PXGEN: a General-Purpose Graphical User Interface for Protein Crystallography Experimental Control and Data Acquisition

Stephen H. Kinder,* Sean M. McSweeney† and Elizabeth M. H. Duke

CCLRC Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, UK.
E-mail: s.h.kinder@dl.ac.uk

PXGEN is a general-purpose graphical user interface for experimental set-up and control of protein crystallography data collection. *PXGEN* is not linked intrinsically to any software package or proprietary hardware and should be transportable to other experimental facilities. The experimental techniques supported are single-wavelength data collection and multiwavelength anomalous dispersion. The graphical user interface runs on a UNIX-based workstation exploiting the host's power to manage multiple programs. *PXGEN* provides a mechanism for making data collection much easier and less error-prone. The design and implementation of *PXGEN* are described, which is now installed on protein crystallography beamlines 9.5 and 7.2 of the Synchrotron Radiation Source at Daresbury Laboratory.

Keywords: protein crystallography; data acquisition; experimental control; graphical user interface; multiwavelength anomalous dispersion.

1. Introduction

At Daresbury Laboratory a number of data-acquisition programs developed independently to control protein crystallography monochromatic data collection (multiwavelength anomalous dispersion included; hereafter MAD), Laue data collection and X-ray absorption edge scans (as a tool for both MAD and monochromator calibration). The separate development paths of these packages resulted in inconsistent user interfaces.

The image-plate detector systems now in use at many laboratories and synchrotrons come with control software available to run on UNIX workstations. Synchrotron experimental facilities generally have a number of additional pieces of equipment not controlled by standard delivered image-plate control software. New hardware and software have been developed at Daresbury Laboratory for local hardware control, *e.g.* client-server software with network communication to dedicated VME processors. A complex hardware and software set-up up has evolved.

UNIX workstations provide the possibility for new and easier-to-use user interfaces hiding the complexity and details of the hardware and software involved. These interfaces can reduce the learning time required by new and revisiting users of experimental facilities. Graphical user interfaces (GUIs) also offer the opportunity to limit the possibilities for user errors, *e.g.* by the use of option menus allowing selection from a known set of possibilities. These factors are particularly important at synchrotron sources.

† Present address: European Molecular Biology Laboratory, c/o Institut Laue Langevin, BP 156, 38042 Grenoble CEDEX 9, France.

These facilities have a rapid turnover of users, as well as constant new visitors, so the efficient use of limited beamtime is essential.

2. Data-collection hardware

The Synchrotron Radiation Source (SRS) at Daresbury Laboratory is a second-generation synchrotron source. The protein crystallography stations are presently using image-plate detector systems. The data-collection hardware available on the protein crystallography station 9.5 (Thompson *et al.*, 1992) at Daresbury will now be summarized. This is relevant to the development of *PXGEN*. The main beamline components of this station are represented schematically in Fig. 1.

Station 9.5 has a 300 mm marresearch® (MAR) image-plate detector system and Enraf–Nonius CAD-4 goniometer. The MAR image plate has a dedicated crate of control electronics and separate processor which communicates with the host system *via* a SCSI bus. The Enraf–Nonius goniometer has its own dedicated control electronics and processor. Communication with the host system is *via* a serial line. A rapidly tunable channel-cut monochromator is controlled by a dedicated DC servo feedback control system with a dedicated processor. Standard alignment of other components (*e.g.* collimator to the beam, mirror, beamline slits) is performed by in-house developed stepper motor controllers resident in a VME system. A 68030 processor provides control of the VME crate. Collection of incident-beam intensity statistics is performed by further VME modules. The complexity of hardware available and

required to be controlled, particularly on SRS station 9.5, is a recent development.

The other Daresbury protein crystallography stations also have MAR image-plate systems. The rotation axis on station 9.6 is as supplied by MAR and that on 7.2 is of in-house design, but the rotation is now controlled by the standard MAR electronics.

