New CCP13 Software and the Strategy Behind Further Developments: Stripping and Modelling of Fibre Diffraction Data

John Squire[1], Hind AL-Khayat[1], Struther Arnott[1], Jane Crawshaw[3], Richard Denny[1,2], Greg Diakun[2], David Dover[1], Trevor Forsyth[4,5], Andrew He[1], Carlo Knupp[1], Geoff Mant[2], Ganeshalingam Rajkumar[1], Matthew Rodman[1,2], Mark Shotton[1,2] & Alan Windle[3]

[1] Biological Structure & Function Section, Biomedical Sciences Division, Imperial College London, London SW7 2AZ
[2] Synchrotron Division, CLRC Daresbury Laboratory, Warrington, Cheshire WA4 4AD
[3] Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ
[4] Institut Laue Langevin, 6 Rue Jules Horowitz, BP156 F-38042 Grenoble Cedex 9, France
[5] Lennard-Jones Laboratories, School of Chemistry and Physics, Keele University, Staffordshire ST5 5BG, UK

Fibre Diffraction Review **11**, 7-19, 2003

ABSTRACT

The aim of the CCP13 project is to provide easily usable software programs for those doing fibre diffraction or solution scattering studies; in fact any studies using diffraction from non-crystalline materials. Over the last ten years CCP13 project scientists have developed a set of programs for stripping and analysing fibre diffraction patterns and also patterns from solution scattering and other low-angle diffraction methods. The suite is now being updated to keep in step with new developments in beamlines, in detectors and in computing power and, in addition, new programs are being developed to allow for systematic modelling of different kinds of structures. The present report discusses the strategy behind the approach, it describes what has been implemented so far and it outlines new developments that are in the pipeline.

INTRODUCTION

Fibre diffraction has made critical contributions to our understanding of biomolecular and synthetic polymer systems and it will without doubt continue to do so as new resources and instrumentation become available to the scientific community. In the hands of adept practitioners the information content of X-ray and neutron diffraction data from fibres and other less than crystalline systems can be much greater than crystallographers generally concede. However, the problem of carrying out satisfactory data extraction and modelling of fibre diffraction patterns has still not been completely solved (Arnott, 1999). Furthermore, the outstanding success of the high-intensity X-ray beams and excellent area detectors at synchrotron X-ray sources, together with new applications of neutron beam sources has led to an accumulation of high quality experimental data which in the recent past could not be analysed satisfactorily due to a lack of appropriate software. In addition, massively increased computing power allows both biological and synthetic polymer questions to be addressed in a way that was not possible a few years ago. The present report summarises both the approach being taken by CCP13 personnel to consolidate and develop its data stripping software activity (Denny, 1993; 1994; 1995a,b; 1996; Shotton *et al.*, 1998; Shotton & Denny, 1999; Stubbs, 1999) and also the further expansion of its software development into the rigorous modelling of fibre diffraction data. The clear aim is to establish generally accepted objective standards of data stripping and modelling so that, as in the case of protein crystallographers using CCP4 programs, the quality and reliability of our diffraction analysis can be objectively assessed by others, and important scientific problems can be tackled with confidence. The new methods are being applied to key fibrous and non-crystalline targets in biology, medicine and materials science.

DEVELOPMENTS IN DATA ANALYSIS

So far, the major effort in CCP13 has concerned the extraction of useful intensity data from fibre diffraction patterns and from small-angle scattering (SAXS) data. The nature of the patterns from fibres

can be quite diverse, ranging from those which exhibit Bragg sampling from polycrystalline specimens, to patterns where the intensity is continuously distributed along layer-lines, reflecting a complete lack of axial alignment across the sample. Most programs in the CCP13 suite are designed to be as problem independent as possible: XFIX, FTOREC, LSQINT, FDSCALE, FD2BSL and SAMPLE are all dedicated to analysing two-dimensional fibre diffraction patterns. CORFUNC is a collective name for a set of programs for calculating correlation functions from SAXS data. XFIT is a one-dimensional peak-fitting program of general applicability, including its application to the analysis of time-resolved diffraction data. XCONV provides utilities for converting image files in a variety of detector formats to either BSL format (the format used by the NCD group at the CLRC Daresbury SRS) or to TIFF format. Achievements to date include the development of data reduction programs incorporating full mathematical and theoretical rigour, the incorporation of interactive and batch methodology for treatment of time-resolved diffraction data and the treatment of complex background data with low pass filters, roving windows and so on. Background treatment methods can also be applied to separation of crystalline from continuous data.

TYPICAL PROCEDURES IN FIBRE DIFFRACTION ANALYSIS

This Section provides an overview of the processes involved in data analysis and modelling and the CCP13 (or other) programs that can be used at each stage. These programs are then described in more detail in a later section.

Data Extraction

Once a fibre diffraction pattern has been recorded, and nowadays this is usually in digital form from an image plate, CCD or multiwire area detector rather than film, the image needs to be corrected for the particular properties of the detector (e.g. spatial distortion, detector response, etc). This is often most conveniently done using the software associated with the detector. Following this there is often a need to add diffraction patterns (e.g. from multiple short experiments of the same kind to build up enough counts), there may be a need to normalise various exposures according to the intensity of the incoming beam which will vary over time, there may be a need to subtract a camera or solution background and it may be necessary to do this over a whole time-series recorded in a single experiment. How this is done depends on the circumstances. Many of these manipulations can be done in the Daresbury NCD program BSL, once the files have been through XCONV to convert them from detector to BSL format (see Fig. 1), but BSL itself does not display the images and is not associated with a GUI. Alternatively, some operations can be done by the ESRF program FIT2D which instantly displays images and includes many useful manipulations (http://www.esrf.fr/computing/espg/subgroups/data_analysis/FIT2D). For example, FIT2D is very good for looking at 1D profiles across selected regions of the pattern. However, there is a new WINDOWS-based version of BSL called BS developed by Natalia Koubassova (http://jan.imec.msu.ru/~natalia/bs_form.htm) which may also be found useful. A new, improved program for these manipulations is currently being developed as part of the CCP13 suite.

Assuming that these preliminary stages have been completed, the next step is to carry out true fibre diffraction analysis. Figure 1 shows a flow diagram of the procedures that can be followed. The first step requires the determination of the tilt and rotation of the fibre relative to the vertical, finding the centre of the pattern and determining any tilt of the detector that may have been present. Once the files are in BSL format, this process can be done in XFIX (Fig. 1) which displays and analyses these parameters. The values coming out of XFIX can then be used in the program FTOREC to produce a corrected, centred, quadrant-folded image. This consists of intensities correctly mapped to reciprocal space (e.g. initially curved layer-lines from a highly tilted fibre will become straight). The image from FTOREC needs to be viewed in XFIX. Subtraction of a smoothly varying background can also be carried out in XFIX or it can be done in LSQINT, which is the next program in the sequence.

