



# Some Practical Aspects of Data Collection and Processing

Sandy Blake

University of Nottingham, UK; A.J.Blake@nottingham.ac.uk

## **Introduction**

Unlike all other stages of a structure determination, the collection of diffraction data occurs in real time, requiring the continuous and exclusive use of valuable equipment. The aim is always to obtain the best possible quality of data from the available sample within a reasonable length of time. This requires not only having the best possible crystal, but also choosing appropriate experimental conditions and data collection parameters.

## **Collecting data with area detector diffractometers**

Until the mid-1990s, nearly all single-crystal data collection for chemical crystallography applications was carried out using four-circle (serial) diffractometers. Image plate (IP) area detector systems made some impact and the first commercial area detector instruments based on charge-coupled detectors (CCDs) were introduced in 1994. The advantages of area detectors over point detectors greatly outweigh any disadvantages. The major obvious feature of an electronic area detector is its ability to record diffraction data over a substantial solid angle. As far as normal Bragg diffraction is concerned, this means the simultaneous measurement of a number of reflections: the number of reflections which can be measured simultaneously depends on the size of the unit cell, the size of the detector and the crystal-to-detector distance. An area detector actually records the whole of the intercepted diffraction pattern and not just the Bragg reflections. This can be useful for special purposes, such as the detection of twinning or the study of incommensurate structures.

## **Experimental conditions**

### **Radiation**

The crystallographer may have some choice over the conditions under which data are collected and one of these is the wavelength of the radiation used, the most common choice in the home laboratory being between copper (1.54184 Å) and molybdenum (0.71073 Å). For the same power settings, Cu X-ray tubes produce a higher flux of incident photons and these are diffracted more efficiently than with Mo radiation: Cu radiation is therefore particularly useful for small or otherwise weakly-diffracting crystals, especially if absorption effects are moderate. In addition, focussing optics provide greater enhancements for the longer wavelength. For crystals with long unit cell dimensions, reflections are further apart when the longer wavelength Cu radiation is used and this can minimise reflection overlap. If you need to determine absolute configuration, and your crystals do not contain elements heavier than, say, silicon, then Cu radiation is essential. On the other hand, absorption effects are generally less serious with Mo radiation and this can be crucial if elements of high atomic number are present. Mo radiation allows collection of data to higher resolution and is likely to cause fewer restrictions if low-temperature or other attachments are required. Some diffractometer manufacturers offer hybrid Cu/Mo, Cu/Ag or Mo/Ag instruments: each is equipped with two different X-ray tubes and their associated optics but only a single (expensive) area detector is required. Some examples are given in the Table below.

Sample	Choice of radiation
poorly-diffracting organic compound	Cu to maximise diffracted intensity
organic compound with $b > 50 \text{ \AA}$	Cu to minimise overlap
absolute configuration on $C_{19}H_{29}N_3O_7$	only readily feasible with Cu
inorganic materials with very high $\mu$	Mo or Ag to minimise absorption
high-resolution studies	Mo or Ag
weakly-diffracting inorganic material	a 'no-win' situation?

Where a sample diffracts too weakly to yield enough observable data on a laboratory source, and it is considered important enough, it could be taken to a synchrotron radiation facility. In this event several advantages can accrue: the wavelength can be tuned specifically to the needs of the sample; the intensity typically represents at least a 100-fold gain over that of a laboratory source; the resolution between reflections is also generally better. The disadvantage is that there is heavy competition for such a rare resource and there is often a long wait for synchrotron beam time. For this reason, it is worth considering whether a more modest enhancement of the source intensity would overcome the problem of weak diffraction: a longer wavelength, a lower temperature or a more powerful laboratory source may suffice.

Lastly, when considering the most appropriate wavelength for a particular experiment, you should also bear in mind the physical restrictions that will preclude data collection in some areas of reciprocal space: these limitations include where the beam stop is positioned and where accessories such as low-temperature devices, cryostats or high-pressure cells are deployed. The restrictions are generally much more serious at longer wavelengths, since higher  $2\theta$  values are required to achieve the same resolution.