3. Data-collection software

The data flow diagram in Fig. 2 shows a number of inter-acting programs in the system controlling data collection on station 9.5. The circles represent programs, the squares are items external to the system, and the parallel lines are data stores within the system. The arrow-headed lines represent directional data flow. Fig. 2 shows the central role which *PXGEN* plays in the overall system. Some programs run on the VME system and communicate with *PXGEN via* UNIX sockets. Fig. 3 is a decomposition of 'Image collection' in Fig. 2.

MAR image-plate detector systems are supplied with a number of pieces of software to provide control. Those of relevance here are represented in Fig. 3. The *mardc* process checks the MAR command file for commands which it then acts upon. It is this program which actually interfaces to the MAR electronics. The *ip_xform* process transforms spiral images as output by the detector into Cartesian coordinate images. These programs have not been modified at all for our purposes here.

This brief description shows some of the software complexity we have to interact with. This complexity is likely to increase as the station is further developed. The role of a GUI in this system is vital to hide the complexity and make the process of collecting data as easy and reliable as possible.

PXGEN presently runs on UNIX systems requiring X-windows and *MOTIF*. At Daresbury Laboratory we use Silicon Graphics systems on our protein crystallography stations but the software has been ported to Hewlett-Packard systems for the ESRF. The GUI is separate from local-hardware-specific data-acquisition software.

4. *PXGEN* design

Initial development has so far centred on station 9.5 at the SRS, Daresbury Laboratory. A modified version of *MADNES* (Messerschmidt & Pflugrath, 1987) was available to control the MAR detector and CAD-4 goniometer combination we are using. However, we did not want a GUI closely linked to *MADNES* and designed in the particular way of the *MADNES* GUI (Skinner, Pflugrath & Sweet, 1992). We wanted to design a user interface which would be capable of using any package to perform the hardware control. Indeed, we also felt there was no need for the user to know which particular package was performing the control; the GUI should hide all this detail. A central theme to the overall design has been to display, in any part of the interface, only the essential information.

MOTIF has now become a *de facto* standard for most workstation manufacturers. We believe it is right to try and fit in with the 'look and feel' of the resident host machine. Usually this is a UNIX workstation running *MOTIF*. The user interface management system *UIM/X* (Sastry & Gallop, 1994) was chosen at an early stage. This allows the design of *MOTIF* GUIs in a 'what you see is what you get' (WYSIWYG) manner. Many built-in tools, editors and a C interpreter make the process very easy using *UIM/X*.

The central starting point for *PXGEN* is a single window giving easy access to the commonly used functions (Fig. 4). Brief context-sensitive help is available at the bottom of the main window. The layout of this main window does not change during use, thus making it easy for the user to remember button locations. The 'Data collection' toggle buttons allow the choice of experiment type to be performed. This selection affects the actual behaviour when the 'Perform exposures' button is pressed and the sensitivity of some items in the 'Edit exposures' interface (see Fig. 6). An insensitive GUI element is greyed out and its function cannot be used.

When a button is pressed and further interaction is required, a separate interface pops up to handle that interaction. The use of a separate window for each function allows specific design for each function. This has been another central design philosophy.
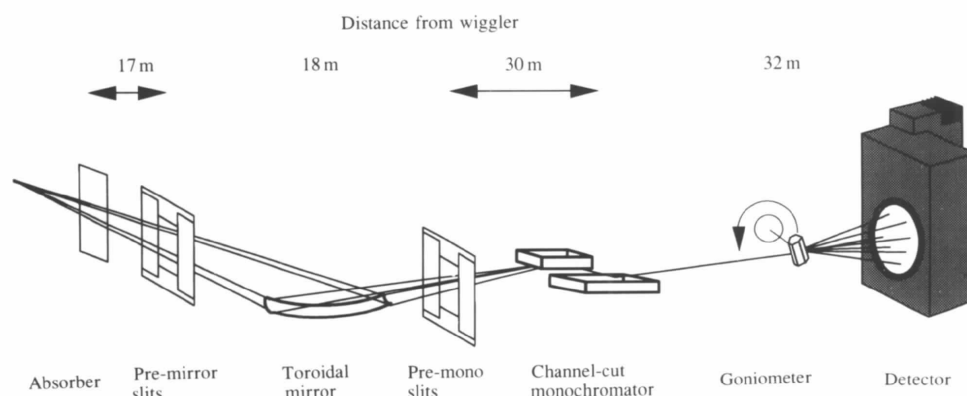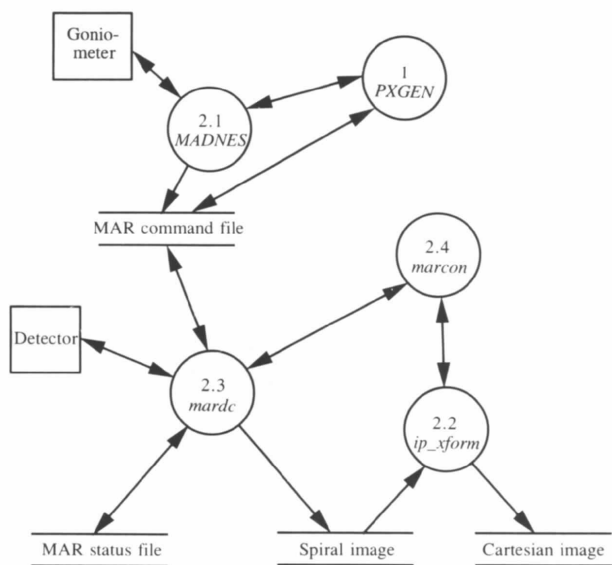