LSQINT is designed either to estimate the integrated intensities within Bragg peaks in patterns from a polycrystalline specimen or to determine smoothly varying intensities along continuous layer-lines. Among other things, account needs to be taken: (i) of the Lorentz correction (outer reflections will spend less time crossing the Ewald sphere and will be under-represented), (ii) of the peak shape and (iii) of the disorientation within the specimen. LSQINT will fit a smoothly varying background using several

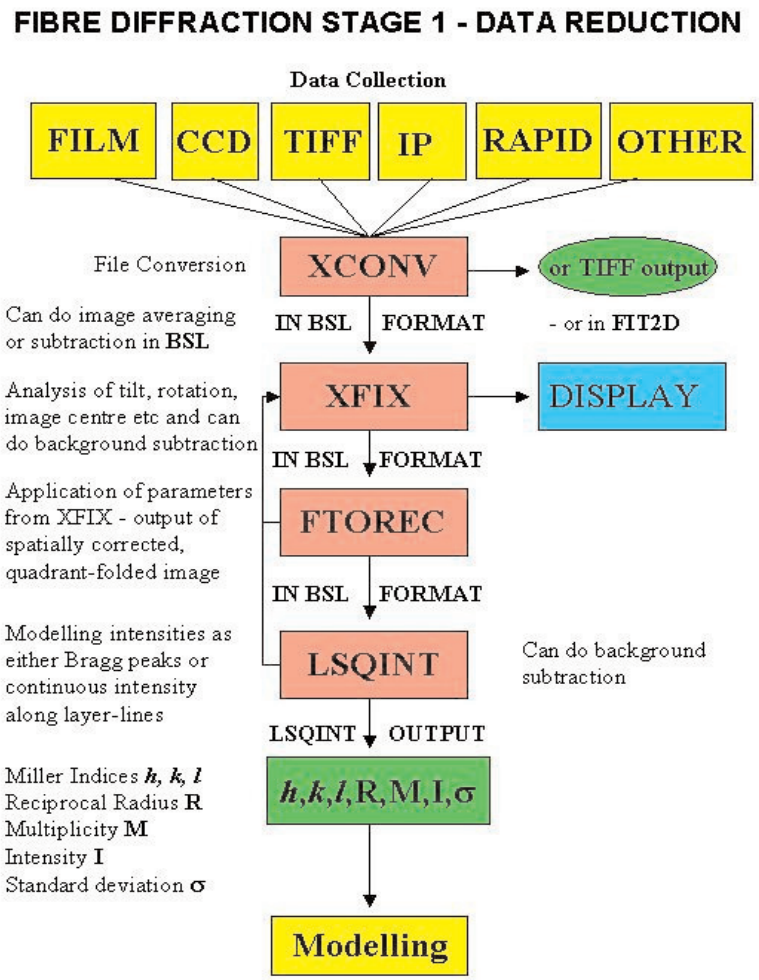

Figure 1: Flow diagram to show the processing procedure for treating a raw fibre diffraction pattern from a variety of detector sources through the file conversion program XCONV, into the analysis and display program XFIX which outputs parameters such as fibre tilt and rotation for input into FTOREC where the pattern is corrected and remapped in reciprocal space. Once again it can be viewed in XFIX. The corrected data from FTOREC can be fitted in LSQINT either as continuous intensities or as Bragg peaks after suitable background removal. Intensities output from LSQINT can then be used in modelling processes outlined in Figure 2.

alternative procedures if this has not been done in FTOREC.

In the case of well-sampled layer-lines (Bragg peaks) the important output parameters from LSQINT (apart from peak shape and disorientation) are:

[1] the ***h***, ***k***, ***l*** values for a particular peak,
[2] the multiplicity **M** of the peak. (The fibre pattern is essentially a section taken through a 3D transform that has been rotated continuously through 360° such that reflections at the same radius along a layer-line are added. The number of such reflections (i.e. the multiplicity) will vary depending on the space group involved and on the values of **h** and **k** for the particular peak and some peaks will be weighted more strongly than others.)
[3] the integrated intensity (**I**) of the peak and the standard deviation (**σ**).

In the case of continuous layer-lines the important features are the layer-line number (***l***), the reciprocal radius along the layer-line (**R**) and the values of **I** and **σ**.

Modelling

Once good intensity values have been determined from LSQINT, the problem is how to determine the associated structure. The strategy to be followed (Fig. 2) depends on the resolution of the information in the diffraction pattern. Patterns containing high resolution diffraction peaks, say from a nucleic acid, polysaccharide or crystalline polymer, need to be fitted at an atomic level, but, unlike protein crystallography where electron density maps are

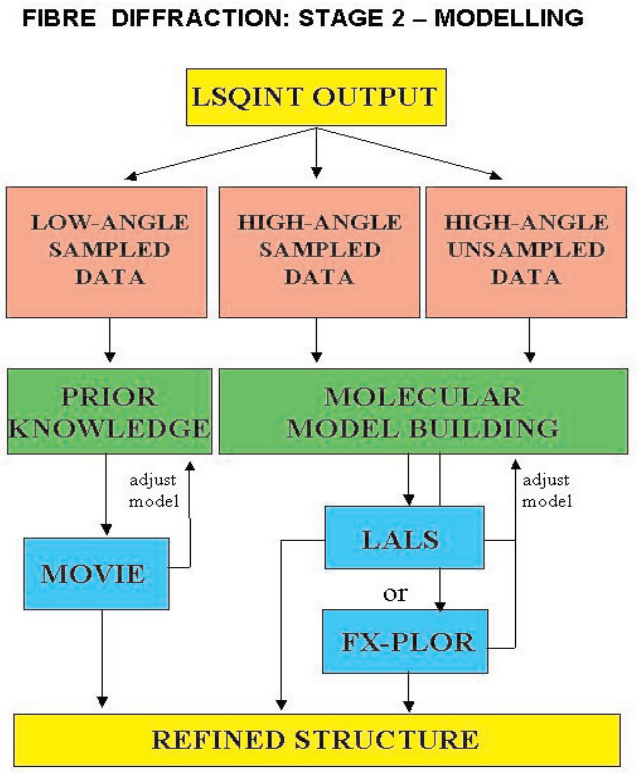

Figure 2: Alternative steps in the modelling procedure depending on whether the analysis is carried out at atomic resolution (high-angle pathways) or at the level of protein sub-domains (low-angle pathway). The programs in blue boxes are still in development.

