Efficient collection of oscillation data - planning, pitfalls, and prospects

Marian Szebenyi

Most macromolecular crystallographers coming to CHESS collect diffraction data using the oscillation method. This is a time-tested and well understood technique, but several factors must be considered to collect good oscillation data in the most efficient way possible, so as to make best use of one's limited synchrotron time. This article considers some of these factors, and mentions how new software at CHESS will help users plan more efficient data collection.

Evaluating the Initial Image.

The first step in data collection is always to mount a crystal and take a diffraction pattern. Often a still exposure is taken first, followed by an oscillation if the still looks promising. What can we conclude from the images shown in Figure 1a and Figure 1b? From the image alone several things can be checked:

Singleness of crystal - anything other than a single pattern of well-defined lunes probably indicates a split, multiple, or twinned crystal. Translating the crystal along the spindle may allow finding a region which is single.

Mosaicity - more spots than expected for the oscillation range are produced if the mosaic spread of the crystal is high. From the image itself one can get some feel for mosaicity, but this should be checked after indexing (see below).

Overloads - saturated pixels are distinguished from others on the displays available at the stations: they appear in color on the Fuji scanner Image_Analyze display and in black on the CCD display. There should be no more than a few percent overloaded reflections in the resolution range of interest. To collect a wide range of intensities, it may be necessary to take multiple passes through the total oscillation range, with different exposure times.

Shadowing - it is possible for equipment such as a cooling nozzle to block part of the detector surface. This is usually obvious, but not always. In the case of a short exposure with relatively few spots (from a small molecule crystal, for example), one may need to look closely to detect the region where data are missing.

Spot separation - successful integration of reflections requires enough separation between them. The required distance depends on the spot size, but is typically about 10 pixels, as in the example of Figure 2a. The 6-pixel separation in Figure 2b. will clearly cause difficulty in integration and should be avoided if possible, either by moving the detector back or by narrowing the oscillation range.

Signal-to-noise - adequate peak to background ratio is needed for good data. Scaling by the image display program may make an image look fine when in fact it is not. A check of the background values may reveal the problem; backgrounds over about 1000 for image plates or 5000 for CCD's are suspicious. A program is being developed to give plots of background and signal-to-noise as a function of resolution, to aid in this aspect of image evaluation.

For the two images shown in figure 1, both crystals appear to be single. Although not shown in these figures, neither had excessive overloads in the resolution range of interest. A shadowed region is visible in Figure 1b, but only a small fraction of the data will be obscured by it. The spot separation is close but adequate for Figure 1a. Figures 1c and 1d, however, reveal a problem. Here is plotted the background (and a few peaks) along radial lines of the images in Figures 1a and 1b, respectively. The background in the Figure 1b image is clearly excessive. This high background is probably due to scattering from frozen solvent, either in or surrounding the crystal. It would be advisable to look for a better crystal, or to try mounting in a smaller loop if external solvent is the problem.

Indexing the Initial Image

Once a visually satisfactory image has been obtained, the crystal should be rotated, usually by 90 degrees, and another exposure taken, to check for anisotropic mosaicity, splitting that was not apparent on the first image, and any crystal centering problem. The latter is probable when diffraction is very weak or absent at the second spindle position but is fine on a repeat of the first exposure.