Figure 1
Schematic layout of SRS station 9.5.

User choices are *via* option menus where possible, reducing the risks of error. These menus allow a single choice from a list of known possibilities, *e.g.* monochromator position units. Number entry checking for basic floating-point or integer validity is performed at the point of entry. All experiment parameters are checked for validity, *e.g.* monochromator positions are achievable, before operations such as data collection are allowed.
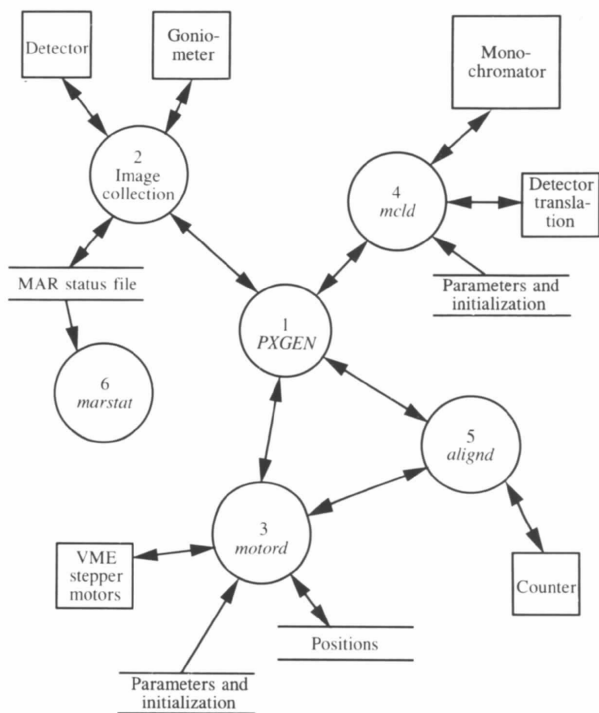
Figure 2
Main programs controlling data acquisition on SRS station 9.5.

Figure 3
Decomposition of process 2 in Fig. 2.

PXGEN and supporting software run in a client-server manner. *PXGEN* is the main client making use of programs such as *MADNES*, *motord* (stepper motor control) and *mcld* (monochromator and detector distance control) as servers. *MADNES* is run under the GUI receiving its commands from the interface rather than a user typing at a keyboard. Its output is similarly directed to *PXGEN*. *MADNES* has been modified only slightly to communicate some special commands, such as exposure completed, directly to *PXGEN* separately from normal output. Other local programs have been specifically written to communicate *via* UNIX sockets. Using this client-server model, specific functions are easily isolated. Errors are reduced, as are maintenance implications and effort. The software solution is easily scaleable and upgradable. The effort required in replacing a specific piece of hardware, and required software, is minimized.

