Automated mounting, centering and screening of crystals for high-throughput protein crystallography

W. I. Karain,[ab] G. P. Bourenkov,[a] H. Blume[a] and H. D. Bartunik[a*]

[a]Max Planck Research Unit for Structural Molecular Biology, MPG-ASMB c/o DESY, Notkestrasse 85, Hamburg, Germany, and [b]Dept. of Physics, Birzeit University, Birzeit, Palestinian Authority. E-mail: Bartunik@mpghdb.desy.de

A fully automated system for screening protein crystals for X-ray diffraction analysis has been designed and is being installed on the beamline BW6 at DORIS in Hamburg, Germany. The system includes robotic mounting of flash-frozen crystals from a storage dewar, centering and alignment of the sample both by optical and X-ray (scattering and fluorescence) techniques, assessment of the diffraction quality of the sample, and SAD/MAD or non-conventional diffraction data acquisition with high-throughput data rates. The system covers all experimental steps required for protein x-ray structure analysis and provides a powerful means for structural genomics projects.

Keywords: structural genomics; crystal mounting robot; synchrotron radiation.

1. Introduction

The average time needed for the *de novo* solution of protein structures has been drastically reduced by a number of recent experimental and theoretical developments. These include in particular fast diffraction data acquisition using synchrotron beamlines with high incident flux density and area detectors with fast read-out, generally applicable techniques of derivatization with Se, Br or Xe, and experimental phasing by SAD or MAD methods. In many cases, where well-diffracting crystals were available, it has been possible to solve new protein structures within less than one day. Specific steps of the experimental procedures, e.g., beamline alignment and diffraction data collection at a number of preselected X-ray wavelengths (Blume *et al*., 2001), have already been automated. A number of groups started to develop crystal mounting robots and optical centering facilities (Muchmore *et al*. 2000; Abola *et al*. 2000; Roth *et al.* 2002).

High-throughput protein crystallography for structural genomics requires full automation, in particular of all steps that are carried out on a synchrotron beamline, *i.e.*, from mounting crystals on a diffractometer to verifying the interpretability of initial electron density maps. The present paper describes a system that automates all experimental steps. Furthermore, the concept includes a solution to the problem of assessing the diffraction quality of the sample, which until now required experienced crystallographers.

2. Experimental set-up

The beamline BW6 at DORIS in Hamburg, Germany, is dedicated to macromolecular crystallography. It is tunable through the X-ray wavelength range from 0.6 Å - 3.1 Å. The diffraction setup, for high-throughput applications in particular, includes a collimator, a two-circle goniostat, a CCD detector (MAR), a PIN diode and a multi-channel analyzer (Amptek) for X-ray fluorescence measurements, and a crystal cooling device (MSC) without the need for liquid nitrogen supply. The PIN diode is mounted near the horizontal plane and measures the backward scattering. The collimator includes motorized x-y slits, which define the cross-section and the divergence of the X-ray beam, and equipment for monitoring the incident intensity. The sample can be translated along the crystal rotation axis (z) and in the plane (x, y) perpendicular to this axis. All components of the beamline and the diffraction setup are controlled by software based on LabView (National Instruments). Two long-distance microscopes with S-VHS color video cameras with different fields-of-view (1.65 mm x 2.2 mm; 9 mm x 12 mm) are used for optical alignment of the sample. The pictures are digitized by a frame grabber.

3. Crystal mounting robot

The robotic system transfers flash-frozen samples to the crystal rotation axis of the goniostat. The samples are mounted on CrystalCap Copper magnetic pins (Hampton Research) with 10 mm long microtubes and Cryoloops. The robot consists of a storage dewar, a plucker that lifts a selected sample out of the dewar, and a gripper with transfer tongs that brings the sample to the rotation axis (Fig. 1). As compared to other designs (Deacon, 2001), the system is very compact and can easily be adapted to other diffractometers and environments. Presently, the crystal transfer time from the dewar to the goniostat is about 5 sec.