### Temperature

If a reliable low-temperature system is available it is almost always worthwhile considering data collection at low temperature (most low-temperature devices will even produce slightly elevated temperatures if this is required for a special experiment, while some have extended upper temperature limits). Low temperature is essential for crystals mounted using protective oil films, compounds melting below about  $50^\circ\text{C}$  and those which are thermolabile. There are more general advantages: at lower temperatures atomic displacements are reduced and the intensities of reflections at higher Bragg angles thereby enhanced, allowing the collection of better diffraction data at higher resolution. Low temperature also minimises librational effects which can otherwise give artificially shortened bond lengths and other systematic errors. One advantage of the reduction in temperature is the relative ease with which disorder can be modelled: for example, popular pseudo-spherical anions such as  $\text{BF}_4^-$ ,  $\text{PF}_6^-$ ,  $\text{ClO}_4^-$  and  $\text{SO}_4^{2-}$  are often badly disordered at room temperature but are either ordered or their disorder is much easier to model at low temperature. The fact that cooling methods typically involve rapid freezing of crystals on the diffractometer raises the question of what proportion of such crystals are determined as metastable rather than as equilibrium phases.

### Other conditions

Considerations of crystal size, methods of mounting, choice of goniometer head and optical centring have been mentioned previously but are worth stressing again as they can seriously affect the outcome of the experiment. The collimator selected should allow the entire crystal to be immersed in the X-ray beam but its diameter should not be excessive, as this will contribute to scattering by the air, resulting in increased background levels. This scattering is more serious with Cu than Mo radiation. The effects of an oversized

crystal are also strongly dependent on the elements in the sample: it has been shown that with only light elements (*e.g.*,  $Z < 8$ ) present, the use of large ( $\sim 2$  mm) crystals does not cause serious problems and may be beneficial because of the enhanced intensities which can be measured, or because the data collection time can be reduced (Gorbitz, 1999).

### **Detector systems**

#### **Image plate (IP)**

The phosphor in these devices stores the image in the form of trapped electron colour-centres. These are later 'read' by stimulation from visible laser light (which causes them to emit their own characteristic light for detection by a photomultiplier) and then erased by strong visible light before another exposure to *X*-rays. The main advantages of image plates are that they are available in large sizes and are relatively inexpensive; they also have a high recording efficiency and a high spatial resolution. The one major disadvantage is the need for a separate read-out process, which requires minutes rather than seconds. Faster image plate systems offer one solution but another method of reducing the time unavailable for *X*-ray exposure is the use of two or more plates, so that one is recording an image while another is being read: however, this adds considerably to the cost and complexity.

#### **Charge-Coupled Device (CCD)**

CCD-based detectors are familiar from camera phones and other mass-market applications. They employ a semiconductor in which incident radiation produces electron-hole pairs; the electrons are trapped in potential wells and then read out as currents. There is no direct recording of *X*-rays: instead, a phosphor is coupled through fibre optics to the CCD chip, which is cooled to reduce the inherent electronic noise level due to thermal excitation of electrons. Efficient recording, a high dynamic range and a low noise level, and a read-out time measured in seconds or fractions of a second, combine to give the CCD some clear advantages as a rapid area detector, but the size is limited by the size and quality of chips available.

#### **Solid-state detector**

Recent years have seen the introduction of so-called 'pixel arrays', a form of solid-state detector which records *X*-rays directly (*i.e.*, without conversion to visible light), and over a larger area. These detectors exploit the Si photoelectric effect and are under active development (see for example <https://www.dectris.com/technology.html>). Features include rapid, noise-free readout, excellent signal/noise ratio, the absence of any dark current and very high dynamic range and quantum efficiency. Unlike with CCDs, energy discrimination and fluorescence suppression are possible. Typically the shutter is open throughout the collection of a run, with the fine time-slicing performed by the detector. For integrating detectors such as the CCD, the reflection intensities will ideally utilise the full dynamic range of the detector, with any overloaded reflections being discarded or remeasured with beam attenuation and/or shorter exposure times. With a photon-counting detector we must also consider the count-rate threshold derived from the dead time for each pixel after it has measured a photon. Exceeding this count rate for a reflection will result in an underestimation of its intensity, but beyond using thin-slicing (with frame widths of 0.1 or 0.2°), the only way of bringing the count rate below the threshold is to attenuate the beam. The narrow frames can be merged in batches of 5 or 10 to give frames of effective width 0.5–2.0°.