5. Experimental set-up

An easy and intuitive to use interface has been designed to allow the set-up of various beamline components prior to data collection (Fig. 5). The central boxed area, shown containing 'Carriage Front' and 'Carriage Back', changes depending on the category of motors selected. Having selected a category, the specific movement required is selected by simply clicking on it. Motor categories are displayed at the top. For each motor (such as Carriage Front) a number of motions of that motor are possible. When a motor is selected, the sensitivity of the movement arrows change. This sensitivity effectively allows or stops the movement arrows being used. For the selected Carriage Front, this can be moved either horizontally or vertically, thus all arrows are sensitive. Once one of the arrows is pressed, that motion proceeds until the arrow is released.
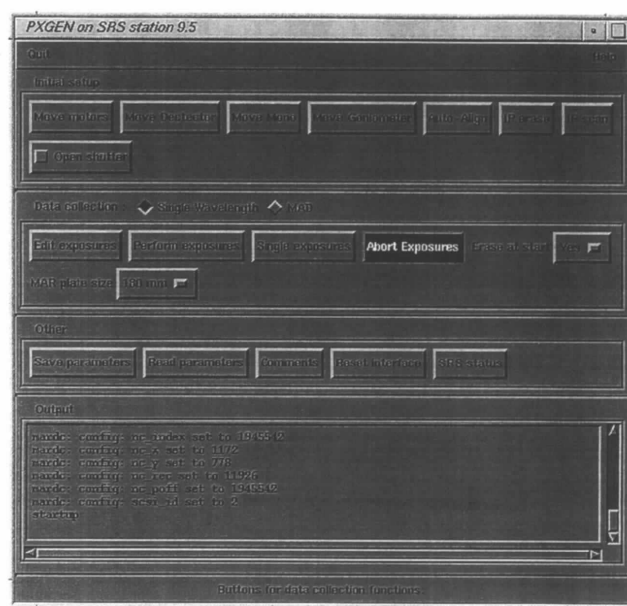
Figure 4
PXGEN main window.

The 'Move style' option menu at the top right allows selection of modes where moves are made to positions or by increments. Several moves can be set up and started together. The text window at the bottom of the interface displays messages from the network connected server.

This interface provides the means to manually align the collimator to the X-ray beam. However, an automatic alignment algorithm has been built in. After initial location of the beam, this then scans a number of axes attempting to maximize the collimator ion-chamber reading. Feedback on the process is provided by a separate interface and plots of ion-chamber readings against position.

Positioning of the 2θ arm, goniometer, monochromator and detector distance prior to data collection are allowed using separate but very similar interfaces. Display and input of positions are made in units selected by the user from option menus. Single oscillations and stills are provided for.

The ability to align a crystal such that the Bijvoet pairs are collected on the same images is standard practice in both MAD and optimized anomalous data collection. We have implemented a procedure to allow this using a single image which makes full use of the CAD-4 diffractometer. This is based upon the use of *Imstills* (Leslie, 1995) for peak searching and *DENZO* (Otwinowski, 1993) for indexing and alignment determination. The procedure is operated from a single interface. A selection of possible crystal orientations for the requested axes orientations is presented.
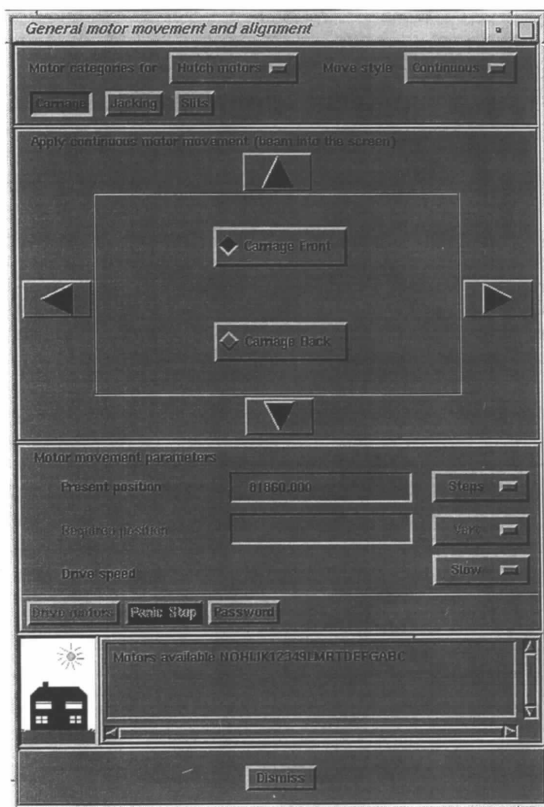
Table 1
Sample MAD data collection.

