Molecular Biophysics, New York City, NY, USA*

STING is a PDB viewer that enables close interaction between sequence and structure windows. STING offers a simple and easy way to map a single amino acid (or nucleotide) letter code to its position in 3D, and *vice versa*. STING offers extensive information about a sequence: residue numbering, sequence gaps and also PDB defined secondary structure regions [Helices (red lines below the sequence) and Extended Sheets (blue lines below sequence)]. Special attention is given in STING to macromolecular interface analysis.

STING is both didactic tool as well as research tool. It is easy to use and requires virtually no training time.

STING has 6 frames:

1. Graphics Frame shows 3D structure.

2. STING's Control Frame displays STING script commands, the purpose of which is to perform a simple structure analysis and also a sophisticated Interface Analysis. It color codes all charged residues or differentially colors all chains, selects and displays ligand, displays ligand and water molecules in vicinity of the ligand, selects and displays ligand pockets and finally, displays the interface between two chains (all these commands are a single stroke = click on STING's Control Frame buttons).

3. STING's Logo Frame offers useful links to other Bioinformatics servers as well and databases.

4. Sting's PDB Info Frame shows information taken from a PDB file about each chain in the file, the number of residues in each chain, the number of ligands and their names, the number of ions and their names and the number of water molecules.

5. STING's Sequence Frame displays the linear protein and/or DNA sequence, color coded according to hydrophobicity and charged groups. The Sequence Frame also shows the numbering of the residues in the sequence, gaps in the PDB sequence, a chain identifier and ranges for secondary structure elements. Each residue in the Sequence frame is "clickable", resulting in a CPK presentation of its position in the Graphics Frame. Blue and red lines below the sequence are also "clickable" resulting in a graphical RIBBON presentation of the specified sequence region.

6. STING's Status Frame shows the residue/nucleotide number and chain identification any time a user slides the mouse over a residue/nucleotide on the Sequence Frame or any time a user clicks the mouse over a certain atom in the Graphics Frame.

Further detail is provided below about STING interface commands.

"Interface on"

The "Interface on" command (button) provides a simple means of getting general information about interfaces between two molecular chains, with a simple mouse stroke. This command will turn on only atoms at an INTERFACE of the first two chains in a PDB file. As of now, this is a hard-wired command; in other words, you will be able to see most interfaces for PDB files having chains A and B, H and L, E and I (which is the most frequently observed case). We are going to implement the ability to choose pairs of chains in the next version of STING. In combination with the command "Color by chain" (issued prior to Interface on) graphical information about interfaces becomes easy to visualize. The interface between two chains is defined based on a distance, set to 8.0 Ångstroms, and measured between any two atoms in different chains. A value of 8.0 Ångstroms was chosen empirically; we first tried a distance of two times 3.3 Ångstroms (the distance from Hydrogen acceptor from one chain, to water molecule, to hydrogen donor on the other chain). This corresponds to a distance of 6.6 Ångstroms, but we found that graphical presentation of the interfaces is much more "complete" if distance of 8.0 Ångstroms is used. We judged completeness by how well an interface is populated by atoms, or by how many holes we have on the chain interface. Obviously, the user should consider Interface graphics presentation more as a guide than as an exact interface definition. The user should also be aware that chain identification in STING's PDB Info Frame follows alphanumerical order. However, appearance of the chains in the PDB file does not follow this order! This is very important to recognize, especially in case of the "Interface ON" STING command, where the first two chains encountered in a PDB file are taken for interface building (the next version of STING will have an option for chain choice). We are currently implementing exact interface definition and its graphical presentation in our package, HORNET to be announced soon, based on buried surface area upon complex formation.

"HOH + Interface"

This command is very useful for the analysis of the interfaces and water molecules captured between Interface Forming Residues (IFR). The availability of such quick identification of these waters may aid in identification of indirect H-bond formation between two chains (obviously with involvement of structural water molecules). HOH molecules visualized by this command are identified here by a distance, set to 3.3 Ångstroms, and measured between a subset of atoms belonging to the interface from one chain, to any HOH molecule (actually, in most cases, an Oxygen atom). This is then done for the other chain (defining its IFR and HOH molecules at the determined distance of 3.3 Ångstroms). Finally, the visualized waters correspond to an intersection of water molecule ensembles defined above. In other words, the only water molecules presented (color coded magenta), are those that satisfy the geometric condition of being 3.3 Ångstroms (maximum) distance from both chains. These HOH molecules are likely to make H-bonds with both chains. The user can easily see a subset of these HOH molecules within the frame of only one interface half by using the option: "Interface: 1st half" and "Interface: 2nd half".

"Interface: 1st half"

This option allows the user to examine in detail only half of a complementary surface (see example in STING's tutorial section). Again, this is a hard-wired scripting (since we do not have a chain independent command, as we do not have it for the "Interface on" button), and so, this feature will not work for chain names other than the first two encountered in a PDB file.