#### **Some characteristics of CCD area detector systems**

Area detectors do not just offer the possibility of collecting diffraction data more quickly, although that is generally perceived as their principal advantage. They also render feasible experiments which are beyond the scope of older instruments, through higher sensitivity

and the recording of the whole pattern. Much of the following applies to any type of area detector, but refers specifically to CCD systems, since these are currently the most widely used in chemical crystallography.

An area detector, whether based on a CCD or an alternative technology, is only one component of a single-crystal *X*-ray diffractometer. It needs to be combined with a goniometer for mounting and moving the crystal sample, a source of *X*-rays, and electronic and computing control systems. Although an area detector records a number of diffracted beams simultaneously, it is still necessary to rotate the crystal in the *X*-ray beam in order to access all the available reflections. In most chemical crystallography systems, the detector is offset to one side rather than being held perpendicular to the incident *X*-ray beam, so that a higher maximum Bragg angle can be observed for a single detector position. Typical designs and configurations give data to a maximum Bragg angle of around 25–30°, appropriate for a 'routine' structure determination with a more than adequate data/parameter ratio if Mo- $K\alpha$  radiation is used.

### ***Crystal screening***

Although area detector data are quite tolerant of poorly centred crystals, it is obviously best to centre crystals as accurately as possible. Initial exposures can be recorded in a matter of a few seconds to give an almost instantaneous indication of the quality and intensity of diffraction by the crystal. It is a good idea to record frames at one or two different  $\varphi$  angles as this can detect crystals that look promising in one direction but show serious problems in another. Such exposures are taken with the crystal in a random orientation, and either stationary or oscillated through a small angle. At this stage, some obvious problems such as poor (or no) crystallinity, splitting of reflections and overall weak diffraction can be identified: an obviously unsuitable crystal can be quickly discarded and a hopefully better crystal selected. Note that such exposures are essentially two-dimensional and may not indicate all possible reflection splitting or other problems: these may only be detected when a series of frames have been collected for unit cell determination (and indexing may have failed – see below).

### ***Orientation matrix and unit cell determination***

Collecting a series of frames covering two or three small regions of reciprocal space takes a few minutes. After obtaining the positions of the reflections, including interpolation between successive frames to obtain precise setting angles, these are stored in a list by the control program. This yields coordinates of observed reflections in reciprocal space, referred to goniometer axes. It may be necessary in some cases to adjust the criteria for the inclusion of reflections in this list: this is most commonly done on the basis of  $I$  or  $I/a$  for more weakly diffracting crystals, where default settings may not yield enough reflections. An incorrect (or uncertain) initial matrix and cell at this stage are not a problem, as they can be revised following the full data collection without loss of data. One advantage in obtaining a good initial matrix and unit cell is that it is possible to check whether the latter corresponds to a known phase. It also provides some reassurance that you will be able to index and process the full dataset.

### ***If indexing fails ...***

If there is no convincing indexation, you should first check the indexing parameters. These may be hidden defaults but they actually control the indexing in various ways, including:

- (i) the maximum indices that can be assigned;
- (ii) the lengths of the axes of the original primitive cell;
- (iii) how far the indices are allowed to deviate from integral values;
- (iv) the minimum fraction of data which must be indexed on any candidate cell.

Parameters (i) and (ii) can be used to exclude ridiculously large cells, but must be reduced with care unless you know the cell beforehand. We recommend starting off with upper and lower limits on cell axes of 3 Å and 120 Å, respectively, and indices of 15 or 20. Parameter (iii) allows less well centred reflections to be included: a value of 0.1 may be too low in many cases, but values of 0.3 or higher are likely to generate multiple false cells. Tweaking parameter (iv) is useful if a twin is suspected and you are trying to index one component, but there are more sophisticated methods for handling such cases. If indexing fails try increasing the tolerance on parameter (iii). A visual survey of the frames may indicate whether increasing (i) or (ii) is likely to be effective.

