

## 25. MACROMOLECULAR CRYSTALLOGRAPHY PROGRAMS

set, and reports data completeness and quality in the file XSCALE.LP. The desired program action is specified in the file XSCALE.INP. It consists of a definition of shells used for analysing the resolution dependency of data quality and completeness, space-group number and cell constants, and one line for each set of input reflections. Each set has a file name, type identifier, a resolution window for accepting data, a weighting factor for the standard deviation of the reflection intensity, a decision constant for accepting Bijvoet pairs, a number controlling the degree of smoothness of the scaling function and an optional file name for specifying the output file.

*Problem:* All reflections beyond the highest shell specified for analysing the resolution dependency of data quality and completeness will be ignored regardless of the resolution window given for each data set.

## 25.2.9.2.4. VIEW

VIEW is used for visualizing data images as well as control images produced by XDS. It responds to navigation commands entered by movements of a mouse, and reports the corresponding image coordinates and their pixel contents upon activation of the mouse buttons. VIEW also allows magnification of selected image portions and changes in colour.

*Problem:* For many detector image formats as well as XDS produced images, the true pixel value is stored in a coded form which is interpreted by VIEW as a signed integer. Numbers less than -4095 displayed by VIEW correspond to large positive pixel values.

## 25.2.9.2.5. XDSCONV

XDSCONV accepts reflection-intensity data files as produced by XSCALE or CORRECT and converts them into a format required by software packages for structure determination. XDSCONV estimates structure-factor moduli based on the assumption that the intensity data set obeys Wilson's distribution and uses a Bayesian approach to statistical inference as described by French & Wilson (1978). For anomalous intensity data, both structure-factor amplitudes  $F_{hkl}$  and  $F_{\bar{h}\bar{k}\bar{l}}$  are simultaneously estimated from the Bijvoet intensity pair by a method similar to that described by Lewis & Rees (1983) – which accounts for the correlation between  $I_{hkl}$  and  $I_{\bar{h}\bar{k}\bar{l}}$ . The output file generated may inherit test reflections previously used for calculating a free  $R$  factor (Brünger, 1992*b*) or may contain new test reflections selected by XDSCONV.

## 25.2.9.3. Remarks

XDS is not an interactive program. It communicates with the input file XDS.INP and during the run accepts only a change in specification of the last image to be included in the data set (DATA\_RANGE=) – a useful option when processing overlaps with data collection. To prevent the program from overtaking the measurements, a maximum delay should be set (MINUTE=) to be slightly longer than the time for generating the next image.

Experience has shown that the most frequent obstacle in using the package is the indexing and accurate prediction of the reflections occurring in the images. Usually, the problems arise from incorrect specifications of rotation axis, beam direction or detector position and orientation, oscillation range, or wavelength. The occurrence of gross errors can be reduced by using file templates of XDS.INP specifically tailored to the actual experimental set-up which require only small adjustments to the geometrical parameters.

However, even small errors in the specification of the incident beam direction or the detector position may lead to indices which

are all offset by one reciprocal-lattice point, particularly if the initial list of diffraction spots was obtained from a few images covering a small range of crystal rotation. For this reason, IDXREF tests a few alternatives for the index origin and reports its results, such as the expected coordinates of the incident beam in the image, which can be checked by looking at a data image with the VIEW program. The user may then repeat the IDXREF step, thereby forcing the program to use a plausible alternative for the index origin.

It is recommended that all program steps are run on a few images to establish whether the indexing is correct and also to find reasonable values describing crystal mosaicity and spot size. Incorrect indexing may be apparent from large values of symmetry  $R$  factors or from comparison with a reference data set reported in the CORRECT step. Also, looking at the file FRAME.pck with the VIEW program should show the last data image processed with most of the observed diffraction spots circled. More accurate estimates of the parameters describing spot dimensions are reported in PROFIT.LP and should be used for updating these values in XDS.INP before starting data processing for all images.

Refinement of parameters controlling the predicted position of spots is carried out in the IDXREF, COLPROF and GLOREF step, which allows the user to adopt a variety of strategies. If all data images are available, spots should be extracted by COLSPOT from images equally distributed in the data set. If IDXREF is able to explain most of the spots, the refined parameters will be sufficiently accurate for the complete data processing, and refinements in the COLPROF step are unnecessary. In other cases, if processing overlaps with data collection or the first strategy was unsuccessful, IDXREF is based on spots extracted from the first few images and provides an initial parameter set which is periodically refined during COLPROF. This allows correction for slow crystal slippage or minor changes of the incident beam direction. Finally, if refinement in GLOREF was successful, the new values may be used to repeat COLPROF (without parameter refinements) and subsequent steps.

## 25.2.10. Macromolecular applications of SHELX

(G. M. SHELDRICK)

## 25.2.10.1. Historical introduction to SHELX

The first version of SHELX was written around 1970 for the solution and refinement of small-molecule and inorganic structures. In the meantime, it has become widely distributed and is used at some stage in well over half of current crystal structure determinations. Since small-molecule direct methods and Patterson interpretation algorithms can be used to locate a small number of heavy atoms or anomalous scatterers, the structure-solving program SHELXS has been used by macromolecular crystallographers for a number of years. More recently, improvements in cryocrystallography, area detectors and synchrotron data collection have led to a rapid increase in the number of high-resolution (<2 Å) macromolecular data sets. The enormous increase in available computer power makes it feasible to refine these structures using algorithms incorporated in SHELXL that were initially designed for small molecules. These algorithms are generally slower but make fewer approximations [*e.g.* conventional structure-factor summation rather than fast Fourier transform (FFT)] and include features, such as anisotropic refinement, modelling of complicated disorder and twinning, estimation of standard uncertainties by inverting the normal matrix *etc.*, that are routine in small-moiety crystallography but, for reasons of efficiency, are not widely implemented in programs written for macromolecular structure refinement. This account will be restricted to features in SHELX of potential interest to macromolecular crystallographers.