Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Mail-in crystallography program at Brookhaven National Laboratory's National Synchrotron Light Source

Howard Robinson, Alexei S. Soares, Michael Becker,‡ Robert Sweet and Annie Héroux*

Biology Department, Brookhaven National Laboratory, Upton, NY 11973, USA

‡ Current address: Argonne National Laboratory GM/CA-CAT, Building 436D, 9700 South Cass Avenue, Argonne, IL 60439, USA.

Correspondence e-mail: heroux@bnl.gov

Received 3 April 2006
Accepted 6 July 2006

For the last six years, the Macromolecular Crystallography Research Resource (PXRR) has been providing mail-in access to several beamlines at the National Synchrotron Light Source (NSLS) for X-ray diffraction. This program is available at no charge for academic users from the US and around the world. This service has fostered a new paradigm for collaborations between the local staff scientists and synchrotron users. It also accelerates the innovation of new productivity tools, which benefit users who come to NSLS beamlines as well as those who use the mail-in program.

1. Motivation

Time constraints, beam availability and travel costs are the most obvious factors motivating experimenters to use the mail-in program rather than to come to the synchrotron in person. However, another motivator has emerged: the ability to have crystals screened and data sets collected promptly has allowed users to adjust crystallization conditions and to send new crystals quickly with improved quality, thus accelerating their research projects.

From the point of view of the Macromolecular Crystallography Research Resource (PXRR), minimizing the scheduling constraints, taking advantage of the available beamtime, decreasing the time spent for training and increasing the number of users served is an efficient approach.

Conceived and operated by the PXRR, the mail-in program has been highly successful. As previously reported in a *Nature* news feature (Schmidt, 2003), the success of our program has been recognized by several other synchrotron facilities, which now have added similar modes of operation to their programs. We relate in this article some protocols, practicalities and productivity metrics.

2. Balancing the load

Experimenters interested in the mail-in program initiate communication with the mail-in manager, by 'phone or e-mail, about the suitability of their project. They provide an abstract and other details of the project *via* our database (PXDB; Skinner *et al.*, 2006), by submitting a web-based project description (http://www.px.nsls.bnl.gov/pxdb). The mail-in staff then consider the project. At this juncture, we either inform the home experimenter that a staff member is ready to proceed as a mail-in experimenter (ME) or that the project is not suitable owing to staffing limitations or for technical reasons. This provides us with the means, early in the process, to regulate our workload depending on staffing levels and the overall activities of the PXRR.

doi:10.1107/S0907444906026321

3. Collaborative framework

It might seem important to establish the exact nature of the collaboration between the home experimenter and the ME at the start of each project. However, we have found this difficult to pin down at the outset. Instead, we find it more useful to provide guidelines to help the home experimenter evaluate whether it is most reasonable to identify the contribution of the ME with an acknowledgement in the paper or whether the home experimenter should consider inviting the ME to co-author a paper. Our published guidelines for the home experimenter (http://www.px.nsls.bnl.gov/publication_policy.html) consider two extremes: at one end are projects that have involved fairly straightforward data collection and analysis by the ME and at the other end are projects where the ME has been heavily engaged, contributing significant effort and scientific judgment. Projects that lie in the grey area between these two extremes are left for the home experimenter to evaluate the acknowledgement/co-author alternatives. For consistency, the MEs pursue all projects as if they are collaborators and then they respect the home experimenter's decision about co-authorship. We have no preconceived limits to the extent of our involvement in a new project. Our contribution is as small or as large as the users demand.

4. Project sequence

After filing the forms describing their group and their projects through the PXDB, the home experimenters can submit mail-in requests for their projects. This initiates communication between an ME and the home experimenters. The home experimenter will then send the crystals in a dewar. It is typical to have more than one project per dewar, sometimes destined for multiple MEs. A mail-in 'visit' starts when the dewar arrives. With a pool of six beamlines (two insertion-device and four dipole) to choose from, an ME often initiates a project the same day it arrives; nearly all crystals are put on the beamlines within 2 or 3 d. The ME screens all the crystals received (three images are taken at 0 and 90°). The ME considers the diffraction quality (resolution limits, spot shape, mosaicity) and decides on the data-collection strategy for the experiment (choice of dipole or insertion-device beamline and collection parameters *etc.*). Further data with a different strategy can always be added at the home experimenter's request.