Wavelength (Å)	Rotation range (°)
0.9323	0–5 step 1 followed by 180–185 step 1
0.9192	0–5 step 1 followed by 180–185 step 1
0.9185	0–5 step 1 followed by 180–185 step 1
0.9323	5–10 step 1 followed by 185–190 step 1
0.9192	5–10 step 1 followed by 185–190 step 1
0.9185	5–10 step 1 followed by 185–190 step 1
0.9323	85–90 step 1 followed by 265–270 step 1
0.9192	85–90 step 1 followed by 265–270 step 1
0.9185	85–90 step 1 followed by 265–270 step 1

6. Data collection

Data collection is set up using the interface shown in Fig. 6. Most fields default to those last entered or sensible defaults. Detector distance and monochromator position 1 default to the current positions. Greyed-out fields are insensitive, for example, if they relate to a change in experiment mode, *e.g.* Mono positions 2–4 are only for use with MAD experiments. When the 'Apply' or 'OK' buttons are pressed, a number of checks are performed to confirm validity of the entered data. An error dialogue appears with explanation if errors are detected. Examples of these checks are for monochromator and detector positions within limits. Once validity has been confirmed the 'Perform exposures' button within the main *PXGEN* window is made sensitive.

The data-collection techniques supported are single-wavelength oscillations and multiwavelength anomalous dispersion (MAD). These techniques are well documented elsewhere (Arndt & Wonacott, 1977; Hendrickson, 1991). The possibilities for MAD data collection are quite complex. Images are collected for the same rotation range at between 1 and 4 wavelengths. This rotation range can be divided into a number of blocks. Each block is collected at each wavelength before moving on to the next block. As well as collecting each block of data, identical blocks can be collected 180° away. These shifted blocks are collected immediately after the corresponding unshifted block. The
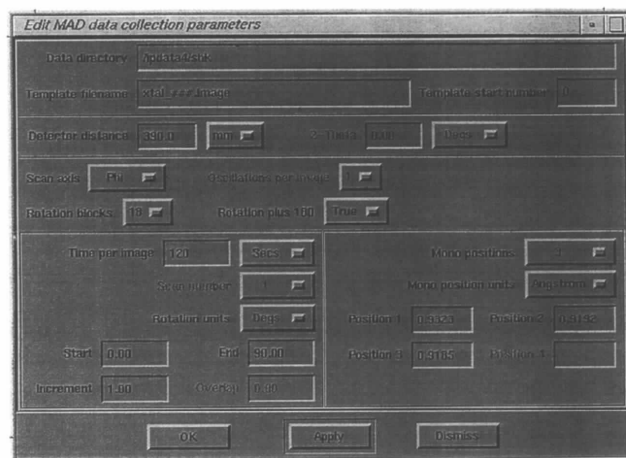
Figure 5
General move motors interface.

Figure 6
Edit exposures interface.

sequence of data collection set up in Fig. 6 is shown in detail in Table 1. *PXGEN*'s role in MAD data collection is to control the above sequence. Images are collected by the same software used for conventional single-wavelength data collection. *PXGEN* ensures the correct wavelength and starting rotation position for each rotation range. The underlying data-collection software is not modified for this purpose.

7. Future plans

The easy selection of an efficient strategy for data collection is an area we intend to address. This area of data collection has already been dealt with by others. *MOSFLM* V5.30 (Leslie, 1992) provides a data-collection strategy for a single rotation axis and programs such as *DS* (Tanaka, 1995) provide simulation for more complex geometry systems. We aim to provide a facility which will suggest to the user various combinations of oscillations which will maximize completeness given an initial crystal orientation.