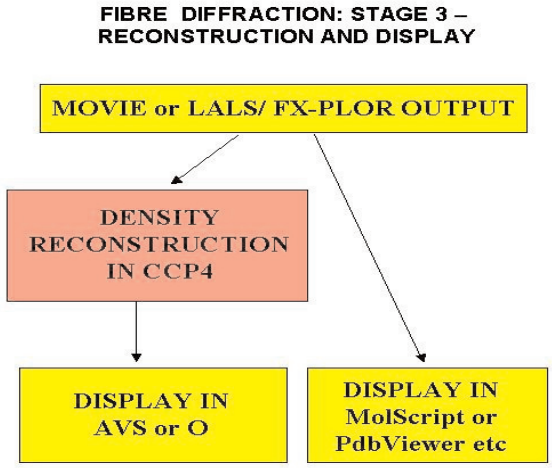

Figure 3: The final steps in the modelling process depend once again on the resolution of the data being processed. Models from high-resolution data can be viewed in conventional molecular graphics programs (e.g. MolScript, PdbViewer). In the case of low-resolution Bragg data the output from MOVIE or another such program can be used as an input to CCP4 programs which can calculate electron density maps at the required resolution that can be displayed in O or on AVS.

computed and fitted, this has to be done by creating a possible model, comparing its computed diffraction pattern with the observations coming from LSQINT, and then refining the model against the data. CCP13 currently uses two alternative programs to do this, LALS and FX-PLOR (see below), each of which also has the facility to minimise the energy of the structure being studied, along with assessing goodness of fit to the X-ray diffraction data using a crystallographic R-factor. Once the best atomic coordinates have been determined, the structures can be displayed using conventional molecular graphics programs (e.g. PdbViewer or MolScript: Fig. 3).

In the case of low-angle X-ray diffraction data, where the resolution is limited, it can be helpful to use protein sub-domains as the repeating units and to parameterize the domain positions, compute the diffraction patterns and vary them to give the optimal fit to the intensities coming from LSQINT (Fig. 2; see Squire *et al.*, 2003 - this volume). However, final testing of any model should be carried out where possible with the full atomic coordinates of the component molecules. In the case of diffraction from muscle, the program MOVIE will take in molecular shapes, change the positions of sub-domains within molecules, and optimise all of these things by simulated annealing against the observed diffraction data as assessed by a computed R-factor (Hudson *et al.*, 1997). MOVIE is currently designed for a specific task, but expansion of these ideas to more general objects (i.e. other than muscle) is under way. Once the best parameters have been determined (Fig. 3), the structure can be transformed or reconstructed using CCP4 programs (http://www.ccp4.ac.uk) and computed density maps can be viewed, for example, in O (Jones *et al.*, 1991) or in AVS.

CCP13 DATA CONVERSION AND STRIPPING PROGRAMS

This section builds on the overview of the last section, it summarises the technical details of the available programs and it also reports on new developments that have already been implemented or are planned. Platforms on which the various programs are implemented are shown in Table 1.

TABLE 1: CCP13 PROGRAM STATUS AND PLATFORMS

Program	Description
XOTOKO	1-D data manipulation
BSL	2-D data manipulation
V2A	vax to unix data conversion
A2V	unix to vax data conversion
OTCON	ascii to otoko data conversion
RECONV	otoko to ascii data conversion
TIFF2BSL	image plate (tiff) to bsl conversion
BSL2TIFF	bsl to tiff conversion
I2A	ieee to ansi data conversion (DEC)
XCONV	file format conversion (GUI-driven)
XFIT	1-D fitting and plotting (GUI-driven)
XFIX	fibre pattern analysis (GUI-driven)
CONV	file format conversion (cmd line)
FTOREC	reciprocal space transformation
LSQINT	2-D integration and background fitting
CORFUNC	correlation function analysis
SAMPLE	Fourier-Bessel smoothing
FDSCALE	scaling and merging of intensities
FD2BSL	intensity to bsl conversion

Program	Solaris 2.7	Irix 6.2	OSF 3.2	Linux	Windows
XOTOKO	31/03/03	30/05/96	29/04/97	31/03/03	-
BSL	31/03/03	21/03/97	27/04/97	31/03/03	-
V2A	19/05/95	-	-	-	-
A2V	19/05/95	-	-	-	-
OTCON	06/06/95	08/07/94	-	08/05/97	-
RECONV	06/06/95	31/10/94	-	08/05/97	-
TIFF2BSL	01/2002	01/2002	-	01/2002	-
BSL2TIFF	21/03/97	-	-	-	-
I2A	n/a	n/a	29/04/97	02/05/97	n/a
XCONV	08/11/02	08/11/02	08/11/02	08/11/02	15/04/03
XFIT	06/12/99	17/08/99	17/08/99	17/08/99	-
XFIX	06/12/99	17/08/99	17/08/99	17/08/99	-
CONV	25/05/99	25/05/99	25/05/99	25/05/99	05/02/03
FTOREC	10/07/99	10/07/99	10/07/99	10/07/99	05/02/03
LSQINT	09/01/03	08/01/03	08/01/03	09/01/03	05/02/03
CORFUNC	25/05/99	22/04/99	-	22/08/99	-
SAMPLE	22/04/99	22/04/99	15/07/99	22/04/99	-
FDSCALE	22/04/99	22/04/99	22/04/99	22/04/99	-
FD2BSL	22/04/99	22/04/99	19/02/02	19/02/02	-