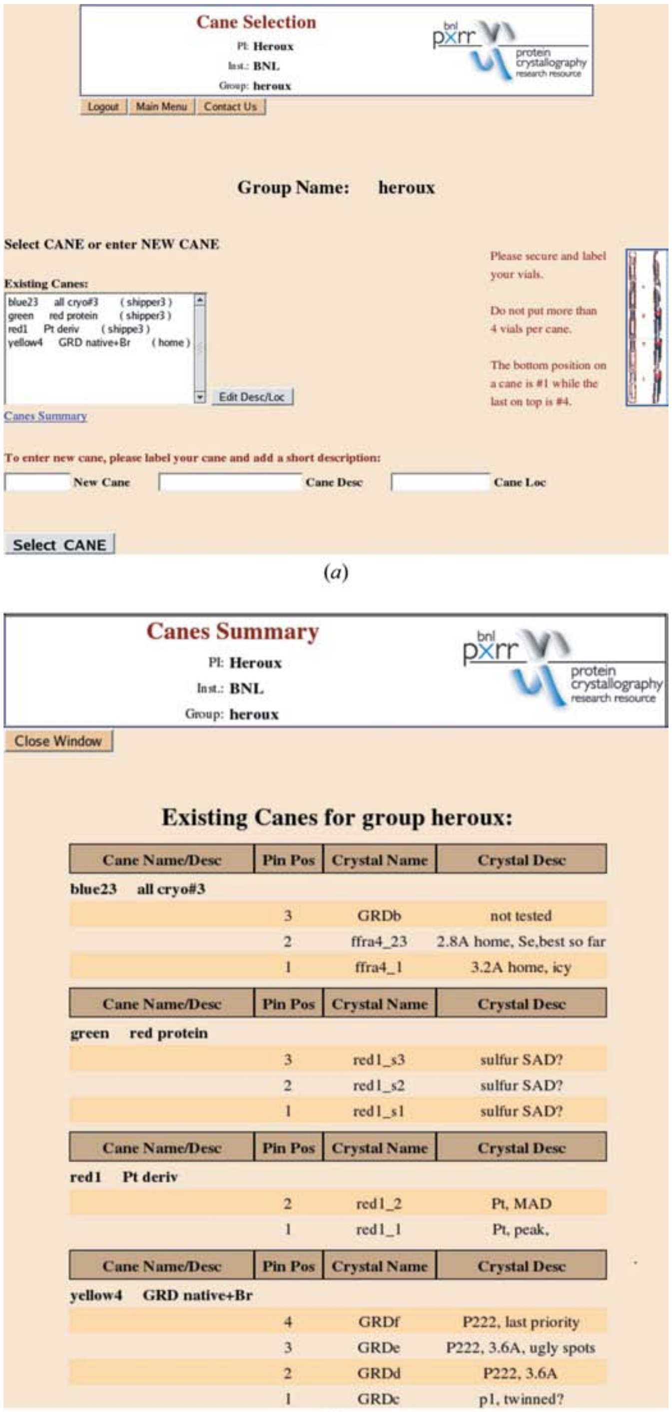
Usually, all the crystals are returned to the home laboratory within a week. This prompt turnaround is largely possible owing to the excellence of the crystallographic software available: *HKL*-2000 (Otwinowski & Minor, 1997) and *MOSFLM* (Leslie, 1992) for dealing with the raw data collected and *SOLVE* (Terwilliger & Berendzen, 1999), *HKL2MAP* (Pape & Schneider, 2004), *Phaser* (Storoni *et al.*, 2004) and *RESOLVE* (Terwilliger, 2000) for solving the structure. The visualization of the models and electron-density maps is made easy with *COOT* (Emsley & Cowtan, 2004) and *O* (Jones *et al.*, 1991). Without these programs, our ability to assess the quality of the data as we collect it would be impossible.

