

Methods and software for high-throughput data collection in macromolecular crystallography: the Brookhaven experience

R.M. Sweet, D. Schneider, H. Robinson, J. Skinner, J. Jiang,
R. Buono, M. Cowan, H. Bosshard, W. Nolan

Biology Department
Brookhaven Nat'l Lab
New York

Macromolecular Crystallography Facilities at the NSLS

The PXRR (summary at <http://www.px.nsls.bnl.gov>) is a group of about 24 workers who operate six beamlines at the NSLS. We provide a common pool of equipment and personnel, especially including overnight PX Operators and comprehensive computing support. We have a long-standing drive to provide visitors to our synchrotron PX facilities with easy-to-use tools that let them focus on the crystallographic problem at hand, and software development is a central aspect of this. We have presented numerous innovations in the field: the first electronic area-sensitive detector at a beamline in the US ('91), on-site computer resources for real-time data reduction ('91), a graphical user interface (GUI) for diffractometer and data-reduction control ('92), an integrated GUI for beamline control and data-collection ('93), automated spectrum analysis during MAD data collection ('95), "one-button" data reduction with graphical analysis tools ('96), integrated data-collection strategy prediction ('97), web-based remote experimental observation of the experiment ('98), web-based experimental control ('99), and automatic experiment logging into a web-viewable log file ('00). Our work may have provided some inspiration for software packages that appeared later.

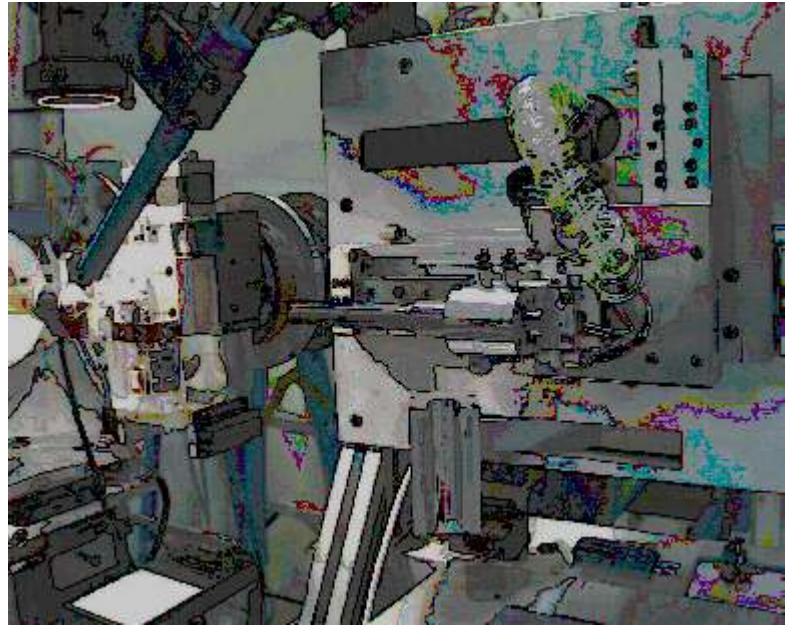
We have been defining a new paradigm for use of the synchrotron by structural biologists. Rapid Access is the norm for ordinary beam line visits. We provide fluid access to all of our work stations – the investigators see a "sea of photons," and it's not unusual either for visitors to switch to a more appropriate beam line during a visit or to operate two at once. We also have implemented a successful mail-in (so-called FedEx) operation wherein investigators send us frozen specimens, ready for diffraction studies. This requires a consistent look-and-feel among the experimental stations; this is accomplished largely through software.

Development work in progress

We have long used GUIs to control the single-crystal diffractometer in our experimental stations, and also to provide integrated control of the beamline. The latest version of this software, CBASS – Crystallography at Brookhaven: Acquisition Software System, is described in the chapter by John Skinner. To facilitate the record keeping necessary for the FedEx work, and to adapt a robotic specimen changer, we are developing a relational database system that provides web access to describe the specimens, and then automatic tracking of experimental procedures.

The robotic automounter project

We're replicating the robotic specimen "automounter" devised by Thomas Earnest's group at the Advanced Light Source of Lawrence Berkeley Laboratory. Dieter Schneider is leading the project, which is a collaboration with Earnest's group. We have tidied up mechanical drawings, added a number of improvements, and have produced the first device, which is in operation now at beam line X12-B at the NSLS. The device is being replicated for use at other beam lines.



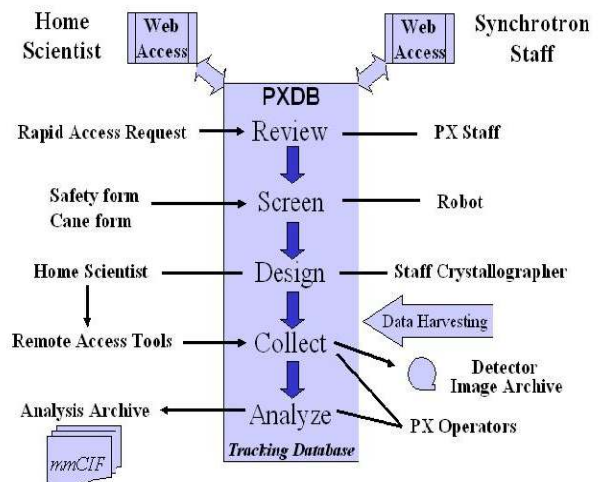
The PXDB

High-throughput macromolecular crystallography requires careful attention to each step in the process. We're developing an experiment-tracking database, which will address a range of goals.

- ! We want to harvest experimenters' personal information from the NSLS user database.
- ! Users will never have to type anything in twice!
- ! We want experimenters to be able to connect effortlessly to their specimens and their data.
- ! We'll harvest information from the experiment-control software as we go.
- ! This should produce a nearly complete summary of the experiment and result at the end.

We have addressed especially steps requiring slow, manual interaction, like experiment planning, documentation of the process, and transferring of information between stages of the process. We have created a relational database system that records almost everything in the synchrotron diffraction experiment, all the way from submitting beam-time request information, inclusion of useful information from previous stages, planning and execution of the experiment, and structure solving and deposition of the result into the Protein Data Bank. The system is accessible to users and our staff through the web where project overviews and various reports can be pulled up. The system is also integrated with the experiment setup and data collection system. An XML import/export tool allows easy transfer of information to/from other stages. The system employs the PostgreSQL database engine. It

Overall plan for use of this database



should be adaptable to ORACLE.

Many aspects of the core infrastructure are in place, and the PXDB was released to users on 1 July 2004. It connects the project to the data, and records information about every image for our reporting and eventual user applications. We also have released a "Rapid Access" application system for users, and these users will find information about their experiment waiting for them when they begin the experiment a few days or weeks later. So far there's no direct connection to the NSLS proposal system, PASS, nor can we get personnel information from the lab-wide Guest Information System. All of this is in progress and there are no serious roadblocks.