"STINGpaint"

STINGpaint is a part of the STING package and it was developed to allow the presentation of residue characteristics in the Sequence Frame (color coding of residues with respect to their hydrophobic-

ity and charge). As a consequence, during development of the STING project, we have slightly expanded on the STINGpaint idea and adopted it for use with Multiple Sequence Alignment (MSA) coloring. STINGpaint is a part of our ongoing work on "MSA_STING", a package that will be able to show both sequence alignments (in the Sequence Frame) and structures (in the Graphics Frame) for respective sequences. STINGpaint now supports the following sequence and MSA formats: Coloring sequence of any PDB entry, Coloring any sequence in FASTA format, Coloring MSA in PRISM output format (PRISM is a sequence/structure /threading/homology modeling program developed by An-Suei Yang in Barry Honig's laboratory), Coloring MSA in PSI-BLAST output format, and Coloring MSA in GCG output format.

Work in progress:

1. MSA-STING, with multiple structure graphics presentation as well as with capacity to perform and display Multiple Sequence Alignments (MSA), is being developed.

2. HORNET, an important step in STING development, is now in the testing phase. This package will offer hydrogen bond network analysis graphical details.

3. Downloading STING. This will be available soon. Individual users will be able to download STING and install it on their UNIX workstations, PC or Mac platforms. The STING version that users would be able to download and install on a local machine would respond to the need to be able to read local PDB format files (such as model structures). In addition, users would be able to access PDB files over the network (without having to store one individual copy of the whole PDB database). Locally installed STING will also have an advantage with respect to speed. Possible delays, due to network performance, might be something users would like to avoid (especially when working on analysis of model structures when the network is not used). The STING version for local installation will also have some extra features such as the possibility of choosing chains for interface analysis, choice of background colors for the Graphics Frame, etc.

For further information, consult http://trantor.bioc.columbia.edu/STING/ or http://www.cenargen.embrapa.br/bbnet.

Searching for the Right PDB File with the 3DB Browser™ and Displaying Structure and Sequence Data with STING

Goran Neshich[1], Roberto Togawa[1], Jaime Prilusky[2], Joel Sussman[3,4], and Barry Honig[5]

[1] Embrapa/Cenargen, Biotechnology/Bioinformatics, BrazilianBioNet (BBNet), Brazilian Bioinformatics Resource Center, Brasilia, Brazil

[2] Bioinformatics Unit, Weizmann Institute of Science, Rehovot, Israel

[3] Department of Structural Biology, Weizmann Institute of Science, Rehovot, Israel

[4] Protein Data Bank, Brookhaven National Laboratory, USA

[5] Columbia University, Department of Biochemistry and Molecular Biophysics, New York City, NY, USA

STING is a WWW tool for the simultaneous display of information about macromolecular structure (in STING's Graphics Frame) and sequence (in STING's Sequence Frame). The user interested in viewing a particular PDB structure can get the extensive information about the particular protein and/or DNA structure including the underlying sequence, the structural water molecules, the ligands and also information about the macromolecular interfaces. STING offers a simple and easy way to map an amino acid (or nucleotide) single letter code to its position in 3D, and vice versa.

The user can either input the four letter PDB ID code and start using STING's features, or perform a search of the PDB based on various search parameters. To this end an effective link has been established between STING and the 3DB Browser. The essential features of this link are summarized below.

1. STING uses the 3DB Browser as the primary search tool for identifying the PDB file to be analyzed.

2. The 3DB Browser recognizes that the search request has come from STING and acknowledges this by displaying STING's logo (honeybee). After a desired PDB file has been located, the 3DB Browser automatically provides the option "STING IT" which will open STING.

3. The 3DB Browser offers, independent of STING, the option to view and analyze a PDB file with STING. The STING option in this case is one of the items listed among many other packages for viewing and analyzing PDB structures.

4. The 3DB Browser also offers a distinct option: "STING IT" at the end of any search.

Further information about the 3DB Browser and STING can be found, respectively at
http://www.pdb.bnl.gov/pdb-bin/pdbmain (main entry page),
http://www.pdb.bnl.gov/pdb-docs/3DB_Browser_help.html (help page),
http://trantor.bioc.columbia.edu/STING (main entry page), and
http://trantor.bioc.columbia.edu/STING/help (help page).

Pattern Searching over the TOPS Protein Topology Database

David R. Gilbert[1,4], David R. Westhead[1], Daniel C. Hatton[1], and Janet M. Thornton[1,2,3]

[1] EMBL Outstation: European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK

[2] Biomolecular Structure and Modeling Unit, Department of Biochemistry and Molecular Biology, University College, London, WC1E 6BT, UK

[3] Laboratory of Molecular Biology, Department of Crystallography, Birkbeck College, University of London, Malet Street, London, WC1E 7HX, UK

[4] Department of Computer Science, School of Informatics, City University, Northampton Square, London, EC1V 0HB, UK

We have developed a system that supports fast pattern searching over TOPS protein topology databases; the facility is available at the TOPS protein structural topology WWW site[1] at http://tops.ebi.ac.uk/tops. The search engine is based on a string-graph algorithm and uses constraint propagation to prune the search space. Users can search on motifs from a library or define their own search patterns. All 15400 descriptions of protein domains currently in the Brookhaven Protein Data Bank (April 1998) have been converted into a database of TOPS diagrams (11MB). Average

match times are 3-5ms per diagram on a Dec-Alpha. For example, it takes 80 seconds to find all matches to a jelly-roll type 1 description (500 hits out of 15400) on a Dec-Alpha. Some queries take advantage of precompiled topological information generated by the TOPS program; a search for Tim barrels takes 15 seconds to find 260 matches in the database. Load times are additionally 3 sec for the atlas and 32 sec for the PDB. Users can search on motifs from a library or define their own search patterns. We are now in the process of enhancing the system to permit users to compare the topology of a given domain with all the other domains in a database.