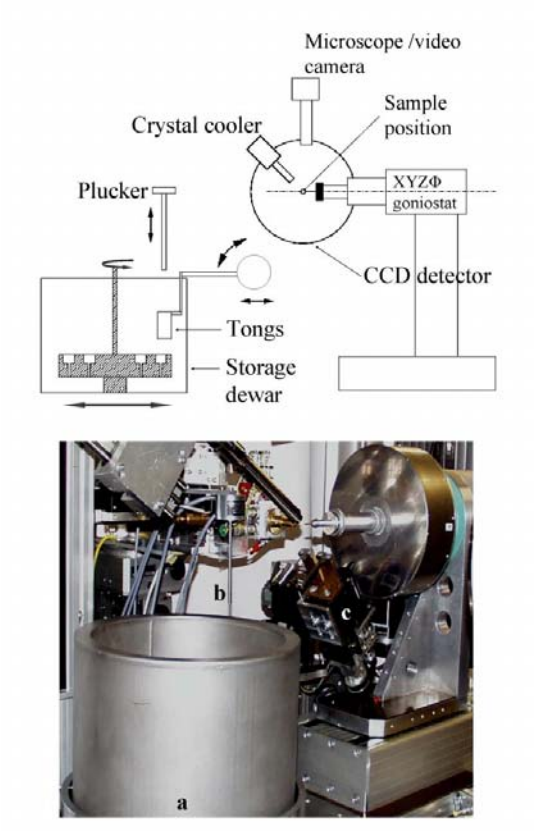

Figure 1 Diffraction set-up and crystal mounting robot with the dewar (a), plucker (b) and gripper (c).

3.1. Storage dewar

A custom made stainless steel dewar and a nitrogen filling system are used for storing the pre-frozen samples at liquid nitrogen temperature. The dewar contains a disk with 48 storage cells. The storage cells are arranged along two concentric circles. Each cell may be loaded with one sample on a magnetic pin. The disk is coupled *via* a pulley system to a stepper motor that is attached to the lid of the dewar. A combined rotation of the disc and a linear horizontal motion of the dewar bring the selected storage cell into position for removal of the sample by the plucker. For replacing samples, the entire disc is transferred to a transport dewar; during this procedure, the samples remain at sufficiently low temperature. If needed, the design of the disk could easily be modified to include up to about 100 storage cells.

3.2. Sample plucker

An electromagnet, which is mounted on a motorized translation table, is used to lift the sample at the angular mounting position out of its cell, and to position it in place for the automated transfer tongs (Fig. 2a). The electromagnet has a long soft iron pin connected to its center. This pin can be immersed in the liquid nitrogen without damaging the electromagnet itself. It is lowered into the dewar until it touches the pin steel base. It is then switched on, thus grabbing the pin, lifting it up, and positioning it in place for the tongs to enclose and hold the pin. The electromagnet is then turned off and moved away.

3.3. Gripper with sample transfer tongs

The tongs that grip the sample and transfer it to the crystal rotation axis consist of conical stainless steel semi-cylinders. They are connected to the jaws of a gripper. The gripper is attached to a rotational axis with a 90° angular motion range, which is mounted on a linear table. To remove the sample from the dewar, the tongs are positioned with the jaws open, inside the storage dewar (Fig. 2a). The sample, positioned in place by the electromagnetic pin, is then gripped by closing the jaws. A 90° rotation combined with a translation places the steel base of the pin in contact with the magnetic base of the goniometer head (Figs. 2c,d). The tongs are dipped in liquid nitrogen before grabbing the samples. This provides the cold temperature environment needed to prevent crystal damage during transfer. All motions are powered by stepper motors. The sample is returned to the dewar after testing for future use by performing the steps in reverse.