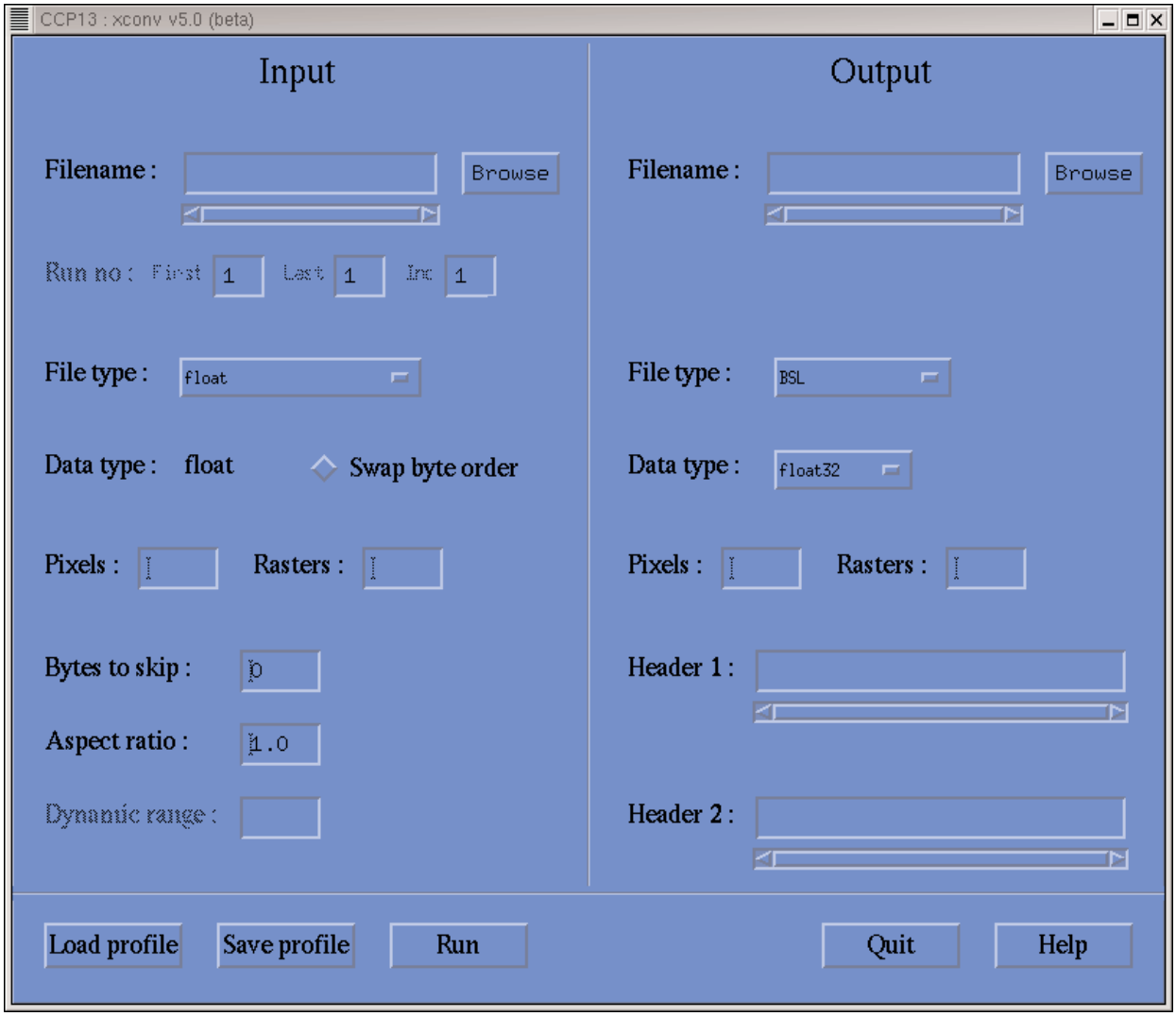

Figure 4: Illustration of the XCONV graphical user interface.

2-D Image Data File Conversion

XCONV *R. Denny (updated Ganeshalingam Rajkumar - December 2002*)

XCONV provides for the conversion of various image data files (e.g. from a variety of detector systems) to either **BSL** or **TIFF** format with the aid of an OSF/Motif-based graphical user interface (GUI). It replaces the CONV program. The XCONV window (Fig. 4) is split into two sections relating to the input image data file(s) and the output BSL file (header file + data file) or TIFF file.
The following input file types are allowed: **float** - float, **int** - unsigned int, **short** - unsigned short, **char** - unsigned char, **smar** - small MAR image plate (1200 x 1200), **bmar** - big MAR image plate (2000 x 2000), **fuji** - Fuji image plate (2048 x 4096), **fuji2500** - BAS2500 Fuji image plate (2000 x 2500), **rax2** - R-Axis II image plate, **rax4** - R-Axis IV image plate, **psci** - Photonics Science CCD, **riso** - RISO file format, **tiff** - TIFF (8,12,16-bit greyscale with all Compression type), **ESRF Id2 (KLORA)** - ESRF Data Format, **LOQ 1D** - one dimensional ASCII data files recorded at LOQ, **LOQ 2D** - two dimensional ASCII data files recorded at LOQ, **SMV** - used by ADSC CCD detectors (8 bit unsigned, 16 bit unsigned, 32 bit signed integer, 32 bit floating), **ESRF Id3** - ESRF Data Format, **BRUKER** - Area detector Frame format Siemens/Bruker file format (8 ,16 or 32 bits per pixel), **Mar345** - MAR345 Image Plate (File extension mar1200, mar1800, mar1600, mar2400, mar2000, mar3000, mar2300, mar3450), **BSL** - the Daresbury NCD BSL file format as described in the BSL manual.

Output formats are: **BSL** - The BSL file format (suitable for input to XFIX, FTOREC and LSQINT), **TIFF8** - 8 bit greyscale tiff image and **TIFF16** - 16 bit greyscale tiff image.

2D Intensity Stripping

XFIX: *R. Denny (updated Mark Shotton - April 99; Ganeshalingam Rajkumar - December 2002*)