A note on TOPS diagrams and patterns

A TOPS *diagram* describes more information about a structure than the corresponding cartoon, or graphical representation (we have described TOPS cartoons in details in an earlier article in this Newsletter[2]). In addition to the Secondary Structure Elements - their type, position on the backbone, and orientation relative to the plane of the page, the diagram describes the H-bonds (parallel or anti-parallel) between strands, and the chiralities of the connections (right or left) between some of the SSEs. For example, the TOPS diagram for 2bop is shown below; it comprises five strands, which form a bifurcated anti-parallel sheet, and three helices. There are two right-handed chirality connections, between strands 1 and 4, and between strands 6 and 8 respectively *(See Figure 1).*

A TOPS *pattern* is like a TOPS diagram, and may describe several (none or more) TOPS diagrams. This is achieved by permitting insertions of some SSEs into the pattern to obtain a diagram. Each backbone connection between every pair of SSEs is annotated with a pair of integers standing for the minimum and maximum number of SSEs that can be inserted. The values of min and max can range from 0 to a large number (in practice 60) which we denote N. For example, a plait pattern (or motif), which matches amongst others 2bop, is shown in Figure 2.

This pattern describes, amongst others, 2bop. In order to illustrate this, consider the plait pattern and the diagram for 2bop to each be 'stretched out' and laid side by side shown in Figures 3 and 4.

Matching a pattern to a diagram

Informally, we match a pattern to a diagram by matching on the SSEs, the H-bonds and the chiralities. When matching on the SSEs we obtain a correspondence. This is a sequence of matching pairs of SSEs, one for each SSE in the pattern, where the first member is an SSE from the diagram and the second is an SSE from the pattern. We can describe the result of matching by the correspondence and also a list of the total number of inserts between adjacent members of the SSE sequence in the diagram. A pattern may match a diagram in none or more ways, and may match none or more diagrams in a given database of diagrams.

For example, there are two ways in which the plait motif can match the diagram for 2bop:

- Character matches: (1,1), (2,2), (4,3), (6,4), (7,5), (8,6). Inserts: (0,1,1,0,0)
- Character matches: (1,1), (3,2), (4,3), (6,4), (7,5), (8,6). Inserts: (1,0,1,0,0)

Using the TOPS query system

The TOPS query system is accessed over the Web using an HTML form, and there is on-line help.

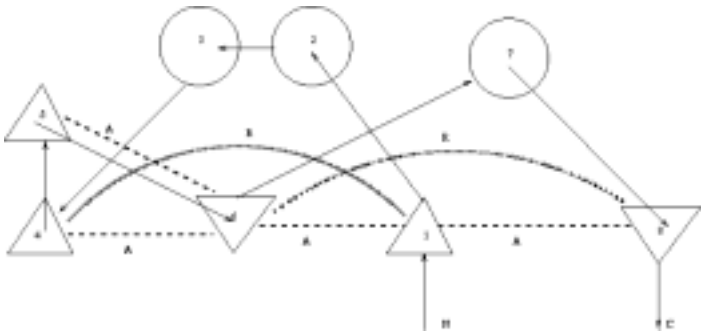


Figure 1. TOPS diagram for 2bop

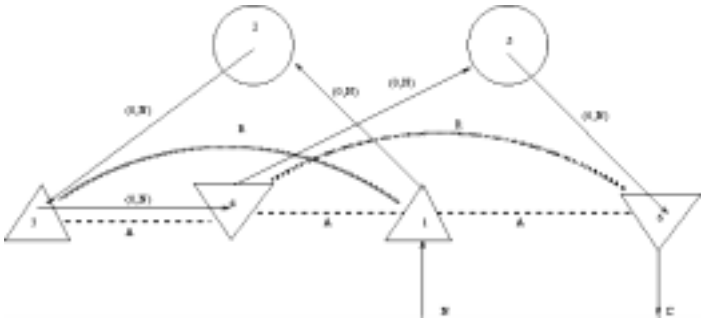


Figure 2. Plait pattern

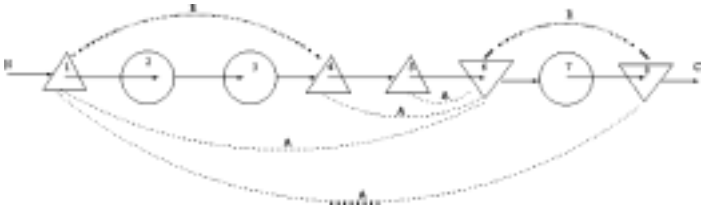


Figure 3. TOPS diagram for 2bop

Figure 4. Plait pattern

Users can select which database (Atlas or PDB) to search, and optionally restrict the search to one given domain. Searches can be sped up by filtering on TOPS precompiled fixed structure information about targets (*e.g.,* if they contain barrels, sheets of various curvatures, or sandwiches).

