Acta Crystallographica Section D
**Biological Crystallography**
ISSN 0907-4449

# Xtrack – a web-based crystallographic notebook

**Mark Harris and T. Alwyn Jones***

Department of Cell and Molecular Biology,
University of Uppsala, Box 596,
751 24 Uppsala, Sweden

Correspondence e-mail: alwyn@xray.bmc.uu.se

Xtrack is a web-based database-backed crystallographic notebook system that allows groups of workers to keep track of crystallographic project data from the crystallization stage through to data bank deposition. A working demonstration version is available from http://xray.bmc.uu.se/xtrack and the code is available for those who wish to set up their own databases.

Received 3 May 2002
Accepted 16 July 2002

## 1. Introduction

Protein crystallographic projects are often worked on by many people over long periods of time and information required for data bank deposition or for re-evaluating earlier decisions is often lost or forgotten. The ability to collect large numbers of data sets in a short time at remote synchrotron sites compounds the problem. To address these issues, an easy-to-use web-based database system was developed to keep track of projects. During the development of the system, another crystallographic database was reported (Haebel *et al.*, 2001).

## 2. Design

The main design aim was to develop a system that was extremely simple to use for beginners and infrequent users, and to have an interface that was not intimidating. It was important to make the system at least as easy to use as a hardcopy notebook, otherwise people would continue to scribble on paper with great plans to transfer the information to the database at some ill-defined later date. To this end, the power of the database query language (SQL) was hidden from users in favour of presenting them with a small number of well defined options and a standard set of data fields. However, in order to be able to develop and expand the system, it was decided that those standard options and fields should not be hard-coded in the program, but should themselves be elements in the database, so that an administrator could easily add new features and fields using standard SQL commands without interrupting database operation. Presentation of data elements associated with a particular collection should be split up into logical pages containing a manageable amount of information and the system should have facilities for extracting data from log files produced by widely used crystallographic programs. To avoid overwhelming the user, only information that is needed for deposition, publication or synchrotron-use reports should be stored, plus information that is not otherwise recorded automatically in files (such as wavelength and straight-through position). Links to data files should be used to provide access to other information. The system should also be protected so that unauthorized users cannot access any data and only registered users assigned to a particular group should be able to read and modify data belonging to that group.

