Acta Cryst. (1998). D**54**, 367–377

Structure Refinement Against Synchrotron Laue Data: Strategies for Data Collection and Reduction

XIAOJING YANG,[a] ZHONG REN[a,b] AND KEITH MOFFAT[a,b]*†

[a]*Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA, and* [b]*Consortium for Advanced Radiation Sources, The University of Chicago, Chicago, IL 60637, USA.*
E-mail: moffat@cars.uchicago.edu

(*Received 9 May 1997; accepted 27 August 1997*)

Abstract

The synchrotron Laue technique has been applied to high-resolution structure refinement of the ribotoxin, restrictocin [Yang & Moffat (1996). *Structure*, **4**, 837–852]. By employing carefully designed data-collection strategies and the data-reduction algorithms incorporated in the software system *LaueView* [Ren & Moffat (1995*a*). *J. Appl. Cryst.* **28**, 461–481; Ren & Moffat (1995*b*). *J. Appl. Cryst.* **28**, 482–493], a set of high-resolution Laue data with a completeness and accuracy comparable to excellent monochromatic data was obtained. Through detailed comparison with the monochromatic data and electron-density maps derived from the Laue data, optimum data-collection and reduction strategies were identified and the application of Laue diffraction techniques to conventional crystallographic refinement was demonstrated.

1. Introduction

Successful structure refinement of macromolecules is based on the accuracy and completeness of measured structure-factor amplitudes over the complete resolution range, to the highest resolution to which crystals diffract. With brilliant synchrotron X-ray sources, data sets with satisfactory completeness and accuracy can be obtained for weakly diffracting small crystals, and even for microcrystals. Data sets for conventional crystallographic studies are often collected at a single X-ray wavelength selected by a monochromator from the broad wavelength range emitted by synchrotron X-ray sources. In contrast, Laue diffraction utilizes many wavelengths simultaneously across the entire X-ray spectrum, which enables strong diffraction patterns to be collected with the briefest exposure (Szebenyi *et al.*, 1988; Moffat, 1997). The Laue diffraction technique has therefore been used largely when data-collection speed is a major concern, as in time-resolved crystallographic studies (Hajdu & Johnson, 1990; Cruickshank *et al.*, 1992; Singer *et al.*, 1993; Cassetta *et al.*, 1993; Srajer *et al.*, 1996; Ren *et al.*, 1996; Genick *et al.*, 1997).

Several data reduction algorithms and associated software have been developed to overcome two major hurdles of the Laue technique, namely the 'energy overlap' and 'spatial overlap' problems (Cruickshank *et al.*, 1987, 1991; Hajdu & Andersson, 1993; Helliwell *et al.*, 1989; Shrive *et al.*, 1990; Hao *et al.*, 1993; Campbell & Hao, 1993; Wakatsuki, 1993); but successful applications are limited in numbers. Although the Laue method has been widely used in time-resolved studies, the reported completeness of Laue data sets seldom exceeds 70% (Hajdu & Johnson, 1990; Schlichting *et al.*, 1990; Stoddard *et al.*, 1991; Vellieux *et al.*, 1993; Singer *et al.*, 1993; Duke *et al.*, 1994; Bradbrook *et al.*, 1995), partly due to the additional difficulties which accompany time-resolved data collection and to limitations of the algorithms used in these particular applications to treat spatial and energy overlaps. Even after deconvolution of doublet spots (Laue spots containing two harmonic reflections), the highest completeness that was achieved with the program *LEAP* is about 62% (Duke *et al.*, 1994). Singer *et al.* (1993) reported a 68% completeness for a Laue data set reduced by the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994; Helliwell *et al.*, 1989; Campbell, 1993) without harmonic deconvolution. Bourgeois *et al.* (1997) very recently reported successful structure refinement against a high-resolution (1.5 Å) Laue data set of cutinase acquired with 150 ps, single pulse exposures. Their Laue images were integrated by a new integration procedure (D. Bourgeois, unpublished results), wavelength normalization was carried out by the *CCP4 Laue* suite (Helliwell *et al.*, 1989; Campbell, 1993), but no harmonic deconvolution was carried out. Their electron-density maps calculated from the Laue data set which is 71.7% complete are generally of satisfactory quality, but a few regions of the backbone exhibit discontinuities in electron density which were attributed to the incompleteness of the data. Series termination errors in a Fourier synthesis resulting from the incompleteness of a Laue data set greatly affect the interpretability of electron-density maps. This problem is accentuated by the non-random distribution of