### ***Re-harvesting the reflections***

If manipulation of the default list of reflections fails to provide a plausible indexing, it may be worthwhile harvesting reflections from the stored frames using your own criteria: these could include modified limits on  $I$  or  $I/\sigma(I)$ , resolution or other factors. Indexing on this list may prove more successful. If indexing does still fail, have a closer look at the frames, as examination of the rocking curves may show splitting or other effects not obvious on individual frames, which may mean that an initially promising crystal is in fact unsuitable. It is probably time to try another crystal if one is available. Even if indexing does work, you should always examine the rocking curves: indexing routines are robust and can often cope with split or broad reflection profiles, so a successful indexing does not automatically qualify a crystal for data collection.

### ***Still having problems?***

If you are faced with a persistent failure to index, and the reasons for this are not obvious, you may be able to call on a program which can display a reciprocal lattice plot of all the reflections harvested from the frames. By rotating the lattice you should be able to decide whether the crystal is single or not. Any significant non-singularity means the crystal should be discarded and another one examined. If all the available crystals show twinning the procedures outlined later can be adopted.

Even if indexing seems to have succeeded you should be wary of unexpectedly large, often triclinic, cells with large uncertainties on their cell dimensions. If a cell is allowed to be large and uncertain enough it will be able to accommodate almost any list of reflections. If indexing appears to succeed but some reflections have indices at or near special values such as 0.50 or 0.33, this may indicate the cell is too small in that direction and should be increased by a factor of two or three, respectively. If the situation is unclear it does not need to be resolved immediately, but can be revisited when the full area detector dataset is available.

### ***After indexing***

A successful indexing is followed by least-squares refinement of the orientation matrix and a Bravais lattice determination. The refinement also acts as a check on whether the indexing is valid: a plausible indexing which fails to give a satisfactory refinement is unsatisfactory. In order to be accepted a refinement must end with the vast majority of significant reflections having (nearly) integral indices, reasonable s.u.'s on cell parameters and suitable values for various quality indicators. It may not always be possible to be certain about the Bravais lattice.

### ***Checking for known unit cells***

There is usually little point in continuing to data collection if the unit cell from the indexing is known, either in the literature or locally. To exclude the first possibility you can use *CellCheckCSD*: it is available within the Bruker *APEX* and Rigaku Oxford Diffraction *CrysAlisPro* software packages, and also as a command-line tool (see

<https://www.ccdc.cam.ac.uk/Community/csd-community/CellCheckCSD>). To avoid repeating an in-house determination you should check against your in-house unit cells, for example by using a program such as LCELLS (O.V. Dolomanov, A.J. Blake, N.R. Champness and M. Schröder, *J. Appl. Crystallogr.* 2003, **36**, 955). Finding that your unit cell is already known is likely to be disappointing, but at this stage you have probably invested only a few minutes of diffractometer time in the crystal; if a known unit cell is not detected until after data collection and structure solution the duplication could waste hours of valuable instrument time.

### ***Unit cell volume***

You can calculate whether your cell volume is compatible with the expected molecular formula, using 18 Å<sup>3</sup> per non-hydrogen atom. This value is valid for a wide range of molecular compounds, although some adjustment might be required for special cases (from 14 Å<sup>3</sup> for some highly-condensed aromatic compounds to 23 Å<sup>3</sup> for some organosilicon compounds). The calculation is **not** valid for non-molecular compounds such as NaCl or FeTiO<sub>3</sub>. A significant discrepancy may indicate that the unit cell is incorrect, that the compound is not as proposed, or that solvent molecules are also present.

### ***Data collection***

It is necessary to set a small number of parameters to control the full data collection: their values are obtained from the preliminary screening measurements and unit cell determination. The most important factors affecting decisions concerning these parameters are the overall intensity level on the frames, the mosaic spread and the crystal symmetry.

### ***Intensity level and drop-off with Bragg angle***

The intensity level will help determine the frame measuring time, which should be long enough to give diffraction to adequate resolution: as a very crude rule of thumb, diffraction maxima should be clearly detectable to at least  $\frac{2}{3}$  of the required  $2\theta_{\max}$ . For strongly diffracting crystals, too long an exposure time may result in detector overloads: if this happens the time should be reduced provided this does not result in the loss of higher angle data. Collecting low-angle data relatively rapidly, while using longer exposure times for the higher angle data and allowing significant overlap of the two regions, is usually worthwhile with copper radiation.