When sending the crystals, either in canes or in pucks, the home experimenter fills out a PXDB form that identifies labels for each sample and provides a brief description of what should be done with each crystal. Keeping track of the locations of the shippers, canes, pucks and data during the visit is the challenge that has driven the development of innovative tools now provided by the PXDB (see Fig. 1). For instance, information from the cane form can be loaded into the beamline-control software, *CBASS*, where simple button clicks can activate crystal screening and sweep collections (Skinner *et al.*, 2006). The home experimenters can monitor the progress of their crystals at the NSLS, tracking data collection from their home laboratory through the sweep query of the PXDB. The sweep query includes links to data-collection journal files in HTML format with pictures of the crystal in the loop, JPEG images of diffraction patterns, all plots from the excitation scan for MAD data collections and annotations and comments added by the ME. We have now implemented an electronic bulletin board to manage and archive project-communication information.

The data from each collection are integrated and scaled by the ME as the collection is proceeding. Quality control is a key feature that we accomplish at various points during the project. During data reduction, we check for signs of radiation damage by monitoring the scaling. For anomalous data collections, we examine the electron-density map after collecting at the peak wavelength to determine whether a SAD experiment can be performed instead of a MAD experiment. If the density cannot be interpreted as protein density, then the subsequent inflection and remote wavelength data are collected. This quality control allows us to optimize our protocols and to determine whether more data are needed. We send all the raw data, images and processed data to the users on DVDs or on firewire disks. We also maintain an FTP site outside the BNL firewall for complete network transfer of data to the experimenter at home, but this can be a very time-consuming task for both parties. We are working on tools that will enable downloads directly through the PXDB web interface. The project is finished when the dewar is sent back to the home investigator.

5. Mixing data collection at the dipole and insertion-device beamlines

Users coming to NSLS in person usually want access to the brightest beamline in order to maximize the number of measurements during their time spent at the NSLS. With mail-in, the MEs can pursue projects at both the slower dipole (DP) and at the faster insertion-device (ID) beamlines. Usually, crystal screening requires the same amount of time at either type of beamline. During the weekday daytime, the PXRR ID lines are scheduled in 2–4 h slots on short (24 h) notice. This provides the MEs with daytime access to collect data at the ID lines, while also making it possible for visiting users from the DP beamlines to obtain ID slots. This flexible approach to scheduling also makes it possible for the staff to pursue short development projects of their own.



Cane Selection

PI: **Heroux**
Inst.: **BNL**
Group: **heroux**

Group Name: heroux

Select CANE or enter NEW CANE

Existing Canes:

blue23 all cryo#3 (shipper3)
green red protein (shipper3)
red1 Pt deriv (shippe3)
yellow4 GRD native+Br (home)

Please secure and label your vials.

Do not put more than 4 vials per cane.

The bottom position on a cane is #1 while the last on top is #4.

To enter new cane, please label your cane and add a short description:

New Cane
Cane Desc
Cane Loc

(*a*)

Canes Summary

PI: **Heroux**
Inst.: **BNL**
Group: **heroux**

Existing Canes for group heroux:

Cane Name/Desc	Pin Pos	Crystal Name	Crystal Desc
blue23 all cryo#3			
	3	GRDb	not tested
	2	ffra4_23	2.8A home, Se,best so far
	1	ffra4_1	3.2A home, icy

Cane Name/Desc	Pin Pos	Crystal Name	Crystal Desc
green red protein			
	3	red1_s3	sulfur SAD?
	2	red1_s2	sulfur SAD?
	1	red1_s1	sulfur SAD?