† Current address: Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, 2153 Sheridan Road, Evanston, IL 60208, USA.

missing data in reciprocal space. In addition, the reliability of structure-factor amplitudes derived from the Laue images may be low, as suggested by high values of merging and refinement R factors (Bartunik & Borchert, 1989; Vellieux *et al.*, 1993; Duke *et al.*, 1994).

With the new Laue data-reduction algorithms encoded in the *LaueView* software (Ren & Moffat, 1994, 1995*a*,*b*), we obtained Laue data sets of restrictocin (Yang & Moffat, 1996) with a completeness greater than 95%, after accurate integration of both streaky and spatially overlapping spots, wavelength normalization and harmonic deconvolution. Other examples of Laue data reduced by *LaueView* are presented in the time-resolved studies on photoactive yellow protein, PYP (Ren *et al.*, 1996; Genick *et al.*, 1997) and carbonmonoxy-myoglobin (Srajer *et al.*, 1996; Teng *et al.*, 1997). Several PYP data sets are complete to 95% in the resolution range of ∞ to $2d_{\min}$; and myoglobin data sets to 85%. Our purpose in using a Laue data set for structure refinement of restrictocin is twofold. First, we aim to refine the restrictocin crystal structure to the highest possible resolution. Second, we intend to test the accuracy of structure-factor amplitudes reduced from the Laue images and to study the behavior of a complete set of Laue data in terms of conventional structure refinement. We describe detailed strategies for Laue data collection and reduction in order to obtain a complete set of high-resolution Laue data suitable for conventional structure refinement. We illustrate the data quality in terms of the conventional statistics in reciprocal space; and through electron-density maps derived from the Laue data, discuss the use of the Laue diffraction techniques for conventional crystallographic studies.

2. Materials and methods

2.1. *Crystals*

Restrictocin is a highly specific *Aspergillus* ribotoxin of 149 amino-acid residues that cleaves a single phosphodiester bond in 28S eukaryotic ribosomal RNA and inhibits protein synthesis (Lamy *et al.*, 1992). Restrictocin can be crystallized by combining vapor diffusion and microdialysis techniques, identical to those used to crystallize the closely related protein, mitogillin (Martinez & Smith, 1991). Restrictocin crystallizes in the monoclinic $P2_1$ space group with cell constants a = 50.24, b = 82.16, c = 38.04 Å and β = 100.5°, and contains two molecules per asymmetric unit (Yang & Moffat, 1996). Crystals with a typical size of 0.3 × 0.3 × 0.8 mm diffract to 1.6 Å resolution with a rotating anode Cu $K\alpha$ X-ray source and to 1.4 Å resolution with 0.9 Å monochromatic radiation at the X4A beamline of the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory.

2.2. *Data collection and reduction of the monochromatic FAST data*

A monochromatic restrictocin data set was collected on an Enraf–Nonius FAST area detector with a four-circle goniometer mounted on an Elliott GX-21 X-ray generator operated at 40 kV and 70 mA. An oscillation angle of 0.2° was used with an exposure time of 30–110 s per frame. The crystal-to-detector distance was 60 mm and the 2θ angle of the detector was set to −20°. Four sweeps of the ω angle, a total of 2050 frames, were collected over a 2 d period. Indexing, geometry refinement and integration of the monochromatic FAST data set were carried out by the program *MADNES*, followed by the program *PROCOR* for profile fitting and Lorentz/polarization factor corrections. A monochromatic scaling scheme in the program *LaueView* (Ren & Moffat, 1995*a*) was used to scale and merge the data from different ω sweeps. The $R_{\rm merge}$ of the data set is 3.52%, with a completeness of 99.3% in the resolution range 15.0–2.2 Å and 92.5% in the shell at the highest resolution collected between 2.2 and 2.1 Å (see Table 1).