## 3. Implementation

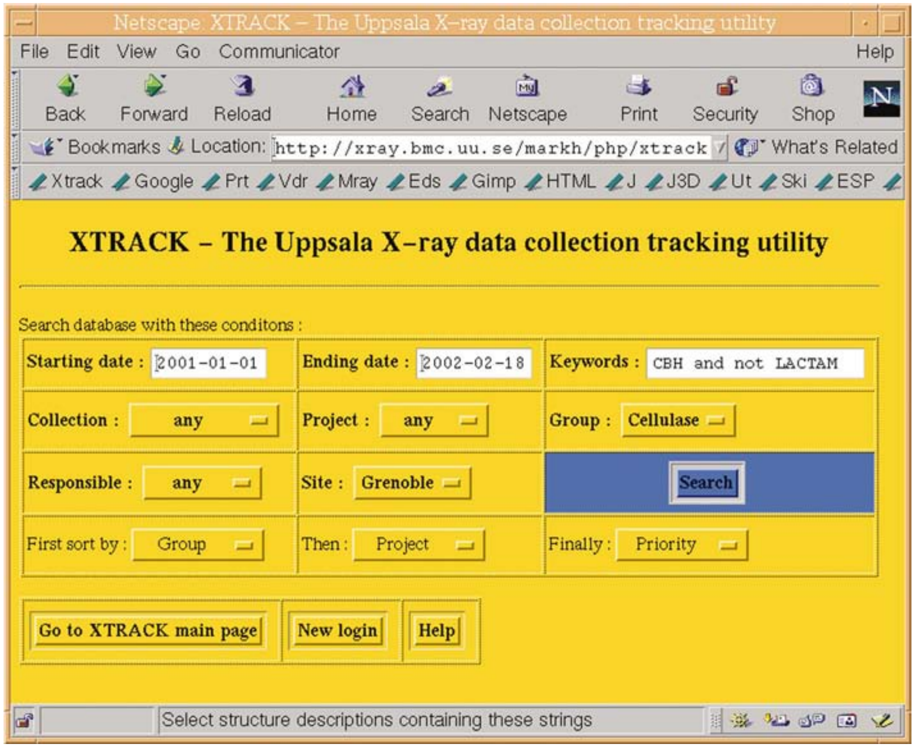
The web interface is a PHP script (The PHP Group, 2002) that communicates with a PostgreSQL database (The PostgreSQL Development Group, 2002) through the Apache webserver (The Apache Software Foundation, 2002). The database contains as much information as possible about prompting, formatting and data-typing of information elements, so that the PHP code can be generic and reusable. Data is organized around a collection, an entity that corresponds to a set of X-ray data collected for a particular structure determination. Each named collection belongs to a particular project worked on by a group in a laboratory. Data for that collection is then divided up into ten pages called Structure, Chemistry, Expression, Crystallization, Collection, Data reduction, Structure solution, Refinement, Analysis and Deposition, each of which contains between five and 20 data items. Each page corresponds to a separate table in the database, with one row of each table being assigned to a single collection, which is identified by a unique sequence ID. The first ten rows of each table are reserved for system information about each column. For example, row 0 contains a nicely formatted string to use as a column header, so that the user does not have to deal with the inevitably terse database column name. Row 1 contains the size of the prompt box that should be presented to the user or, in the case of elements with a finite number of possible entries, a pointer to a list of responses. Row 5 contains a brief help string that appears in the browser's status bar when the user clicks in the input box. Row 3 contains a number of flags, including those that indicate when a data item is related to another item. When two items are related, a JavaScript function can be tied to one of the items such that the related item is updated as the user enters data. Most of these functions can be encoded entirely in the database: for example, selecting a known synchrotron beamline will cause the default detector type to be set to the device that is known to be mounted at that beamline, using the latest information entered by an administrator. Other functions rely on routines hardcoded in the PHP interpreter; for example, the relationship between the wavelength and energy of a beam. Entering data in either of these boxes causes the other box to be updated with the corresponding value. The flag in the database simply names the hard-coded-routine that should be called when either of these boxes are activated. Code for extracting data from log files is also hard-coded in PHP. This is not quite so simple to modify as the database, but a programmer can still easily add new functions without disrupting database use.

## 4. Use

When users log in, they are presented with a choice of searching for existing entries or adding a new one (Fig. 1). Search criteria can be chosen to select a particular collection or to produce a list (for example, all collections performed at a particular synchrotron in the last year). Chosen collections are then listed and can be examined or edited by authorized users (Fig. 2). A typical page selected for editing is shown in Fig. 3. Most items are entered as free text into the input boxes, but where there are a finite number of known responses, a selection list may be offered alongside the free-text box. In free text, anything recognized as a URL (beginning with either http:// or file:/) will appear as a link on subsequent reading. This is how diverse data such as coordinates, maps, diagrams and pictures are included, as well as complete log files. On some pages there is the option to extract data from a log or coordinate file. Currently recognized log files include those from *CNS* (Brünger *et al.*, 1998), *SCALEPACK* (Otwinowski & Minor, 1997) and *REFMAC* (Murshudov *et al.*, 1997), but new modules are being developed regularly. Although normally only data associated with one particular page is displayed, all the data stored for a particular collection can be listed as one long page, or a selected summary of data commonly reported in a publication can be chosen.