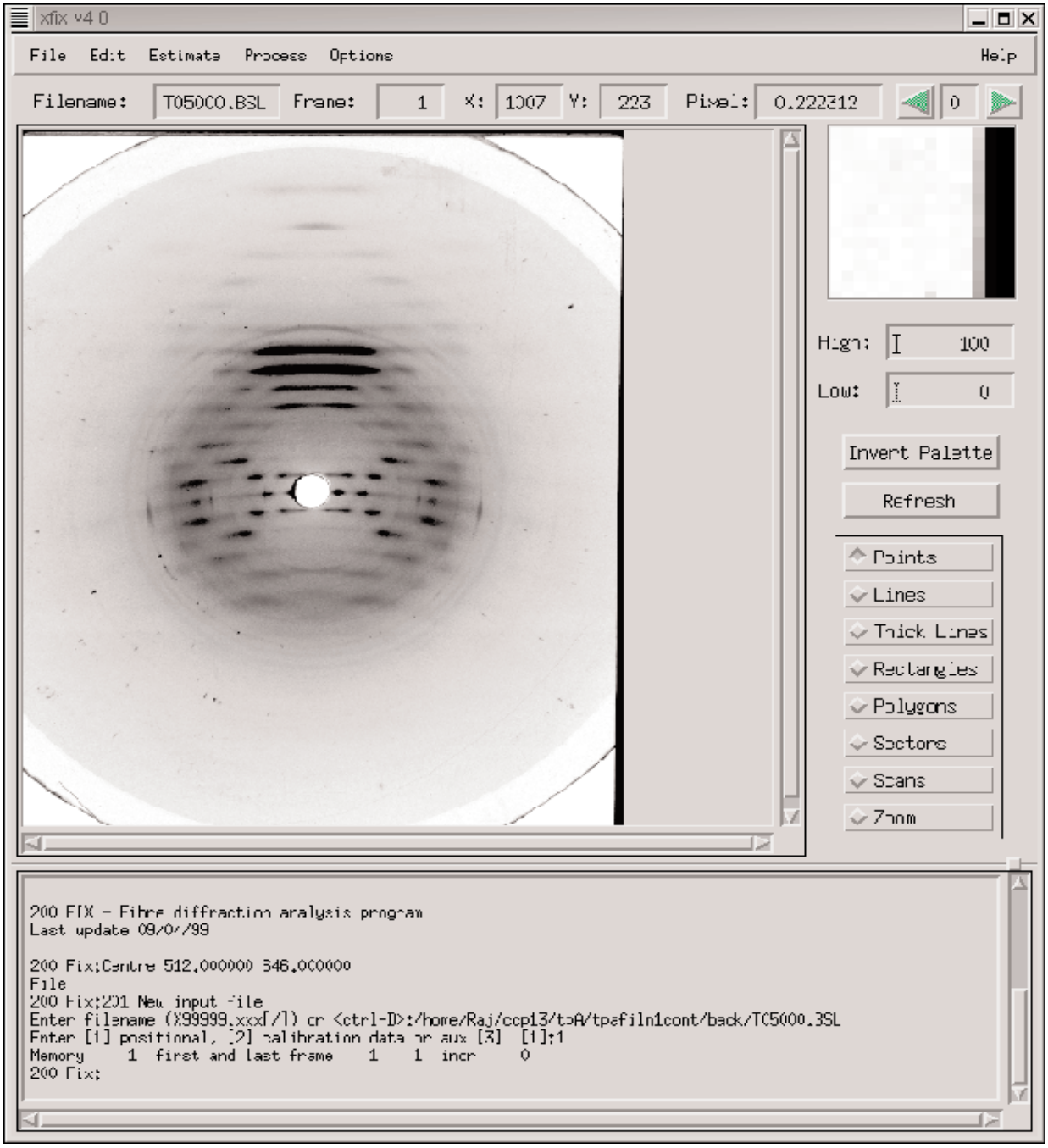

Figure 5: A raw fibre diffraction pattern visualised in XFIX.

XFIX is designed to help ascertain important parameters of a fibre diffraction pattern with the aid of an OSF/ Motif-based graphical user interface (GUI; Fig. 5). The job of the program is to handle the image display, collect information specifying regions of the data and to communicate with the program FIX which performs much of the underlying data-processing. Information such as the pattern centre, detector orientation and fibre tilt can be estimated and refined and putative unit cells can be plotted over the pattern. In addition, more general functions are available such as the capability to plot and fit integrated slices or scans through the pattern. The program also allows the use of three background subtraction tools - the choice of which will depend on the type of data being studied. The background options are: (a) Paul Langan's "roving window" method, (b) a circularly-symmetric background and (c) calculation of a "smoothed" background through iterative low-pass filtering based on the method of Ivanova and Makowski (1998).

FTOREC: *R. Denny (updated Mark Shotton April 99; Ganeshalingam Rajkumar - December 2002*)

This program is designed to take all or part of a diffraction image and transform this portion of the image into reciprocal space, given knowledge of the specimen to film distance, orientation of the image,

XFIX IN ACTION

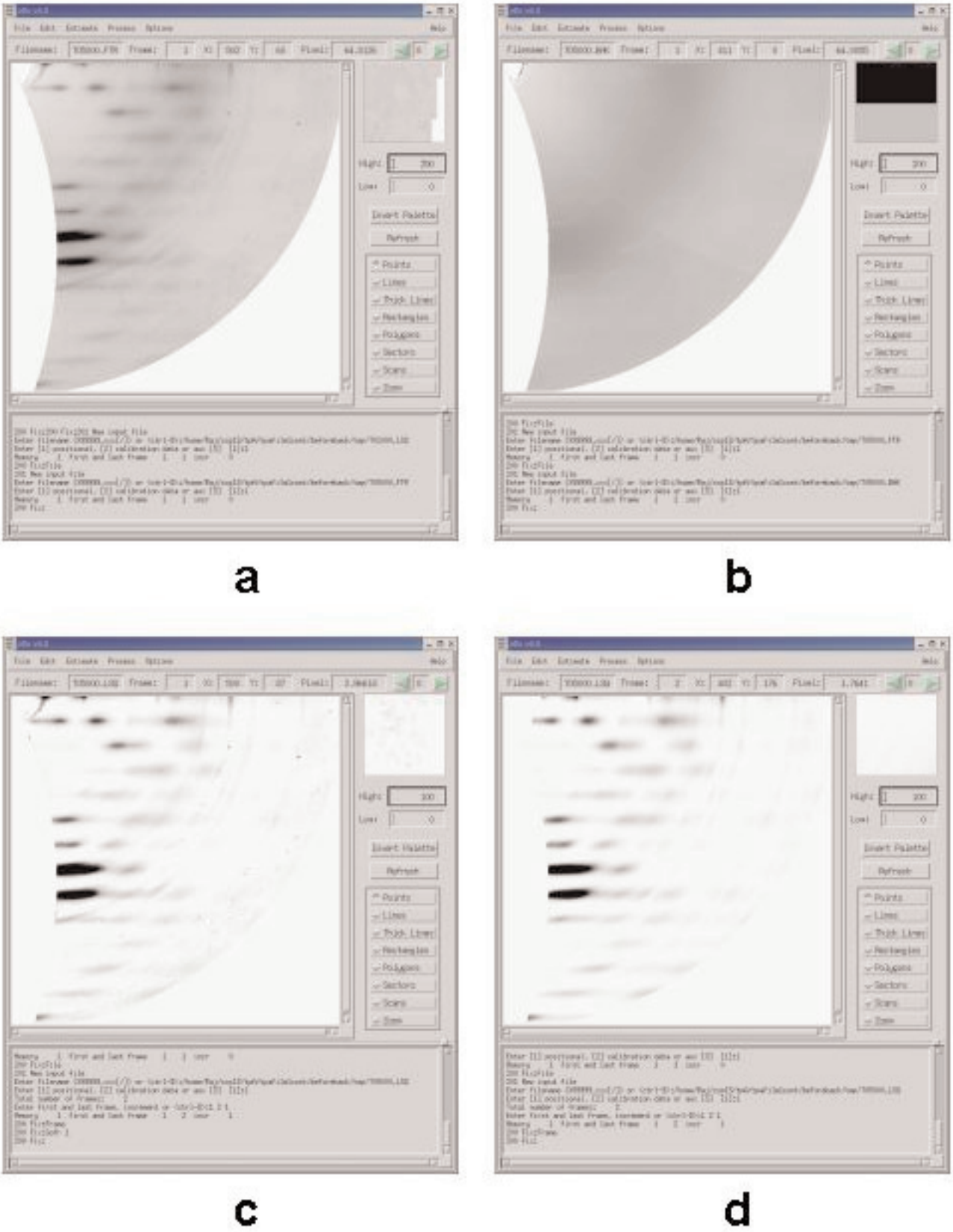

Figure 6: Stages in the processing of the pattern in Figure 5, all displayed using XFIX. **(a)** The spatially-corrected, quadrant-folded version of Figure 5 using parameters determined in XFIX and implemented in FTOREC. **(b)** The estimated background in (a). **(c)** Pattern in (a) after removal of the background in (b). This represents the Bragg diffraction peaks that need to be fitted by LSQINT. **(d)** The best fit to the intensity distribution in (c) as determined using LSQINT, with the LSQINT output redisplayed in XFIX using both the peak intensity values and the estimated peak shapes, fibre disorientation angle etc as needed to get a good fit in LSQINT.

wavelength and tilt of the specimen determined in XFIX. Either polar or Cartesian reciprocal space outputs are available. The effect of this transformation is to straighten out the diffraction pattern and to apply appropriate scaling to detector pixel values as described in Fraser *et al* (1976). The corrected, quadrant-folded image can then be re-displayed in XFIX (Figs. 5, 6)