2.3. *High-resolution Laue data collection*

62 Laue images were collected from one restrictocin crystal on the X26C beamline at the NSLS (Getzoff *et al.*, 1993). The white X-ray beam was focused by a Pt-coated cylindrical focusing mirror, which also imposes a wavelength cutoff at around 1 Å, and collimated to 300 μm in diameter. The storage ring was operated at a current between 240 and 181 mA. An aluminum foil of 150 μm in thickness was inserted into the X-ray beam in order to attenuate the X-ray beam at longer wavelengths and to limit X-ray heating of the crystal. A BioCARS Laue bench camera was used to take these Laue photographs. Fuji HR-III$_{\rm N}$ imaging plates 20.1 × 25.2 cm in size were used to record the diffraction patterns and a Fuji BAS2000 scanner digitized the diffraction images at 100 μm raster size using the Fuji software supplied. Data collection was carried out at room temperature.

In order to overcome limitations of the dynamic range of the image-plate detector and scanner, two exposure times were used. For the first 31 Laue images, a relatively brief exposure time of 0.5 ms per frame was used to obtain accurate intensity measurements for the strong diffraction spots, which occur predominantly at low resolution. This is essential to successful harmonic deconvolution in the later data-reduction stages (Ren & Moffat, 1995*b*). With an 8° angular spacing between images, these 31 Laue images (with three duplicated exposures) cover a spindle angular range of 216°. For the last 31 Laue images, a 20 times longer exposure time of 10 ms per frame was used to measure the weaker spots which occur predominantly at high resolution. An 8° spacing between images was again used but offset by 4°

Table 1. *Comparisons of data-collection parameters for the monochromatic and Laue data sets*

	Monochromatic	Laue
X-ray source	In-house rotating-anode X-ray generator (Elliott GX21)	NSLS X26C beamline
Wavelength (Å)	1.5418 (Cu $K\alpha$)	0.7–2.2
Detector	Enraf–Nonius area detector	FUJI imaging-plate system
Detector size (mm)	62.4 × 46.8	256 × 205
Pixel size (μm)	122 × 91	100 × 100
Crystal size (mm)	0.3 × 0.3 × 0.8	0.23 × 0.2 × 0.6
Number of crystals	1	1
Number of frames	2350 for 3 sweeps	62
Frame width	0.2°	—
Sweep range	140° for each sweep	224°
Crystal-to-detector distance (mm)	60	120
2θ max (°)	49	54 for Laue 31-1 59 for Laue 31-2
Exposure time per frame	30 s	0.5 ms/10 ms
Total exposure time	1175 min	325.5 ms (0.5 × 31 + 10 × 31)
Total elapsed time	2 d	2 h

compared to the first 31 images, and 224° was covered by the second 31 frames (with two duplicated exposures). All image plates were offset vertically by 40 mm to increase the maximum 2θ angle at the detector from 54 to 59° and to record the high-resolution and long-wavelength diffraction. For all 62 images, the crystal-to-detector distance was set to 120 mm, a compromise between the necessity to minimize spatial overlaps and to record data at the highest Bragg angles. If the crystal-to-detector distance is too large and Bragg angle acceptance too small, reflections at higher resolution and longer wavelengths will not be recorded. The wavelength normalization algorithm will then be unable to fully recover the X-ray spectrum at longer wavelengths (Ren & Moffat, 1995*a*). The total exposure time for collecting these 62 Laue images on a single restrictocin crystal was about 325.5 ms; and the total elapsed time was about 30 min for the exposures and about 2 h for image scanning. If a fully automated CCD detector system or an on-line image-plate scanner were employed, the elapsed time of this type of experiment would be limited only by the data readout/storage time. The last few Laue images began to reveal radiation damage as evidenced by elongated spots, which could nevertheless still be integrated successfully (see Table 1).