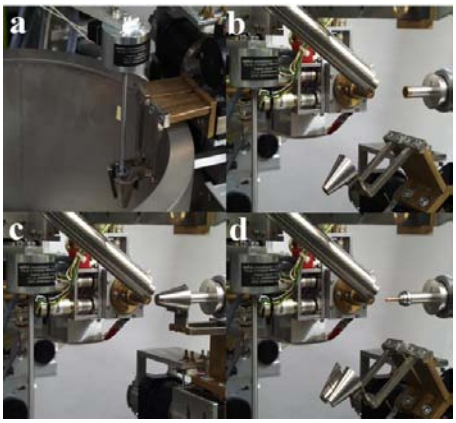

Figure 2 Subsequent steps (a), (b), (c), (d) of the sample transfer from the dewar to the crystal rotation axis.

4. Alignment and centering of the sample

The fully automatic alignment and centering of the sample after its transfer to the crystal rotation axis of the goniostat is achieved in two steps. The first step involves the use of optical methods for identifying the loop containing the sample and moving it into the X-ray beam. The second step aligns and centers the sample itself on the basis of measurements of X-ray fluorescence, if anomalous reference scatterers are present as in the case of SAD/MAD phasing, or X-ray scattering. In this way, one avoids the problems and uncertainties that are associated with attempts to identify and center small objects by optical techniques, only. Optionally, optical filtering routines may be employed after the first step mentioned earlier to identify crystalline samples with well-recognizable edges. In this case, the sample is centered optically, and the alignment is checked and refined with X-rays as described in the second step above. As a further option, semi-automatic centering is possible with a graphical user interface and point-and-click identification of the sample at two different PHI rotation angles. The user moves a rectangle that corresponds to the collimator aperture over the sample position on a monitor screen. This input is used for centering the sample. In the following, the essential aspects of fully automatic sample alignment and centering are described.

4.1. Optical alignment of the loop

The automatic identification and centering of the loop containing the sample are based on a sequence of optical filtering routines. For a test of the algorithm, 20 Cryoloops were mounted randomly on the goniometer magnetic head. A picture was taken for each one, and the pixel value at the tip of the Cryoloop was measured manually. This value was compared with that returned by the algorithm. The results were identical for all the loops with an average centering time of about 1 min. Illumination was kept constant during measurement.

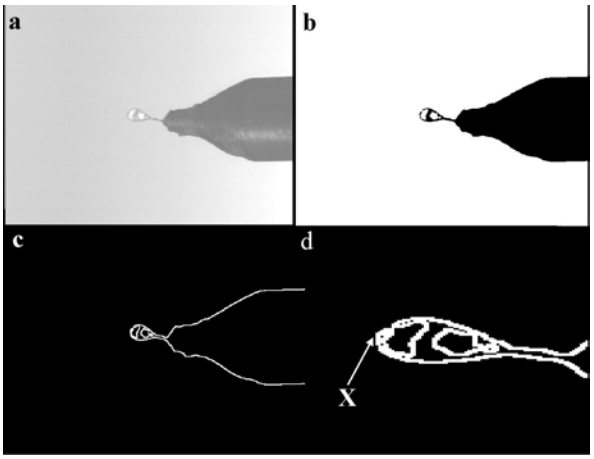

Figure 3 Optical alignment of the Cryoloop. a) Raw image. b) Thresholded image. c) Edge-detection filtered image. d) Detected loop and its tip (X).

4.1.1. Image pre-processing. The image captured by the camera consists of pixels that can store intensity values ranging from 0 to 255 (Fig. 3a). A smoothing of this raw image using a 5x5 Gaussian filter is performed (Gonzalez & Woods, 1992). The Gaussian filtering returns a weighted average of each pixel's neighborhood. This removes random fluctuations in intensity, and has the general effect of blurring the image. The image is segmented between the object of interest and the background by creating a histogram of the

intensity values for the image pixels followed by a thresholding of these intensity values. A intensity level is chosen to represent the boundary between the foreground and the background. All pixels with an intensity equal to or larger than this threshold are set to 255. The rest are set to zero (Fig. 3b). The threshold value is chosen automatically using the triangle method (Zack *et al*., 1977). A line is drawn between the maximum and minimum value of the histogram. The distance L between this line and the histogram frequency is calculated for each bin. The bin with the largest vertical distance L_{max} is chosen as the threshold value (Fig. 4).