Cane Name/Desc	Pin Pos	Crystal Name	Crystal Desc
red1 Pt deriv			
	2	red1_2	Pt, MAD
	1	red1_1	Pt, peak,

Cane Name/Desc	Pin Pos	Crystal Name	Crystal Desc
yellow4 GRD native+Br			
	4	GRDf	P222, last priority
	3	GRDe	P222, 3.6A, ugly spots
	2	GRDd	P222, 3.6A
	1	GRDc	p1, twinned?

(*b*)

Figure 1
Two of the display forms from the PXDB that provide communication between the mail-in experimenter and the investigator at home. (*a*) The root of the cane-description forms, showing the identities of existing canes and offering the possibility of editing these or describing new ones. (*b*) The cane-summary form, showing some examples of notes that can be communicated from the home investigator to the mail-in experimenter.

6. Mail-in productivity

Several criteria could be examined to evaluate the productivity of the mail-in program. An ideal comparison would be to assess the annual number of crystals, data sets, PDB deposits and number of publications produced through the mail-in program with those same figures for the PXRR at large or another facility altogether. However, this comparison is not possible because of the difficulty in harvesting these parameters even with the advent of the database. Users are in the habit of naming their crystals 'test' until they find one they will actually collect data from. We know how many sweeps correspond to crystals named 'test', but we are unable to determine how many crystals they represent.

One measure of productivity is the number of resulting publications (Fig. 2), which is monitored and posted on our website (http://www.px.nsls.bnl.gov). One factor in the notable increase of publications over the 5 y is that the rate of co-authorship has increased from 20% in the beginning to more than 92% in 2005. This fact suggests that the year 2000 paper count, based on voluntary information from researchers and our gleanings through PubMed, is probably under-represented because we were less often involved in the early publications. The 2000 number also reflects less ME availability in the early years; there are now three scientists associated with the mail-in program, whereas originally there was one. An overview of all the PDB depositions per year and beamlines around the world can be found on the BioSync website (http://biosync.rcsb.org).

Another measure of productivity are PDB depositions (also Fig. 2). Here again there has been a notable increase, from two depositions in 2000 to a high of 17 in 2004 followed by 12 in 2005. MAD data sets can be vital for the progress of an investigation, but often these data sets are not the best that

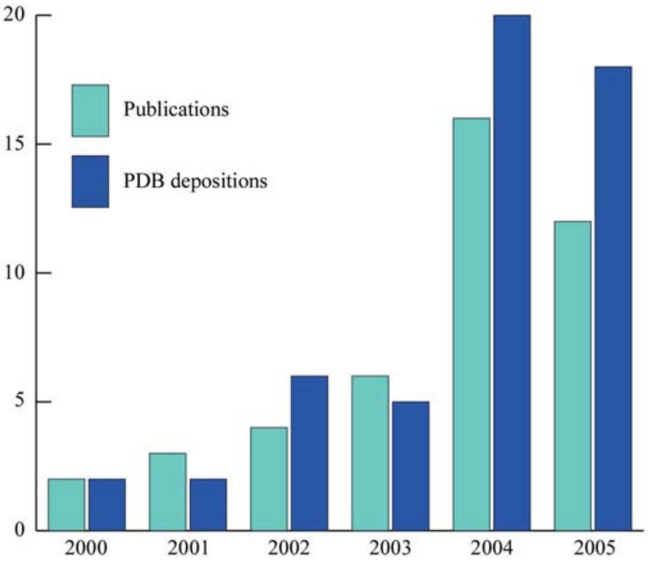


Figure 2
Histogram relating annual productivity of the PXRR mail-in program. Shown are the number of publications and PDB depositions for the years of its existence.