## 5. Platforms

Xtrack is running on a Pentium II running Red Hat Linux 6.2, Apache 1.2, PHP4 and PostgreSQL 7.1 and should port with no


Netscape: XTRACK – The Uppsala X-ray data collection tracking utility
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Shop
Bookmarks Location: http://xray.bmc.uu.se/markh/php/xtrack
What's Related
Xtrack Google Prt Vdr Mray Eds Gimp HTML J J3D Ut Ski ESP

XTRACK – The Uppsala X-ray data collection tracking utility

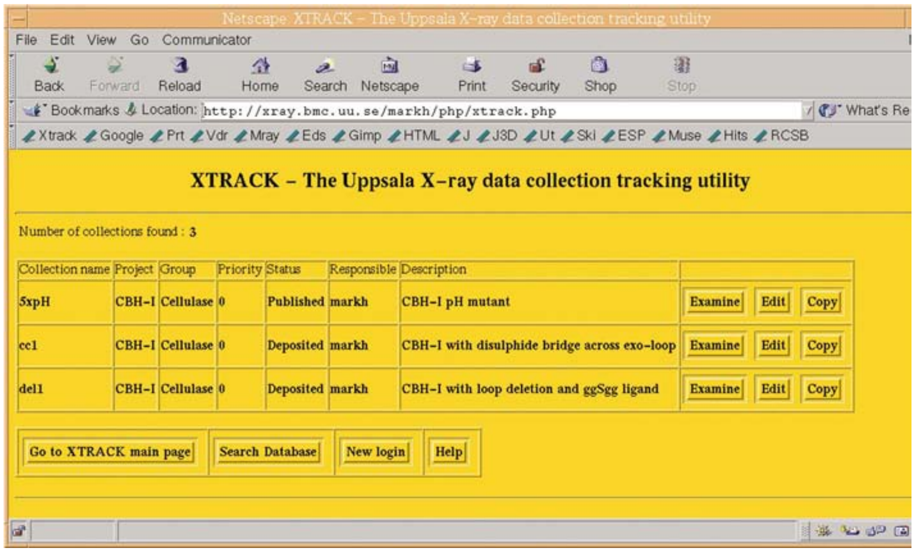
Search database with these conditons :

| | | |
|---|---|---|
| Starting date : 2001-01-01 | Ending date : 2002-02-18 | Keywords : CBH and not LACTAM |
| Collection : any | Project : any | Group : Cellulase |
| Responsible : any | Site : Grenoble | Search |
| First sort by : Group | Then : Project | Finally : Priority |

Go to XTRACK main page | New login | Help

Select structure descriptions containing these strings

**Figure 1**
The Xtrack search interface.


Netscape: XTRACK – The Uppsala X-ray data collection tracking utility
File Edit View Go Communicator
Back Forward Reload Home Search Netscape Print Security Shop Stop
Bookmarks Location: http://xray.bmc.uu.se/markh/php/xtrack.php
What's Re
Xtrack Google Prt Vdr Mray Eds Gimp HTML J J3D Ut Ski ESP Muse Hits RCSB