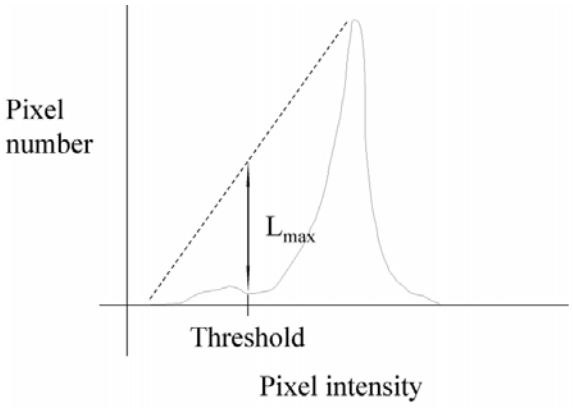

Figure 4 Triangle automatic thresholding technique used for image pre-processing.

4.1.2. Edge detection. Once the object and background are defined, the object edges are detected. This has the advantage of reducing the amount of data needed to represent an object. Edges are regions of high spatial gradients in intensity and can be detected by calculating the gradient at each pixel. This is done using a horizontal and a vertical Sobel filter (Gonzalez & Woods, 1992) . Each filter consists of a 3x3 matrix, or kernel, of integers . The kernels are convoluted separately with each pixel in the image. The new intensity value placed in the pixel, I_{new} , is given by $| I_{new} | =(S_x^2 + S_y^2)^{1/2}$ where S_x and S_y are the values resulting from centering each matrix on every pixel, multiplying the corresponding pixels of the kernel and the image, and summing the results. A thresholding is performed again to enhance the edges (Fig. 3c).

4.1.3. Loop recognition. The loop is detected using a pattern recognition cross-correlation method. A mask approximating a feature common to all loop sizes and orientations is convoluted with the edge-detection filtered image at each pixel. Pixels in the mask have a value of 255 inside, and zero outside. The origin of the mask, O, is shifted to every pixel in the image. The corresponding pixels in the mask and the image are multiplied. The largest result will occur when the mask and the image contain the same, or close to the same, pattern. After examining many pictures, we found that the pin 'shoulders' and Cryoloop 'neck' satisfy this condition (Fig. 5). This will identify the bottom of the Cryoloop neck, thus isolating the Cryoloop and sample region (Fig. 3d). The tip of the Cryoloop is then moved to a predetermined point inside the X-ray beam.

Figure 5 Loop recognition mask.

4.2. Sample alignment based on X-ray fluorescence

Experimental phasing methods using anomalous diffraction at one (SAD) or multiple X-ray wavelengths (MAD) have become so powerful that they will contribute the standard procedure for *de novo* solution of protein structures, in particular in high-throughput applications. Rather general methods of preparing suitable derivatives containing Se or other anomalous reference scatterers are available. The speed with which SAD or MAD phases may be determined on optimized beamlines like BW6 may suggest to apply anomalous phasing even in such cases where traditionally molecular replacement would be attempted. Thus, the problems associated with poor search models may be avoided.

When anomalous scatterers are present in the sample and when their concentration in the cryo-solution is sufficiently low, e.g., after back-soaking, the sample may be reliably and rapidly aligned and centered on the basis of X-ray fluorescence measurements. Assuming that the Cryoloop has been positioned as described under 4.1., the X-ray fluorescence centering is achieved as follows. The incident X-ray beam energy is set to $\approx$ 1 keV above the target absorption edge. This value is chosen instead of the anomalous scattering peak wavelength in order to completely separate the Kα(Lα) emission line from the Compton scattering. The beam aperture is set by the slit system to a vertical lineshape 1mm (vertical) x 0.02 mm (horizontal). The emission line intensity is measured as a function of the sample position and normalized by the ionization current readings. The exposure is timed by the shutter; thus unnecessary irradiation (and damage) of the sample is avoided during the movements and data transfer. The centroid position and size (FWHM) of the horizontal sample projection are calculated from the scan data. After that, the beam shape is changed to a horizontal line, and the procedure is repeated for vertical positioning of the sample in two orthogonal directions.