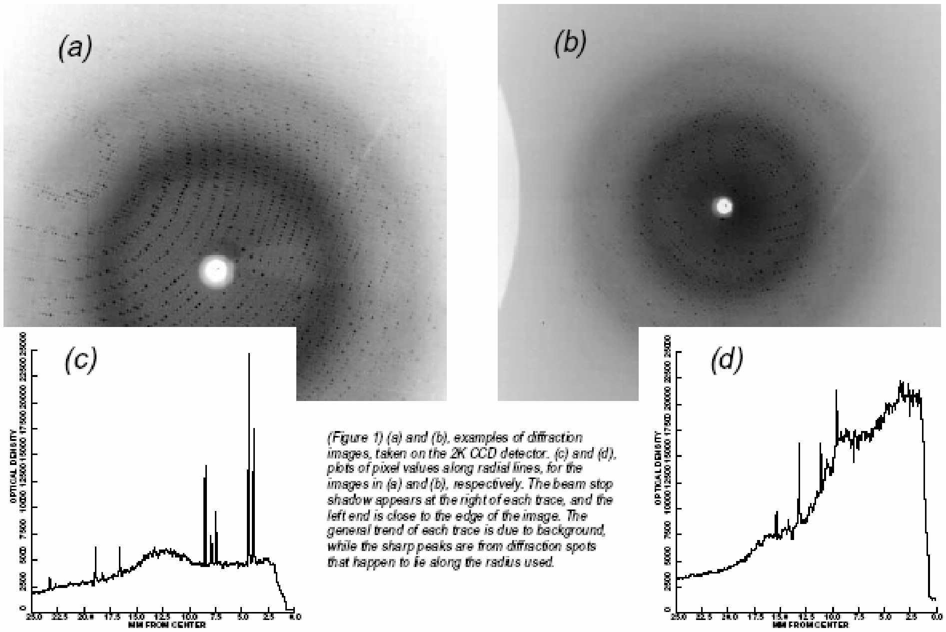
If the second image is good, it is time to index an image. This is easily done using Denzo (part of the HKL program package, by Z. Otwinowski and W. Minor); the only parameters needed are the

Completeness as a Function of Experimental Parameters

Crystal parameters: Space group C2, a = 108.2 Å, b = 63.1 Å, c = 54.5 Å, β = 110.8°
Resolution limits: 30.0 to 3.0 Å   Detector, 1K CCD at 75 mm distance

		Centered CCD			Offset CCD		
	Spindle range	% unique	% anom	Redundancy	% unique	% anom	Redundancy
b* perp to spindle	0-90	95	76	2.5	81	0.1	1.0
	45-135	49	56	3.4	19	0.2	1.0
	0-180	84	95	3.7	82	19	1.5
b* along spindle	0-90	50	49	3.9	45	39	2.3
	45-135	50	49	3.9	45	39	2.3
	0-180	85	83	4.3	85	74	2.3
General position	0-90	78	65	2.9	70	7	1.1
	45-135	65	59	3.2	58	3	1.0
	0-180	96	94	4.0	93	33	5.5

(Table 1) Completeness of data recorded, as a function of experimental parameters. The three crystal orientations shown are 1) c along the spindle (Denzo crystal rotation angles rotx=0, roty=0, rotz=0), 2) b* along the spindle (Denzo angles 0,90,0), and 3) a general orientation (Denzo angles (10,30,20). "% unique" gives the percentage of the unique reflections, ignoring anomalous dispersion, that could be recorded from the crystal by rotating it over the given range (in degrees) of spindle angles. "% anom" gives the percentage of anomalous pairs (Bijvoet mates) that would be recorded during the same rotation. "Redundancy" gives the average number of symmetry-related observations of each unique reflection that would be recorded, assuming that anomalous data are not needed. The redundancy of anomalous measurements (not shown) would be lower. These percentages take no account of losses due to overloaded or overlapping reflections. Data for table from m.simulate program.*



(Figure 1) (a) and (b), examples of diffraction images, taken on the 2K CCD detector. (c) and (d), plots of pixel values along radial lines, for the images in (a) and (b), respectively. The beam stop shadow appears at the right of each trace, and the left end is close to the edge of the image. The general trend of each trace is due to background, while the sharp peaks are from diffraction spots that happen to lie along the radius used.