2.4. *Laue data reduction*

The position of the direct beam center in the first Laue image was determined by refining a few selected ellipses corresponding to prominent zones, which must intercept at the beam center. The first Laue image was then indexed from the positions of four prominent nodal spots on the image. Subsequent images were automatically indexed according to their spindle angles relative to the first reference image. If the crystal orientation changed due to, for example, crystal slippage during data collection, the first subsequent image was manually re-indexed. After indexing, the unit-cell parameters (b, c and β), diffraction geometry (direct beam center, crystal-to-detector distance), detector parameters (pixel size and tilt angles) and crystal orientation were refined individually and jointly. This geometry refinement procedure successfully predicted the diffraction pattern, including the spot position and wavelength. An analytic profile-fitting technique was used to integrate Laue spot of various shapes and to deconvolute spatially overlapped spots (Ren & Moffat, 1995*a*). This analytical profile can describe a diffraction spot with a variety of shapes, in which the spot parameters are distinct but continuous between different detector bins across detector space. About 900 sample reflections (non-spatially overlapping spots with high signal-to-noise ratio and well refined spot position) were selected from each image to extract distinct standard profiles for different bins in detector space. For a test trial, spots in a few sample images are often deliberately integrated to a resolution limit higher than that which the crystal actually exhibits. A histogram plot of $I/\sigma(I)$ *versus* resolution for these sample images is then useful to identify the true, highest resolution limit of the data set. This parameter is critical for the accurate prediction of singles (Laue spots containing one reflection) and multiples (Laue spots containing more than one harmonic) in a Laue pattern. The integrated intensities of the Laue spots from different images were then scaled together. In the data scaling, the Lorentz and polarization factors were corrected and isotropic and anisotropic scale factors and temperature factors were applied. A Chebyshev polynomial of 64 terms was used to model the wavelength normalization curve, also known as the λ-curve. Over-predicted Laue spots and erroneously measured integrated intensities were identified and rejected by visual inspection of scatter plots, such as $(I - \langle I\rangle)/\sigma(I)$ *versus* wavelength, or $(I - \langle I\rangle)/\langle I\rangle$ *versus* wavelength. Data rejection and scaling were carried out alternately. Redundant measurements of a reflection (including the reflection

itself, its Friedel mate and symmetry-related reflections) that are associated with a unique primary reflection were used to establish a set of equations for each reflection, whose solution leads to deconvolution of the structure-factor amplitudes for each component reflection of the overlapped multiples (Ren & Moffat, 1995*b*). Single reflections were also employed to deconvolute multiples since they increase the ratio of equations to unknowns. After harmonic deconvolution, singles and multiples were scaled and merged to yield the final set of Laue data.

2.5. *Structural refinement against Laue data at 1.7 Å resolution*

The starting model for structural refinement against the synchrotron Laue data was a partially refined structure of restrictocin derived from refinement against the monochromatic FAST data at 2.0 Å resolution. The Laue refinement was carried out by a series of protocols in the program *X-PLOR*, including conjugate-gradient positional refinement, grouped and individual temperature-factor refinement, alternated with manual rebuilding against $2F_o - F_c$ and $F_o - F_c$ maps. Two Laue data sets were used in the structural refinement: Laue31–1 reduced from the first 31 images and Laue62 reduced from the combined 62 images, 8% of which was set aside and used to calculate the free R factor in order to guard against overfitting (Brünger, 1992).

3. Results

3.1. *Laue data reduction and analysis*

From the 62 Laue images collected on a single restrictocin crystal, three Laue data sets or subsets were obtained: Laue31–1 for the first 31 images, each of exposure time 0.5 ms; Laue31–2 for the last 31 Laue images, each of exposure time 10 ms; and Laue62 for all 62 images combined. The cell parameters and the diffraction geometry were refined to high accuracy with standard deviations of cell length and angles of 0.34 Å and 0.22°, respectively. The predicted centers of the reflections based on the refined geometric parameters deviate from the observed centers of gravity of reflections by an r.m.s. value of about 20 μm. High accuracy of spatial prediction is important to the subsequent profile-fitting process with an analytical profile. The highest resolution for data sets Laue31–1 and Laue31–2, to which the images were actually integrated, is 1.5 and 1.4 Å, respectively.