Users can also select several output parameters, the order in which they are output for each successful match, and whether to sort on these. Output is incremental unless sorting is selected. Output parameters include domain name, CATH-number, matching node numbers, and sum of inserts between the nodes. There is an HTML link for each domain found by the search to the TOPS Atlas, which may hold the cartoon (graphical) representation of the domain or a representative domain, and also to the CATH database[3] at UCL. Predefined queries can be made for Greek keys, jelly rolls, NAD-binding domains, immunoglobulins, plaits, barrels or various types, trefoils and propellers (or all of these). Alternatively, users can construct their own query patterns based on the sequence of SSEs and associated inserts, and the associated sets of H-bonds and chiralities. There are facilities to define constraints over the total number of inserts, and parallel/anti-parallel H-bonds.

Acknowledgements

The authors wish to thank Alvis Brazma of the EBI for his invaluable suggestions, and the CATH group at UCL, in particular Nozomi Nagano for her painstaking testing of the system and her proposals for improving the user interface.

David Gilbert has been supported by City University, and also by an EPSRC Visiting Fellowship.

References

1. Westhead, D.R., Hatton, D.C., and Thornton, J.M. (1998). *Trends in Biochemical Sciences* **23**, 35-36.

2. Westhead, D.R., Hatton, D.C., and Thornton, J.M. (1998). *PDB Quarterly Newsletter* **83**, 8.

3. Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M.B. and Thornton, J.M. (1997). *Structure* **5**, 1093-1108.

BIOINFORMATICS: Creation of the Swiss Institute of Bioinformatics (SIB)

From the Swiss Institute of Bioinformatics (SIB)

We are pleased to announce the creation of the Swiss Institute of Bioinformatics (SIB) on March 30, 1998.

The SIB is an academic not-for-profit foundation whose mission is to promote research, development of databanks and computer technologies, and to teach and service activities in the field of bioinformatics. The Institute is tightly associated with the University of Geneva, the University of Lausanne, the Swiss Institute for Cancer Research (ISREC), the Ludwig Institute, and Glaxo Wellcome Experimental Research (GWER). It will also operate in association with industrial partners.

The Institute is located in Geneva and Lausanne, starting from five bioinformatics research groups, whose leaders have a well-established international reputation. They include Dr. Ron D. Appel, Director of the Molecular Imaging and Bioinformatics Laboratory at Geneva University Hospital, Dr. Amos Bairoch, Group Leader and Head of the Biocomputing Unit at ISREC, Dr. Victor C. Jongeneel, Director of the Swiss node of EMBNet and of the Office of Information Technology for the Ludwig Institute for Cancer Research, and Dr. Maneul Peitsch, Head of Glaxo Wellcome Experimental Research (GWER) and Director of Scientific Computing, Glaxo Wellcome Research and Development.

The ExPASy molecular biology server, its databases (SWISS-PROT, PROSITE, SWISS-2DPAGE, SWISS-3DIMAGE, ENZYME, CD40Lbase and SeqAnalRef) and software herein, will be managed by the SIB from now on. Academic users will still have free access to any of these products that are partially or completely funded by public grants. The Institute will seek to license its products to commercial users.

All of the SIB groups are expected to continue to develop and enhance their current research activities as well as the many collaborative projects in which they are already involved. The development of the SWISS-PROT protein sequence database is part of such a long term and extremely close collaboration with the European Bioinformatics Institute (EBI) in Hinxton, UK, an outstation of the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany. It is expected that this collaboration will not only be maintained, but also strengthened and broadened. The fact that the Swiss node of EMBNet will be part of SIB is also a good indication of the deep involvement of the SIB in the European bioinformatics infrastructure. Existing collaborations with groups in the USA, such as the National Center for Biochemistry Information (NCBI) at the NIH in Bethesda, Maryland and with many other academic institutions will also be maintained. The same principle applies to existing contacts with Australian and Asian research groups.

The grouping of the five groups in a single organization will foster unprecedented synergy. Thanks to a close collaboration with industrial partners we expect a scaling up of the activities toward the development of integrated databases and software resources in the field of Proteomics. We are confident that SIB will be one of the jewels in the crown of Swiss science.

WWW server: http://www.isb-sib.ch/

For information: info@isb-sib.ch

Development of a New Program for Structure Determination: Crystallography & NMR System

Axel T. Brunger

Department of Molecular Biophysics and Biochemistry, Howard Hughes Medical Institute, New Haven, CT, USA

Two years ago, development was started on a new program for structure determination called Crystallography & NMR System. This program is the result of an international collaborative effort among several research groups. The program has been designed to provide a flexible multi-level hierarchical approach for the most commonly used algorithms in macromolecular structure determination. Highlights include heavy atom searching, experimental phasing (including MAD and MIR), density modification, crystallographic refinement with maximum likelihood targets, and NMR structure calculation using NOEs, J-coupling, chemical shift, and dipolar coupling data.

Crystallography & NMR System will be made available to both academic and commercial users. The program will be provided to academic users with a small administrative fee and to commercial users through a yearly licensing scheme which will support a non-profit support and development group. This non-profit support and

development group, headed by Dr. Paul Adams, has already been initiated at Yale University. Other members of the group currently include Dr. Ralf Grosse-Kunstleve. It is expected that other positions will be added later.

Announcement of the official release of Crystallography & NMR System will be made on the Internet as soon as it is available.