LSQINT: *R. Denny (updated Mark Shotton - April 99; updated for WINDOWS (DOS) Ganeshalingam Rajkumar - January 2003)*

The purpose of LSQINT is to provide an automatic method for the integration of intensities for fibre diffraction data. The program will handle patterns which are largely crystalline in nature, or patterns which have intensity continuously distributed along layer-lines, with sampling only occurring parallel to the c-axis. More than one lattice can be fitted in a single pattern. The approach used is to generate spot profiles and then to fit these profiles to the observed pixel values by a robust linear least-squares method or using the maximum entropy method of Skilling and Bryan (1984). A typical result is shown in Figure 6. Several background subtraction methods are available: (a) a global background can be fitted simultaneously with the spot profiles, (b) a circularly symmetric background can be fitted, (c) a notional roving aperture and filter can be used to define a background, or (d) a background plane can be fitted in a region of the pattern along with the spot profiles, the background at the centres of the regions being joined by an interpolating surface (Shotton *et al*., 1998). Alternatively, the background subtraction can be done in XFIX (see above) prior to entry into LSQINT. Refinement of the cell and profile parameters is also available by a downhill-simplex method. This can be cycled with the profile fitting to produce an overall refinement procedure.

Figure 6(a) shows an original FTOREC-treated fibre diffraction pattern as viewed in XFIX. Figure 6(b) shows the estimated background. Figure 6(c) shows the pattern in 6(a) minus the background - it represents the Bragg peaks to be fitted by LSQINT. Finally Figure 6(d) shows the LSQINT fit to the peaks when replotted in XFIX with the same peak profiles, disorientation etc parameters that give a good fit in LSQINT.

SAMPLE: *R. Denny*

This program performs Fourier-Bessel smoothing of intensities extracted from continuous layer-line data as produced by LSQINT, using the formula described by Makowski (1982). The data are fitted using singular value decomposition to give intensity values at the specified number of knot points.

FDSCALE: *R. Denny*

FDSCALE provides a means of scaling intensities produced in LSQINT from different data sets.

FD2BSL: *R. Denny*

This converts the ASCII output from LSQINT (or FDSCALE) output files to BSL format, allowing flexibility in the output image dimensions and $I/\sigma(I)$ cut-offs for intensity rejection. The output file has one raster per layer-line and a user-defined number of pixels per raster. For Bragg-sampled data, the intensity of a reflection is divided by its multiplicity so that a representation of the cylindrical average of the squared molecular transform is given.

One-dimensional Peak Fitting

XFIT: *R. Denny (updated Mark Shotton - August 98)*

XFIT provides a flexible means of fitting various peak shapes to 1D intensity data and is particularly suited to fitting a sequence of time-resolved data. A polynomial and/or exponential background can be fitted along with the peaks and simple constraints can be applied between pairs of parameters. Once an initial model has been specified, the program can run through a sequence of data frames refining the previous best parameter set to fit the current frame. If the fit worsens unacceptably between consecutive frames, the user is warned and the starting model for the current frame can be specified interactively. An illustration of XFIT in action is given in Figure 7.

Small-angle Scattering

CORFUNC: *T.W. Nye (updated Richard Denny, Mark Shotton - 1999; updated Matthew Rodman - 2002/3)*

This is a suite of programs for calculating one and three-dimensional correlation functions (Γ_1 and Γ_3) from small-angle scattering (SAXS) data. Operation of the programs can be broken down into three logical steps:

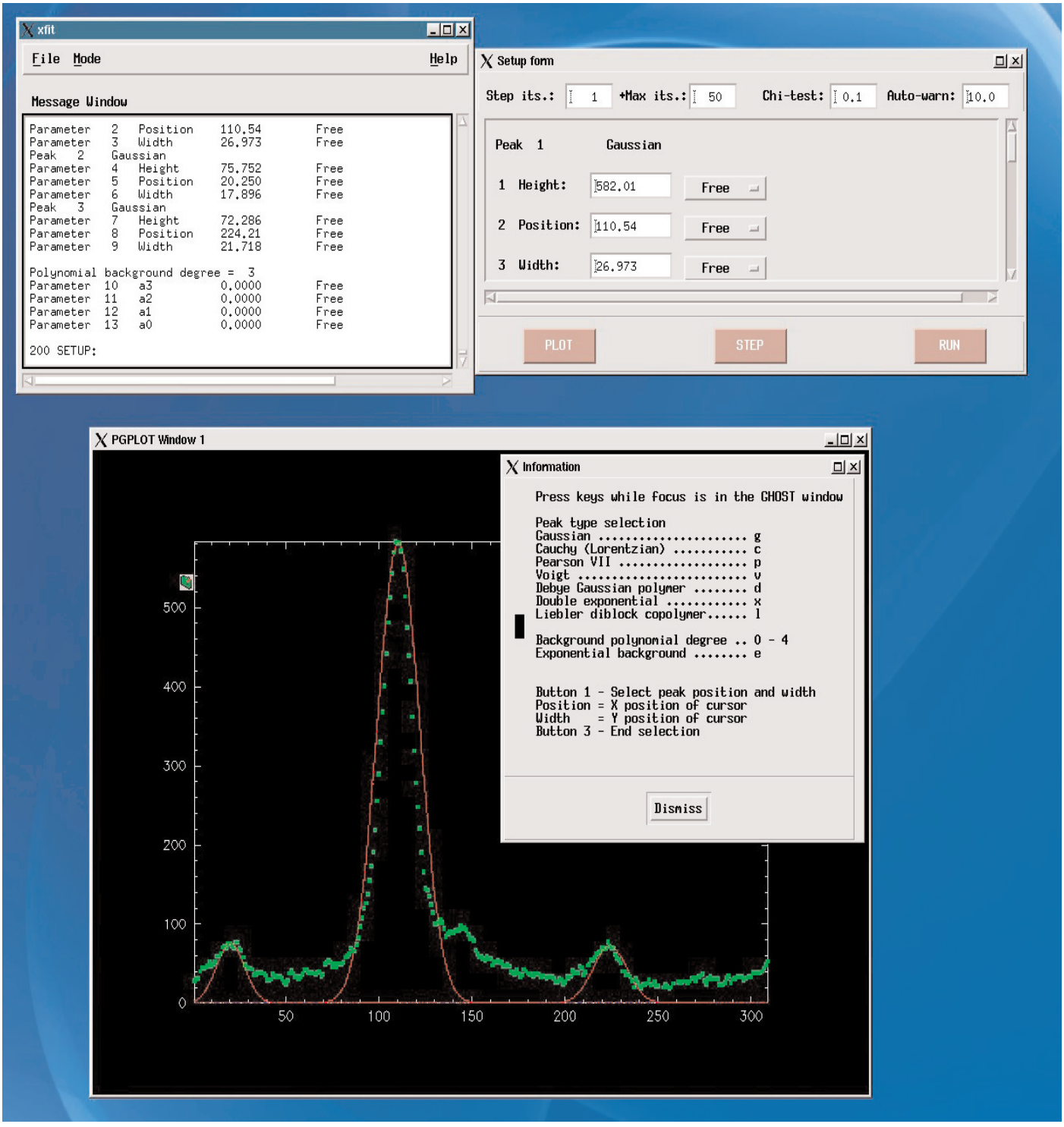

Figure 7: An example of XFIT in action showing modelling peaks roughly in position on a plot of experimental data points (lower window) before any optimisation has been carried out.