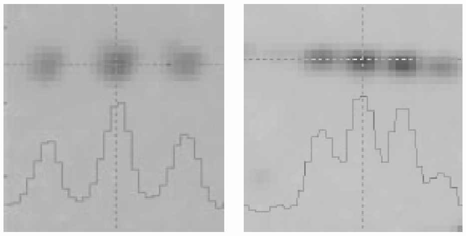
direct beam position and the crystal-to-detector distance. A successful indexing produces the result shown in Figure 3a. The predicted reflections, shown as green, yellow, and red circles, fall on or almost on the actual spots, and very few spots have no corresponding predictions. The predictions in Figure 3b were produced when an incorrect crystal-to-detector distance was supplied. This is the most common cause of a bad indexing. It is largely due to the difficulty of reading an accurate distance on the MacCHESS oscillation cameras; a new camera design will remedy this problem. A distance error makes all the calculated cell dimensions too high or too low; if the correct values are known it is easy to adjust the distance until the calculated values are reasonable. If the distance and direct beam position are correct, and the image has at least a few dozen good spots, the indexing should succeed. If not, the crystal may be twinned, so that the spots are not from a single lattice.

Once a predicted diffraction pattern has been generated, crystal mosaicity may be evaluated. If the predicted lunes are too narrow, the mosaicity is higher than the value supplied to Denzo (assuming the correct oscillation range). The mosaicity may readily be varied until the prediction matches the real pattern. If spots appear in between the predicted lunes and no value for mosaicity will account for them, the crystal is probably twinned, or has a satellite. There is no point in wasting time on such a crystal unless the extra spots are very few or all the crystals of the material are equally bad.

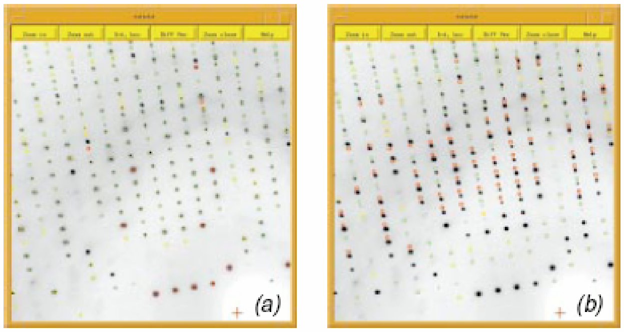
The appropriate oscillation range may be determined by making predictions for various ranges and checking for overlapping reflections. Using a mosaicity a little on the high side for safety, a range that is as wide as possible without generating more than a few overlaps may be selected. In some cases a narrower range than this may be desired, for the reason of reducing background. A few more test exposures may be needed to settle the question. If the unit cell dimensions are not all quite similar, predictions should be made for several spindle settings, as different oscillation ranges may be appropriate at different crystal orientations.

Completeness of Data Set

The ideal data set is 100% complete, with most reflections measured several times. Naturally, this is not always possible. From an indexed image, however, it may be determined how much of the



(Figure 2) Small regions from two diffraction patterns, each showing a row of spots. A plot of relative pixel value along the dashed horizontal line is superimposed on each display. (a), a region with 10-pixel spot separation (good). (b), a region with 6-pixel spot separation (inadequate for good spot integration).



(Figure 3) Portions of a diffraction image (in shades of gray) with predicted reflection positions superimposed (colored circles: green for fully recorded reflections, yellow for partials, red for "problem" reflections). (a), predictions from a good indexing. (b), predictions from an indexing done with an incorrect crystal-detector distance. Display from HKL package.

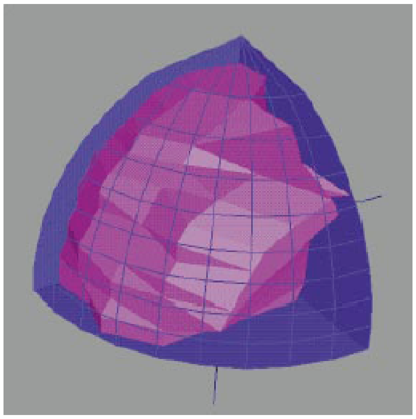
unique data can be collected on the crystal, and what range of spindle angles must be covered to get this fraction. A program now being developed at MacCHESS, "m.simulate", will aid in this determination.