References

Brunger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L. (1998). Crystallography and NMR System (CNS): A New Software System for Macromolecular Structure Determination, *Acta Cryst. D* **54**, 905-921.

HIC-Up Updated

Gerard J. Kleywegt

*Department of Molecular Biology, Uppsala University, Uppsala, Sweden (gerard@xray.bmc.uu.se)**HIC-Up Updated*

The Hetero-compound Information Centre - Uppsala (HIC-Up) Web resource has been updated. It now contains information pertaining to 1,649 hetero-entities encountered in the Protein Data Bank. Besides this, HIC-Up offers several Web-based services for generating dictionaries for other compounds for X-PLOR, CNS, TNT, and O.

Changes since the last update (1997/12/12):

- 216 new compounds added
- server to generate O-style Refi dictionaries (requires O 6.2.2 or newer)
- file naming convention changed (now most files end in ".txt" to prevent accidental MIME-type clashes; also, PDB files are available with ".pdb" extension and ".txt" extension)

The HIC-Up URL is http://alpha2.bmc.uu.se/hicup/

Notes of a Protein Crystallographer — Genomics, Proteomics, and the Secret of Life: A Faustian Dialog

Cele Abad-Zapatero

Abbott Laboratories, Department of Structural Biology, Abbott Park, IL, USA (abad@abbott.com)

The publication of the entire genome of *Mycobacterium tuberculosis* (Cole *et al*., 1998), the tubercle bacillus responsible for the death of so many members of the human species, represents another milestone in the development of what has been called experimental genomics. In a broad sense, this term refers to the efforts to systematically obtain the complete nucleotide sequence of the genome of important and relevant organisms of the earth ecosystem. The ultimate goal is to complete the entire human genome by 2005. The related areas of functional and computational genomics will complement and enhance the impact of the data gathering efforts.

It has been only three years since J. Craig Venter published the random sequencing and assembly of the first complete DNA from the free living bacterium *Haemophilus influenzae* completed by the novel "shotgun cloning" (Fleischman *et al*., 1995). Since then, we have seen in rapid succession entire genomes of several microorganisms representing important biological classes ranging from *E. coli* (Blatter *et. al*., 1997) to *B. subtilis* (Kunst *et. al*., 1997) and to *S. cerevisiae* (Goffeau *et. al*., 1997).

A glimpse at the figures that keep appearing from the complete projects is sobering. For example, the entire genome of *B. subtilis* (an important member of the class of Gram-positive bacteria) contains 4,214,810 base pairs. They code for approximately 4,100 proteins, 42% of which have unknown functions. Among those encoded proteins there are 18 F transcription factors and 77 ATP-binding transport proteins. Many of these new sequences represent important biotechnology findings related to the production of antibiotics (Hoch & Losick, 1997).

The amount of information, the quantity and quality of our knowledge of the biological systems that surround us was unimaginable only a few years ago. The idea of obtaining the three dimensional structure of all (or a sizable subset) of the proteins coded by these millions of bases could wet the appetite of the novice crystallographers and sharpen the tools of the well-seasoned ones (see for example Chayen & Helliwell, 1998; Pennisi, 1998). The possibility of structurally mapping the majority of the macromolecular structures playing critical roles in the human body appears to be within reach in the not too distant future.

Undoubtedly, the knowledge so gained will be extremely valuable for finding novel antibiotics, for understanding the molecular basis of many known and unknown diseases in animals and men, and for designing potent and efficacious drugs against them. Detailed knowledge of the genomes of various microorganisms will pave the way to understanding rare metabolic pathways that could help in solving environmental problems. These are just a few of the immediate applications. One can make conjectures as to the unexpected findings.

All of the futuristic vistas notwithstanding, I can see myself clad in a white robe as a young, Faustian, thirsty for insight, macromolecular crystallographer in front of Mephistopheles having the following conversation:

> I would like to know the secret of life.
>
> I presume you mean beyond the DNA double helix and the genetic code.
>
> Yes of course. Those are past history.
>
> What do you want to offer in return?
>
> I am willing to sacrifice love or happiness for the ultimate knowledge.
>
> Do love and happiness not accompany each other?
>
> Fine! Let us not go into details. I'll put both on the scale.
>
> You macromolecular crystallographers have your own view of what the secret of life is.
>
> I suppose we do. I'll be more specific. Can you put at my fingertips, the three-dimensional structure of all the macromolecular components encoded in the *E. coli* genome?
>
> Of course I can but remember your personal love and happiness must rest on the other side of the scale. And what is the resolution you wish?
>
> Good point. Refined at least to 1.2 Ångstroms resolution.
>
> Certainly, in exchange for your unhappy existence, totally deprived of love. Are you ready? Are you sure of what you ask of me?
>
> Why do you keep insisting on the conditions? What do you

mean am I sure? Are you not supposed to entice me?

I mean that if I were to sacrifice my love and my happiness for the secret of life, I'd better be sure of what I really understand by that.

Am I sure? Are we sure? Albert Szent-Györgyi, the Hungarian biochemist who isolated Vitamin C from extracts of paprika, wrote the following parable in the preface to one of his books:

> If you give a dynamo to a chemist, the first thing he will do is to submerge it in HCl and analyze what substances are deposited during and after the reaction, and which gases are given away. If you give it to a molecular biologist, he will take it apart, disassemble it, characterize each and every one of its parts and then he will put it back together again. Now, he argued, if you were to point out to these scientists that the dynamo works because of changes with time of something called "magnetic flux", they will call you a "vitalist". A person who needs to invoke an "élan vital" to explain biological phenomena.