Only single reflections were used in data scaling and wavelength normalization. A Chebyshev polynomial with 64 terms was used to model the λ-curve in the

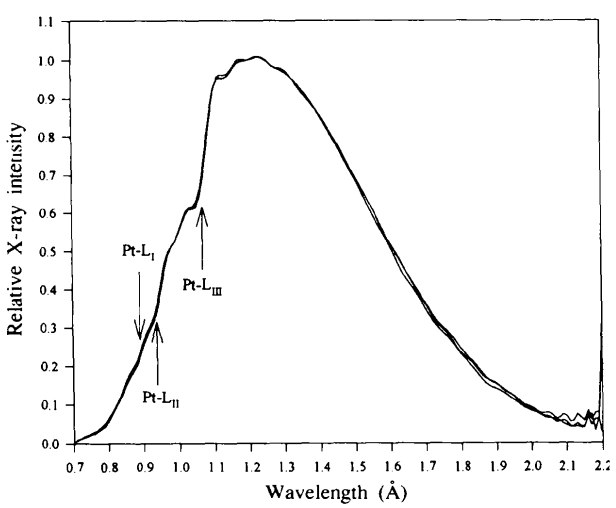
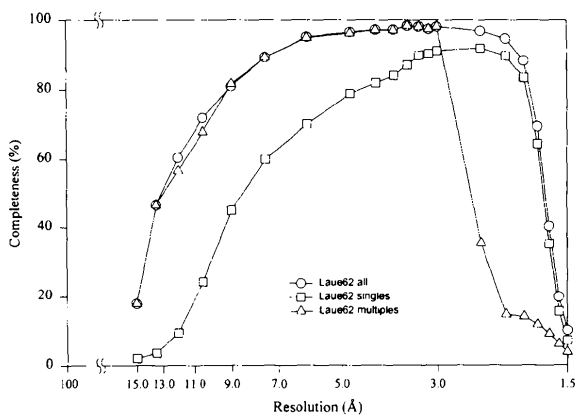
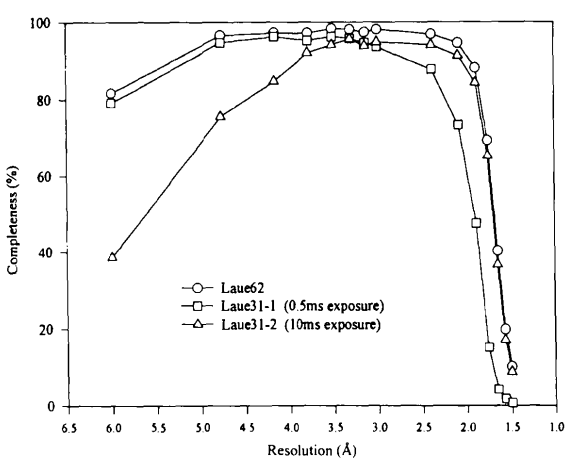


Fig. 1. Three λ-curves of the X26C beamline, NSLS, reduced from the data sets Laue62, Laue31-1 and Laue31-2. The L_I, L_{II} and L_{III} Pt absorption edges are marked.



(*a*)



(*b*)

Fig. 2. Completeness *versus* resolution. Data points are located at the high-resolution side of each bin. (*a*) Singles, deconvoluted multiples and those combined for data set Laue62; (*b*) data sets Laue31-1, Laue31-2 and Laue62.