The fraction of reciprocal space that must be covered to collect all the unique data to a given resolution depends on the crystal symmetry and on whether anomalous data are required. It is common to say that because a crystal is monoclinic, one must collect 180 degrees of data, or because it is tetragonal one only needs 45 degrees. In fact, the rotation range needed to collect the unique data depends on the orientation of the rotation axis relative to the unit cell axes, i.e. on the orientation of the crystal on the camera. In the real world, an additional factor is introduced by the limited area of the detector. For the CCD detectors in particular, recording resolved spots to high resolution may require offsetting the detector perpendicular to the x-ray beam. This results in some combination of a loss of redundancy and a loss of unique data for a given rotation range. Table 1 illustrates the effects of varying crystal orientation and detector position. This is a case where some data are off the edge of the detector if it is not offset, so that with the CCD centered even a 180 degree rotation of the crystal only gives about 90% of the unique data at best. If the crystal is aligned with c* along the spindle, only 90 degrees of rotation are needed to give the maximum completeness, but this maximum is only 76%. If the detector is offset, a complete data set may be obtained, but it requires taking a full 360 degrees of data if anomalous data are needed or if the crystal orientation is not optimum.

Crystal Orientation

Orienting a crystal with a symmetry axis along the x-ray beam can serve to minimize the rotation range required to collect a nearly complete data set. Alternatively, measurement of anomalous data may be facilitated by orienting the crystal to put Bijvoet pairs on each image. In the case of a unit cell with one long axis, placing that axis along the spindle allows wider oscillations to be taken than otherwise. The advantages of using an oriented crystal must be considered in light of the difficulty in scaling frames from a rotation series on such a crystal together, particularly in the lower symmetry classes. A data set from a second, differently oriented, crystal will probably resolve this problem. An additional consideration is that, for some symmetries, data collected by rotation about a symmetry axis will be incomplete no matter how many degrees of rotation are taken, due to the "missing cone" problem. Limitations in detector area may also become more important for oriented crystals. Figure 4 shows the fraction of unique reflections collected in a 360 degree rotation of a small molecule crystal. The figure represents a portion of reciprocal space: the blue surface encloses the total unique volume (to the limiting resolution of the crystal) for this orthorhombic cell; the magenta surface encloses the points corresponding to the unique reflections which were actually measured. Along the left-hand edge, the magenta surface is just inside the blue, showing complete coverage, but at the lower right a substantial number of the unique reflections were not collected. The crystal was oriented with b* near, but not on, the spindle axis; the CCD detector was offset, in order to get the desired resolution. The missing regions are due to a combination of limited detector size, "missing cone" effect, and a cooling nozzle shadow that was not obvious during data collection (due to the small number of spots per image). Although this image was generated using the reflections actually collected, the missing regions due to crystal orientation and detector geometry could have been predicted ahead of time using m.simulate, and the desirability of taking more data on a second, differently oriented, crystal would have been clear. In future, users

CONTINUED ON PAGE 52



(Figure 4) Fraction of the unique data collected on a small molecule crystal, displayed as two volumes in reciprocal space. The blue surface defines the volume occupied by all possible unique reflections for this orthorhombic cell, to the appropriate limiting resolution. The magenta surface encloses the unique reflections actually collected, in a 360 degree rotation, for this crystal. Figure drawn using the program Geomview from The Geometry Center at the University of Minnesota.

Continued from page 50

will be able to check the potential completeness of their data before taking it. An additional capability planned for m.simulate is that of reading in an earlier dataset and telling whether the current crystal will fill in gaps or merely replicate earlier data.

At CHESS, considerations of desirable crystal orientations are currently moot, as reorienting of crystals is limited to what can be done on the goniometer arcs. This may change in the future, however, particularly on the F2 station, and it is sometimes possible to influence a crystal's orientation during the mounting process.

To optimize the anomalous signal from a crystal not oriented with a mirror plane perpendicular to the spindle, it may be desirable to use the "inverse beam" approach: after a few degrees of data have been taken the crystal is rotated 180 degrees and the same amount of data collected. The second set of images will contain the anomalous mates of reflections on the first set. Note that this will only be true for all reflections if the detector is centered.