(i) Extrapolation of the experimental SAXS curve to $q = \infty$ and $q = 0$ (TAILFIT, TAILJOIN).
(ii) Fourier transformation of the extrapolated data (TRANSFORM).
(iii) Interpretation of Γ_1 based on ideal lamella morphology (EXTRACT).

The individual programs involved are given in brackets; they are run in sequence by the CORFUNC shell script.

CORFUNC has recently been modified so that it is more flexible in handling ASCII input; it is now less strict about numeric format and file termination. Also, the error-handling mechanism in the underlying EXTRAPOLATE.F program has been substantially re-written with the aim of being less ambiguous; the previous mechanism was somewhat degenerate (that is, many diverse error sources mapped to the same error code and message). More

importantly, however, is the change made to CORFUNC's convergence behaviour in the module TRANSFORM.F. Here, the maximum **q**-value permissible has been dramatically extended to enable data processing with larger **q**-ranges. This has led to only a minor increase in transform convergence times, inconsequential when compared to overall process duration.

Future developments planned for Corfunc involve adding the option of employing a Hilbert transform. This type of analysis is used to generate segment density distribution functions of adsorbed polymers from SANS data. Currently there are no public domain programs to do this. The work will thus bring CCP13 to the attention of a new audience and will be undertaken with the kind assistance of Steve King (ISIS).".

A useful tutorial on analysing solution scattering data is given at:
http://www.srs.dl.ac.uk/ncd/Saxs_Training_School_overheads/SoftwareTrainingWorkshop.html

DEVELOPMENTS IN HIGH-RESOLUTION MODELLING

FX-PLOR: *A Patch for X-PLOR: (Richard Denny, Mark Shotton - 1999/2000)*

In full consultation with Prof. A. Brunger, CCP13 has written and developed a patch (Denny *et al.*, 1997) that enables X-PLOR (Brunger *et al.*, 1987) to be used with fibre diffraction data. The CCP13 patch follows that created by Wang & Stubbs (1993) for the use of continuous fibre diffraction data. This allows (a) connectivity between symmetric units of crystallographic or strict non-crystallographic symmetry to be maintained and (b) diffraction data to be treated as composite intensities or overlapping Bessel function contributions. For the poly-crystalline part of the patch where the situation is more likely to require approximate helical symmetry and strict crystallographic symmetry, the symmetry linkage routines were modified so that a continuous chain can be built passing from asymmetric unit to asymmetric unit or through the bottom and top of the unit cell. The target functions were also redefined to take account of reflection overlap and the residual function changed to deal with composite intensities rather than structure factor moduli. The patch is a CCP13 creation and does not work without X-PLOR itself, which must be acquired using the normal routes for academic or commercial users.

LALS - Linked-Atom Least Squares analysis

Linked Atom Least Squares (LALS) has proved to be a very successful method of analysing fibre diffraction data from microcystalline fibres (Arnott & Wonacott, 1966; Arnott *et al.*, 1969). It allows the refinement of a stereochemically valid model against observed diffraction intensities with constraints and restraints to force non-crystallographic repeat geometry and intra- and inter-molecular contacts. A Windows version of LALS has been developed and will be available for incorporation in the CCP13 suite in the coming year. It is our intention to maintain a single distributed executable which runs also under Linux using WINE.

Several support programs are being developed for LALS mainly concerned with display and manipulation of the molecular model. Once these have been further developed in practice they too will be incorporated into the CCP13 suite.

Outputs can be viewed using MolScript or PdbViewer.

DEVELOPMENTS IN MODELLING MACROMOLECULAR ASSEMBLIES

Over the last decade the low-angle diffraction facilities at Daresbury have gradually improved and developed so that lines 2.1, 8.2 and now 14.1 are excellent cameras for low-angle and high-angle fibre diffraction work and line 16.1, with its high intensity, together with the RAPID detector system (see below), currently represents, because of the speed of the detector, the best available beamline worldwide for very fast time-resolved low-angle X-ray diffraction experiments on dynamic processes. Such processes include phase transitions in synthetic polymer systems (seconds) and molecular movements in muscular contraction (milliseconds and less). Other beamlines at third generation sources (e.g. line ID02 at the ESRF and the BioCAT beamline at the APS), because of their superb brilliance, can produce very high quality diffraction patterns on image plates or on CCD detectors, but the readout speed remains relatively slow.

The Daresbury detector group has pioneered the development of very fast 2-dimensional multiwire

proportional counters. The resulting RAPID detector allows previously impossible time-resolved low-angle diffraction studies to be carried out. For example, the whole of the low-angle diffraction pattern from muscle (spacings between, say, 600 Å and 40 Å) can be followed with 1 ms time resolution during typical contraction cycles (Harford & Squire, 1992; 1997; Squire, 1998; 2000). The most intense parts of the pattern can be recorded with 100 µs time resolution (Irving *et al.*, 1992; Lombardi *et al.*, 1995). Now that such experiments can be done, and even more sophisticated experiments are in the offing, it is crucial that the data being recorded so successfully should be rapidly stripped and analysed to produce definitive results about the particular specimen involved.

A key aspect of the on-going CCP13 programme is the development of software to model low-resolution structures (e.g. macromolecular assemblies), to compute their transforms and to refine them against the observed, peak-fitted intensity data. Some highly effective, subject-specific modelling programs are already being developed (e.g. the program MOVIE which carries out simulated annealing modelling of muscle diffraction patterns; Hudson *et al.*, 1997; Squire *et al.*, 1998; Squire 2000), but what is needed is more general modelling software that will be applicable to most kinds of fibrous molecular assemblies. This is a demanding task, but with the advent of much faster computers and parallel processing options, the problem of refining models against observed data is now becoming tractable and is one of the aims of the current CCP13 project.

Future Modelling Strategies

Another approach, complementary to that already used, but possibly more powerful when fibre disorientation causes layer-line overlap at higher radii, is to use the correctly mapped whole diffraction patterns as targets against which to do 'whole pattern fitting' by modelling (e.g. Lorenz *et al.*, 1993). Thus a model, as described above, could also include all the fibre disorder factors as parameters, so that the whole diffraction pattern, including background and overlapping layer-lines, is modelled and refined in one process.