Table 2. *Data-reduction statistics of Laue data sets, Laue31-1, Laue31-2 and Laue62*

	Laue31-1	Laue31-2	Laue62
Singles			
R_{merge}† on I (%)	9.97	14.04	11.85
Number of measurements	112076	167095	278514
Number of unique reflections	16891	26755	29183
Overall redundancy	6.63	6.24	9.54
Highest resolution (Å)	1.5	1.4	1.4
Wavelength range (Å)	2.1–0.7	2.1–0.7	2.1–0.7
Multiples			
Number of deconvoluted	8569	10550	11691
Not in common with singles	2592	3321	2982
Possible multiples	5977	7229	8709
All			
R factor‡ between singles and multiples (%)	0.9	1.7	1.4
Number of unique reflections	19398	29177	31217

† $R_{merge} = \sum_{hkl}\sum_i |I_i - I|/\sum_{hkl}\sum_i I_i$, where I_i is one measurement of a unique reflection hkl and I is the mean of all the measurements for the reflection hkl. ‡ R factor $= \sum_{hkl}|F_1 - kF_2|/\sum_{hkl}F_1$ where k is a scale factor. The summation is over all common reflections of two data sets.

wavelength range of 2.2–0.7 Å. With an algorithm that employs a resolution-dependent bandpass (Z. Ren *et al.*, manuscript in preparation), the whole curve was treated as one segment, rather than two segments spanning λ_{min} to $2\lambda_{min}$ and $2\lambda_{min}$ to λ_{max} (Ren & Moffat, 1995*a*). Three λ-curves were derived independently, one from each data set; all display the Pt L_I, L_{II} and L_{III} absorption edges derived from the Pt-coated focusing mirror in the X26C beamline (Fig. 1). Despite the large difference in exposure times, these three curves agree with one another closely and indicate that both the beamline and the wavelength normalization procedure are stable.

When the scale factors and λ-curve were applied to the single reflections, a unique set of merged single reflections was obtained. As expected, the completeness at low resolution for such single-only data sets is not satisfactory (Fig. 2*a*): from ∞ to 6.0 Å resolution, the completeness is less than 50%. To retrieve those low-resolution reflections suffering from energy overlap, the harmonic deconvolution protocol was then carried out. The data set of deconvoluted multiples shows a completeness of over 90% in the resolution range 20.0–3.0 Å, which is highly comparable to the completeness of the monochromatic data set (Fig. 2*a*).

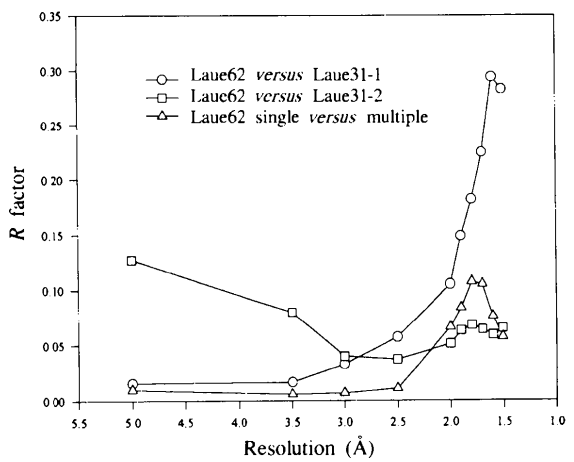


Fig. 3. R factors *versus* resolution between data sets Laue62, Laue31-1 and Laue31-2; and between single reflections and deconvoluted multiples. Data points are located at the high-resolution side of each bin. R factor $= \Sigma_{hkl}|F_1 - kF_2|/\Sigma_{hkl}F_1$, where k is a scale factor.

Fig. 2(*b*) presents the completeness of the three data sets, Laue31–1, Laue31–2 and Laue62. The Laue31–1 and Laue31–2 data sets complement each other at different resolution ranges and result in the more complete joint set of data, Laue62.