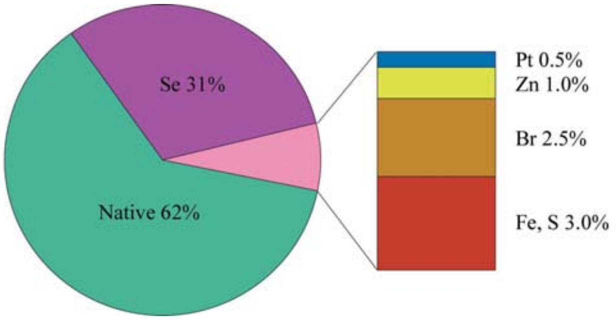


Figure 3
Distribution of the type of experiment performed at the PXRR beamlines during 2005.

might be possible from crystals of the particular molecule. Therefore, the structure that results directly from these data often is not deposited. Instead, in preparation for publication and deposition, researchers follow the MAD experiment with a higher quality data or native data set to be used for the refinement. That data set may or may not be collected at the PXRR beamlines. Thus, a structure that was initially solved here may not be reported as such and in any case may possibly occur in a later year than the original data collection.

Another metric of effort, if not productivity, is the amount of beamtime used. In 2005, 18% of the time available at the six PXRR beamlines (X12C, X12B, X25, X26C, X29 and X8C) was used for mail-in projects. 1200 mail-in crystals were mounted and tested and data were collected for 700 crystals. As seen in Fig. 3, approximately 62% of those crystals required a native data set; the rest were either SAD or MAD experiments, mostly at the selenium edge. These crystals represent projects pursued by 48 groups; most of these groups sent projects more than twice in the year, for a total of 164 visits.

7. Benefits to the PXRR

Our experience is that the MEs spend an average of about 20 h on each project. In 2005 there were about 100 active projects, so we estimate the combined efforts of our three scientists as MEs was about one full-time equivalent scientist. This includes time devoted to data collection, processing and analysis, as well as to management of images and communication with the home experimenter. The limited amount of staff time available to meet an increasing load has driven the development of technologies to improve the efficiency of our efforts. An important example is the development of better information management with the PXDB database system, which came on line at the beginning of 2005. The mail-in program has helped to stimulate the development of automounters at our beamlines for screening large-volume projects. So far, one DP beamline and one ID beamline are equipped with the automounters. Operational innovations have followed on the heels of increased demand for the mail-in projects. Most important among these innovations is the flexible way in which our ID beamtime is managed.

The mail-in program is a driving force for quality control across our beamlines. The MEs are daily users of the beamlines: broken parts are fixed, new ideas are adopted and everything works better. All our beamlines can perform anomalous diffraction experiments at X-ray energies that range from the bromine (13.5 keV) to the iron (7.1 keV) edge, so an experiment can be moved from one beamline to the other. We encourage our on-site users to follow this model and become multi-beamline users during their visits. This cross-beamline user experience has driven the unification of the data-collection environment. The consistency of operation of the beamlines and the active communication among the technical and scientific staff enable an easy transition from one beamline to the other.

Finally, a significant advantage to the PXRR is that our mail-in scientists are able by this mechanism to establish firm intellectual contact with their clients. Many projects have developed a life of their own, leading to new insights and new avenues of investigations.

Financial support comes principally from the Office of Biological and Environmental Research of the US Department of Energy and the National Center for Research Resources of the US National Institutes of Health. We would like to thank Richard Buono, Heinz Bosshard and Matthew Cowan for their strong interactions in regards with the database, John Skinner for his constant support and development of *CBASS*, Jiansheng Jiang and Martin McMillan for maintaining the crystallography software on all the computers and finally the PX-operator group for their prompt assistance to all users.

References

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
Schmidt, T. S. (2003). *Nature (London)*, **423**, 799–800.
Skinner, J. M., Cowan, M., Buono, R., Nolan, W., Bosshard, H., Robinson, H. H., Héroux, A., Soares, A. S. & Sweet, R. M. (2006). *Acta Cryst.* D**62**, 1340–1347.
Storoni, L. C., McCoy, A. J. & Read, R. J. (2004). *Acta Cryst.* D**60**, 432–438.
Terwilliger, T. C. (2000). *Acta Cryst.*, D**56**, 965–972.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.