Once aligned and data collection is underway there needs to be a means of keeping track of data quality. This is particularly important at synchrotron sources where beamtime is at a premium. This may be as simple as watching an image display but we need to think about some quantitative analysis of images or regions of them during data collection. With the availability of high-speed dedicated computers this may well be full-blown continuous data processing. Certainly we plan to implement automatic generation of scripts to run data-processing packages such as *DENZO*, *XDS* (Kabsch, 1988) or *MOSFLM*.

The general nature and transportability of *PXGEN* has been mentioned earlier. To help and make the use of other underlying control software easier, a lower layer of software, between the GUI and the underlying control software, has been implemented. *PXGEN* sends generic commands to this layer which parses them into commands for the specific control software in use. For use elsewhere it is this parser layer, written in C, which would need modifying if, for example, there was a local or other third-party program for the control of data acquisition. The parser would need modifying to translate generic *PXGEN* commands into commands required by the other program. The GUI itself could remain as it is with a single environment variable controlling site-specific configuration at run time. A definite requirement in this area at Daresbury is to use *PXGEN* as a front end to future image-plate and CCD detectors. Two such detectors to be used at Daresbury are the new faster scanning MAR image-plate detector and a three-plate carousel image-plate system under development at the Laboratory of Molecular Biology, Cambridge. Different control hardware for equipment such as monochromators will also need to be taken into account.

The general area of status display and feedback during data collection has been recognized as an area which needs some attention. The various areas now present require rationalization and improvement. The ability to automatically display images during data collection is planned. Although little image interaction will be built in to begin with, this area may well grow with time and user requests.

8. Conclusions

The development of a general-purpose experimental control and data-acquisition GUI for protein crystallography has been reported. Its design and use has been discussed. Plans for its future development are in hand. We feel it provides a general base for supporting data collection at many facilities. It is not specific to synchrotron X-ray sources and could be used at neutron sources or on laboratory-based sealed X-ray tubes. *PXGEN* is in routine use at Daresbury for standard single-wavelength and MAD data collection and has received excellent user feedback. A World Wide Web server is available, http://www.dl.ac.uk/SRS/COMPUTING/PXGEN, providing further information and up-to-date developments. Availability of *PXGEN* can be given upon request: we would be keen to collaborate on further development of the package and porting it to other sites.

References

Arndt, U. W. & Wonacott, A. J. (1977). *The Rotation Method in Crystallography*. Amsterdam: North Holland.

Hendrickson, W. A. (1991). *Science*, **254**, 51–58.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Leslie, A. G. W. (1992). *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, No. 26. CCLRC Daresbury Laboratory, Warrington, UK.

Leslie, A. G. W. (1995). *Imstills*. Unpublished work.

Messerschmidt, A. & Pflugrath, J. W. (1987). *J. Appl. Cryst.* **26**, 457–462.

Otwinowski, Z. (1993). *Oscillation Data Reduction Program.* Proceedings of the CCP4 Study Weekend: Data Collection and Processing, 29–30 January 1993, SERC Daresbury Laboratory, pp. 55–62. CCLRC Daresbury Laboratory, Warrington, UK.

Sastry, L. & Gallop, J. R. (1994). *Graphical User Interface Development Tools – Report on the Review and Workshops on the Current State of the Art*, July 1994. CCLRC Rutherford Appleton Laboratory, Chilton, Didcot, Oxfordshire, UK.

Skinner, J. M., Pflugrath, J. W. & Sweet, R. M. (1992). *Development of a GUI for an Existing Command-Line Driven Program.* Proceedings of Summer 1992 Meeting/Share 79, 16–21 August 1992, Atlanta, Georgia. Share Inc., 401 North Michigan Ave, Chicago, Illinois 60611-4267, USA.

Tanaka, I. (1995). *J. Appl. Cryst.* **28**, 225–227.

Thompson, A. W., Habash, J., Harrop, S., Helliwell, J. R., Nave, C., Atkinson, P., Hasnain, S. S., Glover, I. D., Moore, P. R., Harris, N., Kinder, S. & Buffey, S. (1992). *Rev. Sci. Instrum.* **63**, 1062–1064.