### ***Mosaic spread***

The mosaic spread is established from the widths ( $x, y$ ) of the reflections on individual frames, plus an estimate of  $z$  from the width of the rocking curves. While there is little point in taking small steps through broad reflections, this may be highly advantageous for a crystal with narrow reflections. Broad reflections are not necessarily problematic, provided they can be separated from neighbouring maxima, and a greater scan width giving fewer frames may be appropriate and give the dataset more rapidly. Even if the scan parameters have been suggested by the diffractometer's strategy module, check that the combination of exposure time and scan width gives the required intensity.

### ***Crystal symmetry***

It is useful, but not essential, to establish crystal symmetry. The diffraction symmetry determines the minimum fraction of the whole-sphere diffraction pattern to be measured in order to achieve a complete dataset. If there is any doubt as to the correct diffraction symmetry, the lower symmetry should always be assumed. Although it is possible to calculate the optimum set of frame runs to achieve this completeness as efficiently as possible, for routine work it is sensible to collect the whole sphere of data for triclinic

crystals and at least a hemisphere for all other crystal systems. Multiplicity of observations is essential for correcting data, for example for the effects of absorption, and for the identification of outlier reflections.

### **Other considerations**

Although non-optimal settings of data collection parameters can cause problems, area detector systems are quite tolerant of such mistakes. In particular, collecting area detector data without a valid orientation matrix is rarely catastrophic provided valid indexing is eventually possible. Although other data collection parameters such as the rotation axis (*e.g.*,  $\omega$  or  $\varphi$ ) can also be varied, a typical data collection setup procedure is fairly simple and straightforward. Area detector systems can therefore be operated by less experienced workers with minimal risk of compromising data quality or completeness. Unless the data collection software takes a full set of pictures of your crystal, you must record its colour, shape and dimensions. If a list of indexed crystal faces is required for a numerical absorption correction, these will have to be measured. If you expect even very minor icing of your crystal during the data collection, make these measurements before you start. Finally, you could check that the X-ray generator settings are correct, that the flow of cooling water is adequate and stable and that any low-temperature device is operating at the desired temperature and has sufficient cryogenic fluid to last through the data collection.

### **Data reduction and correction**

Once the frameset has been acquired as described in the section on data collection, integrated intensities must be extracted from the raw frames, a process which is computationally intensive and impossible without access to a sufficiently precise orientation matrix. While the matrix found from the initial indexing may be adequate, and it may even be possible to initiate data reduction to run in parallel with the acquisition of the frames, the outcome may not be satisfactory and the best procedure may be to harvest reflections from the entire dataset, re-index (to check the unit cell) and re-refine the matrix using this longer list. Integration software uses the orientation matrix to determine the reflection positions, and the estimation of intensities can exploit the three-dimensional information available for each reflection through some form of profile fitting. There may also be facilities for updating and refining the orientation matrix during the integration process, to allow for uncertainties or gradual/sudden changes in the crystal orientation, but a high-quality matrix which is valid for all frames represents an ideal situation.

An integration program will also require the relevant calibration and correction files in order to apply any corrections that were not applied to the frames as they were being collected. Reflection spot shapes are determined for different regions on the face of the detector. These model shapes are then used for determining the area of integration for each reflection. The model profile shapes are also used to calculate correlation coefficients which can be used to reject data and to fit profiles to weak reflections.

### **Integration input and output**

Important input parameters include the reflection widths, which may be refined or fixed. In both cases trial integration runs can be used to obtain reasonable initial values. It may be better to err on the wide side but if the width is too great the integration boxes of neighbouring reflections will overlap and produce incorrect intensities. Once the integration program is running, it should produce some form of diagnostic output: examination of this (usually with the aid of a manual or other documentation) provides an indication of how the integration is proceeding. However, the volume of the raw output can be daunting, and if effective visualisation tools are not available many users will only

refer to the output if they subsequently encounter problems. Users are more likely to notice and act on the information if it is presented in an accessible graphical form.