I truly believe the book is *"Introduction to a Submolecular Biology"* (1960). I apologize to the reader because much to my disappointment, I have not been able to find the complete quote and citation of this comparison. However, what I wrote reflects the essence of the text as it appears in my old notes.

Any person contemplating a musician playing an instrument can certainly relate to the following analogy. We can take a musical instrument (*e.g.*, an oboe) from the hands of the musician playing it. We can disassemble it and make a detailed analysis of each individual part; then, we can put it back together again. Our language might betray us when we leave the instrument on the table saying that the instrument is "lifeless" without the musician playing it. In fact, it is not the person as a unique demiurge that originates music. Rather, it is the flow of air circulating within the cavities and interstices of the instrument, as channeled and diverted by the action of its valves and keys, that produces the sound that we call music.

Similarly, as important as the molecular components are, we should not forget the interplay of flows and forces that support life. The currents of multiple ions, protons and electrons; the pressures created by chemical gradients; the pulses of electric currents and polarization potentials; the molecular migrations and myriad of feedback loops. All are critical elements of the living cell, it is precisely the cessation of these fluxes and rhythms that mark the absence of life. Our integral and constitutive parts will remain long after the ephemeral wind of life has ceased. Our challenge is to understand how our beautiful macromolecular structures permit, facilitate and maintain that fragile and intangible state that we call living.

References

Fleischman, R.D., *et. al.*, (1995). Whole-Genome Random Sequencing and Assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496-512.

Blattnet, F.R., *et. al.*, (1997). Science **277**, 1453-1462.

Kunst, F., *et. al.*, (1997). The Complete Genome Sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249-256.

Goffeau, A. *et. al.*, (1997). The Yeast Genome Directory. *Nature* **387**, 5-105.

Hoch, J.A., Losick, R. (1997). Panspemia, spores and the *Bacillus subtilis* Genome. *Nature* **390**, 237-238.

Chayen, N.E. & Helliwell, J.R. (1998), *Physics World*, **11**, no. 5, 43-48.

Pennisi, E. (1998), *Science*, **279**, 978-979.

Letters to the Editor

Dear Editor:
*In the April 1998 PDB Newsletter, on page 7 in his article "Notes of a Protein Crystallographer - Molecular Docking and the Broken Heart", Dr. Cele Abad-Zapatero makes a statement I must take issue with: Citing a paragraph of Jeanette Winterson's novel "Written on the Body", Dr. Abad-Zapatero claims: "The soliloquy comes from the brain of the main character in the novel, as *SHE* debates whether staying close....". Now, one of the most intriguing qualities of Jeanette Winterson's novel is, that at no place is there a statement made as to the actual gender of the main character. There may be ambiguous hints, indications, and numerous games to be played by our fantasy, but whatever we conclude as far as the gender of the poor lover is concerned, the result will be a matter of interpretation based on our very own personal experiences and mindset. So, let science be science, and literature be literature. Otherwise we might fail to do justice to either.*

With sincere regards,
Bernhard Rupp
Macromolecular Crystallography Group
Biology and Biotechnology Program
Lawrence Livermore National Laboratory

Dear Dr. Rupp,
Thank you very much for your pertinent comment regarding my article entitled "Molecular Docking and the Broken Heart". I am certainly familiar enough with the work of Ms. Winterson to realize that the main character of "Written on the Body" is "genderless". I debated how to phrase the specific sentence that you refer to and, in the end, I decided to write it as a female; as if coming from the author herself. Yes, it is inaccurate. The alternative would have been to write the text as:

"The soliloquy comes from the brain of the main character in the novel, as s(he) debates whether staying close to her (his) lover will heal her (his) broken heart or could, in fact, result in an expensively ruinous experiment"[...]

Although more precise, I thought it was a bit clumsier in the literary sense. Besides, since I never declared the name of the lover, the quote might be taken as being sexless from the other side.

As for the second comment in your letter, I would like to point out that it was Ms. Winterson who related her feelings to the scientific problem of docking molecules. The issue of the relation between science and literature is too complex to be discussed here. In the future, we might meet in person and we can have a full discussion over a cup of coffee. Thank you for your comment. Best regards.

Cele Abad-Zapatero

PDB Poetry: 1BAD Story

Contributed by Jaime Prilusky, BioInformatics Unit, Weizmann Institute of Science, Rehovot, Israel

Overheard at a pet shop:

'I would like to buy 1PET.'

'On sale today, we have 5CTS, 4APE, 1DUC for 4CTS, or 1DOG for 3CTS.'

'Humm, I wanted 1FAT fox.'

'Good! We have 1FOX for 6CTS, fat, but 2BAD.'