Figure 6 shows as an example the results obtained for a crystal of poplar plastocyainin A (PCa) with the dimensions 0.2 x 0.13 x 0.09 mm. The PCa structure contains one copper atom *per* 100 amino acids and 30% solvent. The incident beam wavelength was set to 1.05 Å (f_{Cu}" = 2.5 e^-). With an exposure time of 1 sec, up to 3000 counts in the emission line were accumulated providing extremely high contrast. The same exposure time would be sufficient even if the count rates are smaller by a factor of 30 (due to smaller sample size and/or a lower concentration of anomalous scatterers), *i.e.* practically for any sample that is suitable for an MAD experiment. The minimum linear crystal dimension on this beamline is about 20µm. There is practically no upper limit. The total radiation dose absorbed by the sample during the centering (~10^5 Gy) is too small to cause noticeable radiation damage. Further speeding up of this procedure is possible by adopting peak-finding techniques from step-scan diffractometry.

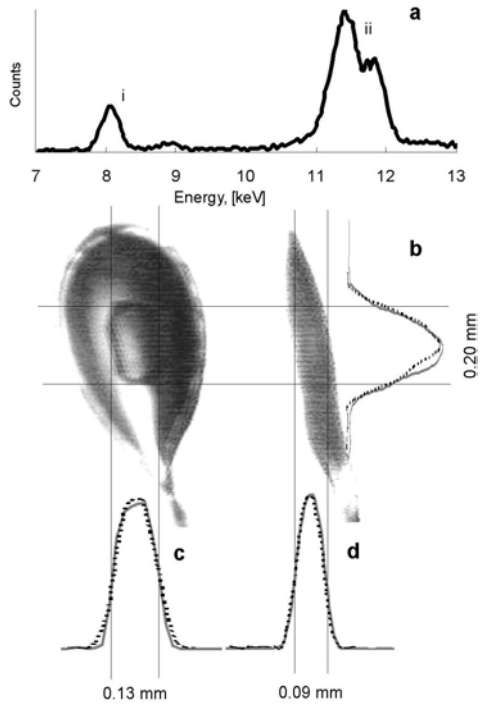

Figure 6 Sample location and centering based on X-ray fluorescence or X-ray scattering. (a) The Cu-Kα emission line (i) is separated from the Compton and elastic scattering peaks (ii). (b) Horizontal scan using an incident beam cross-section (H x V=) 0.02 mm x 0.9 mm. (c), (d) Vertical scans at PHI and PHI+90° with (H x V =) 0.8 mm x 0.02 mm. Solid lines refer to fluorescence, dashed lines to scattering.

4.3. Sample alignment on the basis of X-ray scattering measurements

When X-ray fluorescence cannot be used, the location and at the same time the crystallinity of the sample are identified by two complementary methods based on X-ray scattering. The one method aims to maximize reflection intensities, as measured by the CCD system, by step-wise translating the sample across the incident beam. Due to possible variations in the crystalline order within the sample volume, a large number of reflection spots distributed over the detector surface and knowledge of the orientation matrix are required. Figure 6 shows results for the example of a PCa crystal. The requirement for the crystal size and the achievable accuracy in the sample alignment is practically the same as for the (much faster) X-ray fluorescence method. The other X-ray scattering method is described below.