A suitably constrained unit cell refinement (**Error! Reference source not found.**) should be carried out, either as part of the data reduction or separately, and should include a high proportion of the significant reflections (some software uses all reflections). Although the absolute number of reflections is always high, there may be examples where unit cells refined against a small proportion of the total data should be regarded with caution.

### Corrections

A description of the required corrections to integrated data have already been outlined and these are applied during the integration procedure. Lorentz and polarisation factors (which are instrument-specific) must be accounted for in all diffractometer measurements. There are various methods available for absorption corrections. Numerical corrections can be made on the basis of indexed faces, but routines which exploit the redundancy present in the data are more widely used. Note that redundancy is often rather low for triclinic crystals, and in such cases the corrections need to be assessed particularly carefully. Alternatively, empirical corrections are always available if all else fails. The range of correction factors output may differ significantly from the range predicted from the cell contents and crystal dimensions, due in part to the fact that the corrections can encompass systematic effects other than absorption.\*

### Output

Part of output from the data reduction routine usually includes analyses of data significance ( $I/\sigma$  ratios as a function of Bragg angle and other variables), data coverage, and redundancy and consistency among equivalent data under any assumed diffraction symmetry. Examination of any such output is strongly recommended: it might cause you to question your assumptions about the diffraction symmetry, the validity of the orientation matrix, the quality of the crystal or the completeness or resolution of your data. You may decide that you need to re-process the frames. In the most serious cases, and if the crystal is fortunately still on the diffractometer, you may even decide that you would feel safer collecting some more data. However, in most cases there will be no significant problems, and the dataset is ready for structure solution.

**Input file contains 5746 reflections for this component**

**Maximum allowed reflections = 25000**

**Wavelength, relative uncertainty: 0.7107300, 0.0000089**

**Orientation ('UB') matrix:**

0.0375682 -0.0244249 0.0626351

-0.0522494 0.0170492 0.0547843

-0.0979531 -0.0184620 -0.0322734

A	B	C	Alpha	Beta	Gamma	Vol
8.8205	28.535	11.582	90.000	104.686	90.000	2820.0

**Standard uncertainties:**

0.0009 0.003 0.002 0.000 0.002 0.000 0.9

**Range of reflections used:**

**Worst res Best res Min 2Theta Max 2Theta**

8.8119 0.7685 4.622 55.084

**Figure 1.** The crystal system is monoclinic, so the values of two interaxial angles (in red) are constrained to appropriate values

\*Other possible corrections: (a) Extinction predominantly affects strong, low-angle reflections and is normally corrected for by refining a single correction factor during structure refinement; secondary extinction is wavelength-dependent, being worse with Cu than with Mo radiation. (b) Thermal diffuse scattering (TDS) can artificially enhance the intensity of some high-angle reflections; the fact that TDS decreases with temperature provides yet another incentive to collect low-temperature data. (c) Multiple diffraction effects are more likely to occur if a prominent lattice vector is aligned with the rotation axis: they are most obvious and serious where they cause significant intensity to appear at the position of a systematic absence. If their significance is not noted they can cause problems with space group determination, especially if they

affect screw axis absences. They can also be recognised by their anomalously narrow reflection profiles. (d) Some data reduction programs will attempt to compensate for the effects of crystals which are larger than the X-ray beam: this does not appear to be problematic for crystals containing light elements, but avoiding such a situation is better than trying to correct for it.

### ***A typical experiment?***

The level of detail in the descriptions of data collection and processing may have obscured a key feature of CCD-based diffractometers, namely their simplicity of use: below is an outline of one experimental strategy where no particular problems were encountered. The times shown in brackets for each part of the procedure are *very* approximate.

1. Mount and optically centre the crystal (1 minute).
2. Using a small number of frames, screen the crystal for quality and known unit cells (2 minutes).
3. Collect some frames at higher angles to determine scattering power and its falloff (several minutes).
4. Check the resulting indexing, orientation matrix and Bravais lattice (1 minute).
5. Check that the proposed **exposure times** and **frame widths** are appropriate (1 minute).
6. Establish the **fraction of the sphere** to be collected (1 minute).
7. Record the crystal colour, shape and dimensions (1 minute); index faces if required (a few minutes).
8. Collect the data (several minutes to hours or days, depending on the quality of the individual sample).
9. Optionally, re-determine the matrix using all available/significant reflections (a few minutes).
10. Optionally, survey the frames visually as a check on the orientation matrix (a few minutes).
11. Process the data, applying corrections as required (several minutes).