Web Sites

Referenced in this Newsletter

3DB BrowserTM
http://www.pdb.bnl.gov/pdb-bin/pdbmain

AutoDep 2.1
At Brookhaven: http://www.pdb.bnl.gov:8080
At EBI: http://autodep.ebi.ac.uk/

CATH Protein Structure Classification
http://www.biochem.ucl.ac.uk/bsm/cath/

Chime Software
http://www.mdli.com/download/chimedown.html
http://www.mdli.co.uk/download/chimedown.html

Embrapa/Cenargen Bioinformatics
http://www.cenargen.embrapa.br/

HIC-Up
http://alpha2.bmc.uu.se/hicup/

International Union of Crystallography
http://www.iucr.ac.uk/

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

PDB Lite
http://www.pdb.bnl.gov/pdb-bin/pdblite

PDB Mirror Sites
http://www.pdb.bnl.gov/pdb-docs/mirror_sites.html

PDB Quarterly Newsletters
http://www.pdb.bnl.gov/pdb-docs/newsletter.html

RasMol & Chime: Molecular Visualization Freeware
http://www.umass.edu/microbio/rasmol/

STING
http://trantor.bioc.columbia.edu/STING/
http://www.pdb.bnl.gov/STING

Structure Factor Files
http://www.pdb.bnl.gov/pdb-bin/ftp_index.pl?dir=ftp/structure_factors

Swiss Institute of Bioinformatics
http://www.isb-sib.ch/

TOPS Protein Topology Home Page
http://tops.ebi.ac.uk/tops/

Validation for Layered Release
http://www.pdb.bnl.gov/pdb-docs/validation.html

WHAT IF Home Page
http://swift.embl-heidelberg.de/whatif/

WHAT_CHECK Explained
http://www.sander.embl-heidelberg.de/rob/checkhelp/

What is Layered Release?
http://www.pdb.bnl.gov/pdb-docs/what_is_LR.html

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://www.expasy.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 Cores	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://bioinfo.mbb.yale.edu/MolMovDB/
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://shelx.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/

Affiliated Centers and Mirror Sites

Forty-two 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. Official PDB Mirror Sites are marked with a grey bar ( ) and are listed with their sponsoring center.

ARGENTINA

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)

AUSTRALIA

ANGIS
The Australian National Genomic Information Service
University of Sydney
Sydney, Australia
Shoba Ranganathan (61-2-9351-3921)
shoba@angis.org.au
http://www.angis.org.au

PDB Mirror Site:
http://molmod.angis.org.au/pdb/
Shoba Ranganathan (mmnclp@angis.org.au)

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

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

BRAZIL

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@cenapad.ufmg.br
http://www.1cc.ufmg.br/

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

CANADA

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

CHINA

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)

FINLAND

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

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

FRANCE

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

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

PDB Mirror Site:
http://pdb.igh.cnrs.fr/
Denis Pugnére (pdbhelp@igh.cnrs.fr)

GERMANY

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

GMD
German National Research Center for Information Technology
Sankt Augustin,Germany
Theo Mevissen (49-2241-14-2784)
theo.mevissen@gmd.de
http://www.gmd.de

PDB Mirror Site:
http://pdb.gmd.de
Theo Mevissen (theo.mevissen@gmd.de)

MPI
Max Planck Institute for Biochemie Computer Center
Martinsried, Germany
Wolfgang Steigemann (49-89-8578-2723)
steigemann@biochem.mpg.de
http://www.biochem.mpg.de

INDIA

PUNE
Bioinformatics Center University of Pune
Pune, India
A. S. Kolaskar (0212-355039-350195)
Kolaskar@bioinfo.ernet.in
http://bioinfo.ernet.in

PDB Mirror Site:
http://202.41.70.33/
A.S. Kolaskar (kolaskar@bioinfo.ernet.in)
Sunita Jagtap (sunita@bioinfo.ernet.in)

ISRAEL

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
Marilyn Safran
(pdbhelp@pdb.weizmann.ac.il)

ITALY

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

JAPAN

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

*JAICI

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

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

THE NETHERLANDS

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

POLAND

WARSAW UNIVERSITY
Interdisciplinary Centre for Modelling
Warszawa, Poland
Wojtek Sylwestrzak (48-22-874-9100)
W.Sylwestrzak@icm.edu.pl
http://www.icm.edu.pl

PDB Mirror Site:
http://pdb.icm.edu.pl
Wojtek Sylwestrzak (W.Sylwestrzak@icm.edu.pl)

SWEDEN

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

TAIWAN

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

UNITED KINGDOM

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

*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)

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)

*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

UNITED STATES

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

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

PDB Mirror Site:
http://www.pdb.bu.edu/
Esther Epstein (esther@darwin.bu.edu)

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

*MAG
Molecular Applications Group
Palo Alto, California, USA
Margaret Radebold (650-846-3575)
bold@mag.com
http://www.mag.com

*MSI
Molecular Simulations Inc.
San Diego, California, USA
Stephen Sharp (619-799-5353)
ssharp@msi.com
http://www.msi.com

NCBI
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

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

PDB Mirror Site:
http://pdb.ncsc.org/
Linda Spaminato (info@ncsc.org)

*PANGEA SYSTEMS, INC.
Oakland, CA 94612
Greg Thayer (510-628-0100)
gregt@pangeasystems.com
http://www.pangea.com

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

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

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://pdb.bmb.uga.edu
John Rose (rose@BCL4.biochem.uga.edu)

This page intentionally left blank.