5. Assessment of the diffraction quality

Even samples that optically look like perfectly shaped crystals may turn out to have poor crystalline order. On the other hand, samples of rather crappy optical appearance in fact may be well-diffracting single crystals. Whether samples are suitable for protein crystallography can only be determined on the basis of X-ray scattering exposures. Furthermore, inspection of such images by expert crystallographers is usually required for assessing the diffraction quality.

We developed a new method that may be used to permits to automatically assess the diffraction quality of the sample in about 40 sec, which is the typical amount of time needed to collect two X-ray images for a sample. The approach is based on detecting periodic features in real space by analysing spatial frequency peaks and correlations in Fast-Fourier-transformed X-ray scattering images. The method, which may also be used to automatically locate and center the sample, has been programmed and successfully tested in applications to protein crystals of strongly varying diffraction quality. The method, which will be described in detail elsewhere (Karain *et al.*, 2002), forms part of the suite of routines used for automatic screening of samples.

6. Diffraction data acquisition

As soon as the sample has been identified as a single crystal, and auto-indexing has been performed, full data acquisition may be carried out automatically, at least in space group P1. For optimising the strategy, a determination of the space group is required, which is not yet feasible by an automatic procedure. Using software for communicating between the MAR-CCD data acquisition routines and the program controlling the beamline, a series of complete diffraction data sets at defined X-ray wavelengths and other parameters (like the crystal-to-detector distance, collimator apertures, etc.) may be recorded on BW6 in unattended mode. Thanks to a development by DESY, the beam-shutter reopens automatically after re-injection. During data collection, the X-ray optics of the beamline and the diffraction set-up are continuously and automatically realigned in order to maintain optimum flux density at the sample position. Due to a fast (Myrinet) connection between the CCD data acquisition processor and the file server (both running under LINUX) of the computer network that is used for storage, processing and backup of data, high total data rates up to 10 MB/s may be achieved (Blume *et al*., 2001). The read-out speed of the MAR-CCD currently limits the data rate to 1 MB/s. The backup system of the beamline has a capacity of more than 2 TB.

7. Conclusions

The facilities, that have been developed and whose installation on the beamline BW6 is being completed, make it possible to run all steps required for crystal screening and subsequent diffraction data acquisition in fully automatic mode. Starting from pre-frozen samples, the transfer to the crystal rotation axis, the centering in the X-ray beam, and data collection may be carried out in an unattended mode. We developed a solution also to the problem of assessing the diffraction quality of the sample. The need for intervention by expert crystallographers thus should be limited to particularly complicated cases. The present achievements already are of considerable importance to structural genomics projects.

References

Abola, E., Kuhn, P., Earnest, T. & Stevens, R. C. (2000*). Nature Struct. Biol., Structural Genomics Supplement*, **7**, 973-977.

Blume, H., Bösecke, P., Bourenkov, G. P., Kosciesza, D. & Bartunik, H. D. (2001). *Nucl. Instrum. Methods A*, **467-468**, 1358-1362.

Deacon, A. (2001). *Installation and testing of an automated crystal mounting system.* http://smb.slac.stanford.edu/research/robot.

Gonzalez, R. & Woods, R. (1992). *Digital Image Processing*. Addison-Wesley Publishing Company.

Karain, W. I., Bourenkov, G. P., Blume, H. & Bartunik, H. D. (2002). Manuscript in preparation.

Muchmore, S. W., Olson, J., Jones, R., Pan, J., Blum, M., Greer, J., Merrick, S. M., Magdalinos, P. & Nienbar, V. L. (2000). *Structure*, **8**, R243-246.

Roth, M., Carpentier, P., Kaikati, O., Joly, J., Charrault, P., Pirocchi, M., Kahn, R., Fanchon, E., Jacquamet, L., Borel, F., Bertoni, A., Israel-Gouy, P., & Ferrer, J. L. (2002). *Acta Cryst*. D**58**, 805-814.

Zack, G. W., Rogers, W. E. & Latt, S. A. (1977). *J. Histochem. Cytochem.* **25(7)**, 741-753.

 Received 26 March 2002 • Accepted 16 July 2002