Preparing to Take Data

Enough information is now available to determine the experimental parameters for data collection. These are:

Exposure time - set to give few overloads in the resolution range of interest and a reasonably low background. Multiple passes with different exposure times may be necessary to get a wide resolution range. The minimum exposure time per degree is set by the maximum speed of the spindle motor. For strongly diffracting crystals, it may be necessary to attenuate the x-ray beam to avoid overloading.

Oscillation range - set to minimize number of exposures, while allowing few overlapping reflections and keeping background low. May vary with spindle setting.

Detector distance and offset - set to avoid having spots too close, while collecting data as close to the limiting resolution of the crystal as possible.

Limits of total oscillation - set by range needed to get the most complete data set possible for crystal's orientation. More than the minimum range may be taken if high redundancy is wanted.

Crystal orientation - set, if desired and possible, to minimize number of exposures or maximize quality of anomalous data. Except for rotation about the spindle, can only be controlled (at CHESS, now) to a limited degree, and would usually not be changed.

Exposures per fill - depends on exposure time, readout time for the detector being used, and of course the time remaining in the fill when data collection is initiated. If using the "inverse beam" method to optimize the anomalous signal, the time to rotate the spindle 180 degrees must also be included.

Special Cases

Short-lived crystals. The foregoing describes an approach of careful checking before starting data collection. This is appropriate when the crystal on the camera is frozen (as is now standard at CHESS) and will not be harmed by waiting for the 10-15 minutes it takes for full evaluation and planning. When the crystal is not frozen, or is subject to rapid decay even when frozen, it is better to just make a quick examination of an image for crystal splitting, spot overlaps, etc., and proceed directly to data collection. Then, while the next crystal is being mounted, the images just taken can be examined for mosaicity, etc. and parameters adjusted for other crystals in the batch.

Multiwavelength data collection. The efficient collection of data at multiple wavelengths, for MAD phasing, involves the same considerations as the monochromatic case. Aside from changing wavelengths between exposures and periodically taking energy scans, the data collection process itself is the same, and the same criteria are used to select good crystals, set the oscillation range, and so forth. Because of the importance of Bijvoet pairs in MAD phasing, it is necessary to take particular care with crystal orientation and detector offset, and data may be collected using the "inverse beam" approach. See the article "How to Go MAD at CHESS" on the next page for more details. In addition, the extra time needed for collecting each rotation range three or four times makes it especially important to optimize all experimental parameters, if the data set is to be completed in the time available.

Unusual data collection modes. For very weakly diffracting crystals, long exposures, on the order of an hour, may be required. If normal oscillation exposures are taken and the beam dumps halfway through one of them, the image will probably be too weak to be useful. A "long-exposure" mode is available, in which only one pass is taken through the oscillation range, with the spindle rotating in a series of small steps rather than continuously. In order to compensate for variation in beam intensity with time, exposure at each position is for a fixed number of counts, not a fixed time. With this mode, if an exposure is terminated prematurely one has a narrower oscillation range than desired, but reasonable exposure of the reflections that are present. Use of very long exposures requires special attention to background; in particular, the 2K CCD has relatively high dark noise, which may be a problem in such cases. Otherwise, the same considerations apply as for shorter exposures.

The approach of taking very narrow oscillation ranges ("fine phi-slicing") has been tested at CHESS. In this case, an initial exposure with a wide oscillation range should be taken to evaluate crystal mosaicity. Selection of exposure time is done using a narrow-oscillation image. Total rotation range is determined as usual. When considering the fine phi-slicing approach, the readout time of the detector becomes very important, as does the available disk storage for data frames.

Summary of Taking Data Efficiently

- Mount a crystal and take initial shot(s).
- Check for crystal problems, good exposure time, good spot separation.
- Index image, check mosaicity and oscillation range.
- Check potential completeness of data.
- Set experimental parameters and take data.
- Process data as soon as you can - plans are nice but the proof of the pudding is in the eating!