CONCLUSIONS

In summary, good software to extract intensity data from diffraction patterns produced by non-crystalline materials is available as part of the CCP13 package. Although these programs work extremely well, we are currently producing a front end JAVA driven program that will integrate all aspects of the measurement software from raw data to reduced intensities whether they relate to a crystalline system or to continuous diffraction (i.e. it will link XCONV, XFIX, FTOREC and LSQINT into a common program). We believe that this front-end overlay will reduce the "energy barrier" for new users as well as providing addition useful functionality. At the same time existing programs are being modified and updated and new software is being developed to model the observed intensities at high or low resolution in terms of realistic molecular structures.

In the final analysis, a primary objective of CCP13 is to solve structures by fibre X-ray diffraction in as rigorous and objectively tested a manner as is being done, for example, in protein crystallography. We need to define objective assessment procedures that can be demonstrated to be effective and which meet with general acceptance so that structures solved from fibre diffraction patterns are considered to be just as reliable as those of globular proteins treated using CCP4 tools.

Acknowledgements

The work of CCP13 described here is supported by a project grant (#25/B15281) from the BBSRC & EPSRC, and directly employs Jane Crawshaw, David Dover (60%), Andrew He, Ganeshalingam Rajkumar and Matthew Rodman (50% BBSRC/EPSRC, 50% CLRC).

REFERENCES

[1] Arnott, S. (1999) Lessons for today and tomorrow from yesterday. *Fibre Diffraction Review* **8**, 21-27.
[2] Arnott, S. & Wonacott, A.J. (1966) Atomic co-ordinates for an α-helix: refinement of the crystal structure of α-poly-l-alanine. *J. Mol. Biol.* **21**, 371-83.
[3] Arnott, S., Dover, S.D. & Wonacott, A.J. (1969) Least-squares refinement of the crystal and molecular structures of DNA and RNA from X-ray data and standard bond lengths and angles. *Acta Cryst.* **B25**, 2192-2206.
[4] Brunger, A.T., Kuriyan, J. & Karplus, M. (1987) Crystallographic R-factor refinement by molecular dynamics. *Science* **235**, 458-460.
[5] Denny, R.C. (1993) Integration of fibre diffraction patterns. *Fibre Diffraction Review* **2**, 5-8.
[6] Denny, R.C. (1994) CCP13 Program Developments. *Fibre Diffraction Review* **3**, 12-16.
[7] Denny, R.C. (1995a) CCP13 Program Updates. *Fibre*

Diffraction Review **4**, 8-10.
[8] Denny, R.C. (1995b) Assessment of the Torus algorithm for global optimisation. *Fibre Diffraction Review* **4**, 11-14.
[9] Denny, R.C. (1996) Fibre diffraction spot profiles and the Lorentz correction. *Fibre Diffraction Review* **5**, 24-27.
[10] Denny, R.C., Shotton, M.W. & Forsyth, V.T. (1997) X-PLOR for polycrystalline fibre diffraction. *Fibre Diffraction Review* **6**, 30-33.
[11] Fraser R.D.B., Macrae T.P., Miller A. & Rowlands R.J. (1976) Digital processing of fibre diffraction patterns. *J. Appl. Cryst.* **9**, 81-94.
[12] Harford, J.J. & Squire, J.M. (1992) Evidence for structurally different attached states of myosin cross-bridges on actin during contraction of fish muscle. *Biophys. J.* **63**, 387-396.
[13] Harford, J.J. & Squire, J.M. (1997) Time-resolved studies of muscle using synchrotron radiation. *Reps. Progr. Physics* **60**, 1723-1787 (1997).
[14] Hudson, L., Harford, J.J., Denny, R.J. & Squire, J.M. (1997) Myosin head configurations in relaxed fish muscle: resting state myosin heads swing axially by 150Å or turn upside down to reach rigor. *J. Mol. Biol.* **273**, 440-455.
[15] Irving, M., Lombardi, V., Piazzesi, G. & Ferenczi, M.A. (1992) Myosin head movements are synchronous with the elementary force-generating process in muscle. *Nature* **357**, 156-158.
[16] Ivanova, M.I. & Makowski, L. (1998) Iterative low-pass filtering for estimation of the background in fiber diffraction patterns. *Acta Cryst.* **A54**, 626-631.
[17] Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst.* **47**, 110-119.
[18] Lombardi, V., Piazzesi, G., Ferenczi, M.A., Thirlwell, H., Dobbie, I. & Irving, M. (1995) Elastic distortion of myosin heads and repriming of the working stroke in muscle. *Nature* **374**, 553-555.
[19] Lorenz, M., Popp, D. & Holmes, K.C. (1993) Refinement of the F-actin model against X-ray fiber diffraction data by the use of a directed mutation algorithm. J. Mol. Biol. 234, 826-836
[20] Makowski, L. (1982) The use of continuous diffraction data as a phase constraint. II. Application to fiber diffraction data. *J. Appl. Cryst.* **15**, 546-557.
[21] Shotton, M.W. & Denny, R.C. (1999) CCP13 software development. *Fibre Diffraction Review* **8**, 14-19.
[22] Shotton, M.W., Denny, R.C. & Forsyth, V.T. (1998) CCP13 Software development. *Fibre Diffraction Review* **7**, 40-44.
[23] Skilling, J. & Bryan, R.K. (1984) Maximum entropy in image reconstruction: general algorithm. *Mon. Not. R. Astr. Soc.* **211**, 111-124.
[24] Squire, J.M. (1998) Time-resolved X-ray Diffraction. In '*Current Methods in Muscle Physiology: Advantages, Problems and Limitations' IUPS Commission on Muscle Physiology (H. Sugi, Ed.)* 241-285. Oxford University Press.
[25] Squire, J.M. (2000) Fibre and Muscle Diffraction" in *Structure and Dynamics of Biomolecules (E. Fanchon, Geissler, L-L. Hodeau, J-R. Regnard & P. Timmins, Eds)* pp. 272-301. Oxford University Press, Oxford, UK.
[26] Squire, J.M., Cantino, M., Chew, M., Denny, R., Harford, J.J., Hudson, L. & Luther, P.K. (1998) Myosin rod packing schemes in vertebrate muscle thick filaments. *J. Struct. Biol.* **122**, 128-138.
[27] Squire, J.M., Knupp, C., AL-Khayat, H.A. & Harford, J.J. (2003). Millisecond Time-Resolved Low-angle X-ray Fibre Diffraction: A Powerful, High-Sensitivity Technique for Modelling Real-Time Movements in Biological Macromolecular Assemblies. *Fibre Diffraction Review* **11**, 28-35.
[28] Stubbs, G. (1999) Developments in fiber diffraction. *Curr. Opin. Struct. Biol.* **9**, 15-619.
[29] Wang, H. & Stubbs, G. (1993) Molecular dynamics in refinement against fiber diffraction data. *Acta Cryst* **A49**, 504-513.