Points **2–6** represent **decision points** where you must decide whether it is worthwhile continuing with the current crystal, and if so how the data should be collected. However, if indexing fails at points **2–4**, you might reasonably decide to continue in the hope of succeeding at point **9**. If difficulties are encountered during processing, revisiting points **9** and **10** is recommended.

Once a frameset has been processed to yield a file of reflections, analysis of these will establish the likely space group(s) from consideration of systematic absences and statistical intensity distributions.

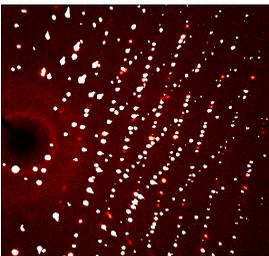
### ***Twinning and area detector data***

A comprehensive treatment of twinning appears elsewhere in the Course, but some comments regarding non-merohedral twinning are appropriate here. As mentioned above, it is not necessary to have an orientation matrix before initiating data collection on an area detector diffractometer, and one consequence of this is that framesets can sensibly be acquired on crystals which (may) have more than one component. Hopefully it will be possible to identify and index the components later, allowing the frames to be processed such that the intensity data corresponding to each component can be extracted. At one extreme, twinning may lead to a complete failure to index using the normal indexing routines; at the other, it may not be recognised until problems are encountered with structure solution or refinement; however, the procedures required to address the problem are identical. It is obviously helpful to recognise twinning as early as possible, otherwise you might waste a lot of time, for example by pursuing other solutions to an unsatisfactory refinement. At the refinement stage, previously undetected non-merohedral twinning may generate a number of symptoms, including:

- stubbornly high  $R$  indices, with no obvious cause such as disorder
- high difference Fourier residuals, again with no obvious cause
- individual reflections with  $F(\text{obs})^2 \gg F(\text{calc})^2$ , with certain indices most affected
- the lowest relative  $F(\text{calc})^2$  ranges show extreme values for certain indicators.

The first indications of non-merohedral twinning may be visible in the diffraction pattern: deviations from a regular lattice of well-shaped diffraction maxima may indicate non-merohedral twinning, although they could also indicate other problems. These features might include adjacent but incompatible (*i.e.*, mutually inclined) reciprocal lattice rows; a minority of reflections which do not correspond to the orientation matrix which fits the majority; and reflections which show splitting, overlap, irregular spacing or strange peak shapes (Figure 2). Although each case has to be assessed individually, the following is a general procedure for dealing with twinned area detector data:

- examine frameset for visible indications of twinning
- use pseudo-precession photographs or other visualisation aids
- identify major twin component, perhaps visually
- or using special software such as *DirAx*, *CrysAlisPro*, *CELL\_NOW*, etc.
- index the major twin component and save refined orientation matrix 1
- identify minor twin component
- index the minor twin component and save refined orientation matrix 2
- repeat the last two steps for any further minor components
- determine and view the relationships between the different matrices
- generate predicted patterns using matrices and check these against frames
- export orientation matrices for use by your data reduction program
- process the frameset using these orientation matrices
- output separate and combined datasets for solution and refinement.



**Figure 2.** Diffraction pattern from a crystal suspected of being a non-merohedral twin.

### ***Some other special cases (in brief)***

**Incommensurate structures.** A single unit cell and orientation matrix are not adequate in such cases because different parts of the structure exhibit different repeats. The stored frames can be processed to extract all the required data.

**Collection of powder or fibre data.** As the whole diffraction pattern is available, data can be extracted in different ways, for example by integrating intensity around a powder ring.

**Studying phase changes.** It is much easier to follow a phase transformation when the whole diffraction pattern is recorded.

**Diffuse scattering.** This can provide information about local structure (including disorder) in addition to conventional crystallographic data.

**Exploring lattice defects.** Again, the ability to monitor what is happening between the Bragg positions is valuable.