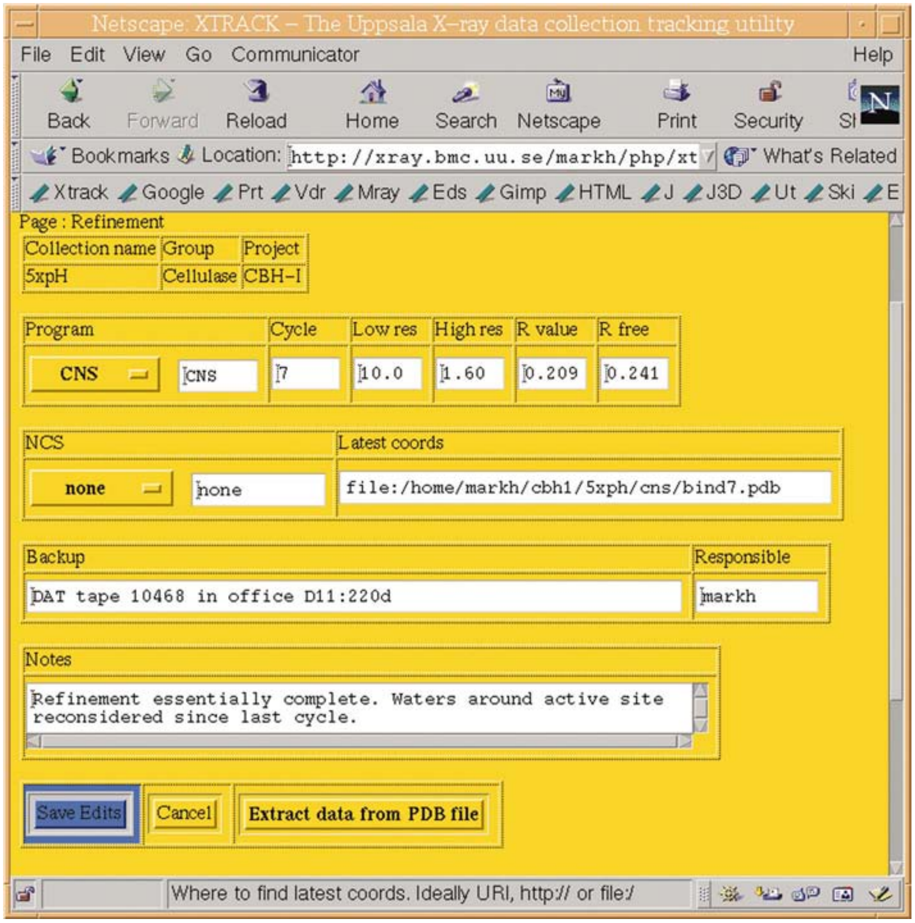
XTRACK – The Uppsala X-ray data collection tracking utility

Number of collections found : 3

| Collection name | Project | Group | Priority | Status | Responsible | Description | |
|---|---|---|---|---|---|---|---|
| 5xpH | CBH-I | Cellulase | 0 | Published | markh | CBH-I pH mutant | Examine Edit Copy |
| cc1 | CBH-I | Cellulase | 0 | Deposited | markh | CBH-I with disulphide bridge across exo-loop | Examine Edit Copy |
| del1 | CBH-I | Cellulase | 0 | Deposited | markh | CBH-I with loop deletion and ggSgg ligand | Examine Edit Copy |

Go to XTRACK main page | Search Database | New login | Help

**Figure 2**
A search result.


Netscape: XTRACK – The Uppsala X-ray data collection tracking utility
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security St
Bookmarks Location: What's Related
Xtrack Google Prt Vdr Mray Eds Gimp HTML J J3D Ut Ski E
Save Edits
Cancel
Extract data from PDB file

Location: http://xray.bmc.uu.se/markh/php/xt

Page : Refinement

| Collection name | Group | Project |
|---|---|---|
| 5xpH | Cellulase | CBH-I |

| Program | | Cycle | Low res | High res | R value | R free |
|---|---|---|---|---|---|---|
| CNS | CNS | 7 | 10.0 | 1.60 | 0.209 | 0.241 |

| NCS | | Latest coords |
|---|---|---|
| none | none | file:/home/markh/cbh1/5xph/cns/bind7.pdb |

| Backup | Responsible |
|---|---|
| DAT tape 10468 in office D11:220d | markh |

Notes

Refinement essentially complete. Waters around active site reconsidered since last cycle.

Where to find latest coords. Ideally URI, http:// or file:/

**Figure 3**
A data-input form.

changes to any system that supports PHP4 and PostgeSQL, all of which are freely available. Users can access the system from any browser with no need for plug-ins or extensions. JavaScript should be enabled in order to take advantage of adaptive default values and data-validity checking, but the system can still be used without it. Cookies are not currently used, but may be considered later for optional saving of preferences and passwords.

## References

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., 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). *Acta Cryst.* D**54**, 905–921.

Haebel, P. W., Arcus, V. L., Baker, E. N. & Metcalf, P. (2001). *Acta Cryst.* D**57**, 1341–1343.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

The Apache Software Foundation (2002). *The Apache Software Foundation*, http://www.apache.org/.

The PHP Group (2002). *PHP: Hypertext Preprocessor*, http://www.php.net.

The PostgreSQL Development Group (2002). *PostgreSQL*, http://www.postgresql.org.