PDB ™ 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: Chase Manhattan Bank
Account name: Brookhaven Science Associates, LLC
Brookhaven National Laboratory
Account number: 615-775942

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

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

One (1) release of the PDB™ on 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
Entry Error Reports ..... errata@pdb.pdb.bnl.gov
Order Information ..... orders@pdb.pdb.bnl.gov
User Group ..... jmoult@carb.nist.gov
Listserver Subscriptions ..... listserv@pdb.pdb.bnl.gov

To subscribe, the text of your message should be **subscribe PDB-L Your Name**
subscribe NEWSLETTER-L Your Name
for PDB Mailing List or Newsletter Mailing List, respectively

FTP Directory Structure for Entries

The PDB FTP server is updated weekly. Files are available by anonymous ftp to ftp.pdb.bnl.gov and on the Web at http://www.pdb.bnl.gov.

The PDB releases entries during the early morning hours each Wednesday, New York time. Entries that have been placed on hold by their authors are made available on the first Wednesday following their hold expiration date.

Entry files as 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

layer1
layer 1 entries in compressed and uncompressed format

layer2
layer 2 entries in compressed and uncompressed format

ndb
entries received from NDB in compressed and uncompressed format

newly_released/
entries released since the last CD-ROM

nmr_restraints/
compressed NMR restraint files

obsolete_entries/
withdrawn and/or replaced entries

reports
all report files

structure_factors/
compressed structure factor files

current_release, fullrelease, layer1, layer2, ndb, and newly_released are divided into multiple subdirectories

Scientific Consultants

John P. Rose, *University of Georgia, Athens, Georgia, USA*
Mia Raves, *Utrecht University, The Netherlands*
Barry Honig, *Columbia University, New York City, NY*
Goran Neshich, *Embrapa/Cenargen and BBNet/BBRC, Brazilia, Brazil*
Gert Vriend, *European Molecular Biology Laboratory, Heidelberg, Germany*
Manfred Hendlich, *University of Marburg, Germany*
Jiri Vondrasek, *Czech Academy of Sciences, Prague, Czech Republic*
Alexander Wlodawer, *NCI-FCRDC, Frederick, MD*
Eric Martz, *University of Massachusetts, Amherst, MA*
Peter Murray-Rust, *University of Nottingham, Nottingham, UK*
Eldon L.Ulrich, *University of Wisconsin - Madison, Wisconsin, USA*

Clifford Felder
Kurt Giles
Harry M. Greenblatt
Jaime Prilusky
Marilyn Safran
Vladimir Sobolev

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*

Betty R. Deroski
Arthur Forman
Sabrina Hargrove
Jiansheng Jiang
Jiri Koutnik
Patricia A. Langdon
Michael D. Libeson
Dawei Lin
Nancy O. Manning
John E. McCarthy
Christine Metz
Michael J. Miley
Steve Murtagh
Keith Peters
Janice Picarelli
Jeremy Praissman
Regina K. Shea
Janet L. Sikora
Lu Sun
S. Swaminathan
Latrice Turpin
Wenjuan 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-AC02-98CH10886) and user fees.

Instructions to Authors

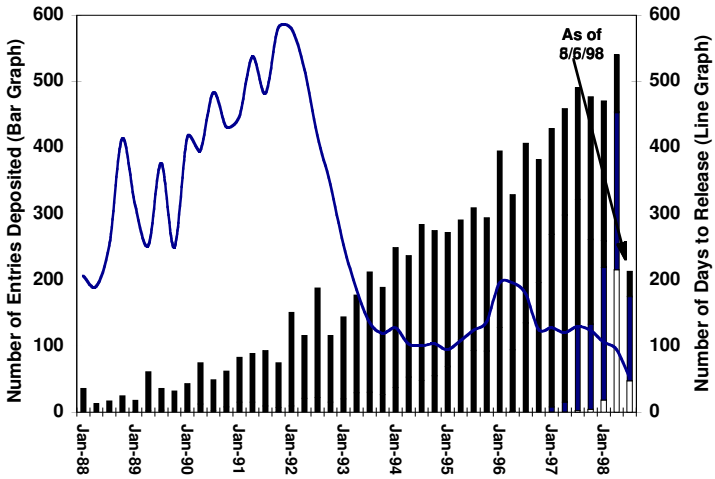
Contributions to the PDB Quarterly Newsletter may be sent by e-mail or diskette to:
Nancy O. Manning, *Editor*
oeder@bnl.gov
References should be in the format used by the Journal of Molecular Biology.
Images must be in Tiff format, 300dpi, and sized as they will appear in print.
Deadlines for contributions are:
March 1, June 1, September 1, and December 1.

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

First Class

Number of Entries Deposited (Bar) and Average Time to Release (Line)

Accumulated and Averaged on a Quarterly Basis



Bar Graph - Number of Entries in the Following Categories:

OnHold	*- (light blue) On-hold per depositor request*
Processing	*- (white) Being processed*
Released	*- (black) Released*

Line Graph - Average Number of Days to Release
The data were accumulated and averaged on a quarterly basis. The average turn-around times for entries now being processed are estimated based on the average of the last 12 months.

Data for the last quarter are accumulated until the date specified on the graph.

See http://www.pdb.bnl.gov/pdb-docs/EntryTurnAround.html for regularly updated plot.

PRINTED ON RECYCLED PAPER