To assess data precision, the R factors *versus* resolution between data sets Laue62, Laue31–1 and Laue31–2 were also examined (Fig. 3). The low R factors of less than 7.5% between Laue62 and Laue31–1 in the lower resolution range (15.0–2.5 Å) and between Laue62 and Laue31–2 in the higher resolution range (3.0–1.5 Å) are consistent with our conclusions from the completeness comparison (Fig. 2*b*). For the data set Laue62, most low-resolution reflections come from Laue31–1 and most high-resolution reflections from Laue31–2. The distribution of R factor *versus* resolution between singles and deconvoluted multiples in Laue62 is also shown in Fig. 3. Not surprisingly, the R factors are very low, since those single observations of 'possible multiple' spots play an important role in the energy deconvolution procedure (Ren & Moffat, 1995*b*). Such 'possible multiple' reflections can be collected as singles in some images and as multiples in other images. For example, in data set Laue31–2, 10 550 reflections were deconvoluted, of which 7229 were derived from measurements of both 'possible multiple' singles and multiples, and 3321 were derived only as multiples (Table 2). Structure-factor amplitudes for these 3321 reflections can only be derived

from a successful energy-overlap deconvolution procedure. Data-reduction statistics for the three Laue data sets including singles and multiples are listed in Table 1. These data sets are highly redundant, and the low merging R factors indicate good agreement between redundant measurements.

3.2. *Performance of the Laue data set in structural refinement*

We initiated the high-resolution structure refinement with a restrictocin model that had been refined against the FAST monochromatic data at 2.0 Å resolution. Switching reflection data sets during a structure refinement process often results in an increase of the refinement R factors. This, however, did not occur when we replaced the monochromatic FAST data set with Laue31–1. The refinement was continued with the data set Laue31–1, containing more reflections, but the R factor (21.1% in the resolution range 8–2.07 Å) showed no increase, confirming that the quality of the Laue data was at least as good as this monochromatic data.

We then extended the resolution limit to 1.7 Å resolution, the highest resolution for which the completeness in the final shell between 1.8 and 1.7 Å is acceptable (30% for Laue31–1 and 57% for Laue62) and reliable Laue structure amplitudes were obtained. After several cycles of positional refinement against Laue31–1 alternating with grouped B-factor refinement, the R factor dropped from 21.9 to 19.5%, followed by individual B-factor refinement, with which an R factor of 18.0% was obtained. To include more reflections, we then used the more complete Laue data set Laue62. With 3434 more reflections (12.1% of the data set), an R factor of 17.7% and a free R factor of 23.7% were obtained. The final refined restrictocin model contains 285 residues, three phosphate groups and 204 waters. The r.m.s. deviations from ideal protein stereochemistry for bond length, bond angle and dihedral angle are 0.018 Å, 2.02 and 27.1°, respectively [see Table 6 of Yang & Moffat (1996) for refinement statistics]. Fig. 4 shows the real-space correlation coefficient (Jones *et al.*, 1991) for each residue in the two molecules of restrictocin in one asymmetric unit.

To further examine the accuracy of our Laue structure-factor amplitudes, a Luzzati plot of refinement R factor *versus* resolution is shown in Fig. 5. Unlike more common Luzzati plots in which the R factor for the bin at lowest resolution is much higher than that for the second bin, the refinement R factor in the lowest resolution bin (8.0–3.5 Å) is the lowest, 12.9%, and provides further evidence for the accuracy of the harmonic deconvolution algorithm of the *LaueView* (Ren & Moffat, 1995*b*). In the higher resolution range, the Laue data set behaves more like a conventional monochromatic data set, in which a progressive increase in R factor is observed, because of both the weaker diffraction and the lower redundancy.

Electron-density maps are perhaps the most informative way to illustrate the agreement between the observed structure-factor amplitudes (F_o) and the calculated amplitudes (F_c) from a refined model. The $2F_o - F_c$ map and $F_o - F_c$ omit map show no discontinuity in either the backbone or side chains, in contrast with the maps obtained by Bourgeois *et al.* (1997). Fig. 6 shows some typical regions of the omit $F_o - F_c$ map of restrictocin at 1.7 Å resolution, where F_o's are derived from the data set Laue62, F_c's and phases are calculated from the refined model (except for the marked residues in the map). Most carbonyl O atoms show very well defined density (Fig. 6*a*); density for side chains such as Arg and Glu is unambiguous (Fig. 6*b*). A segment of α-helix clearly shows the main-chain and side-chain atoms (Fig.

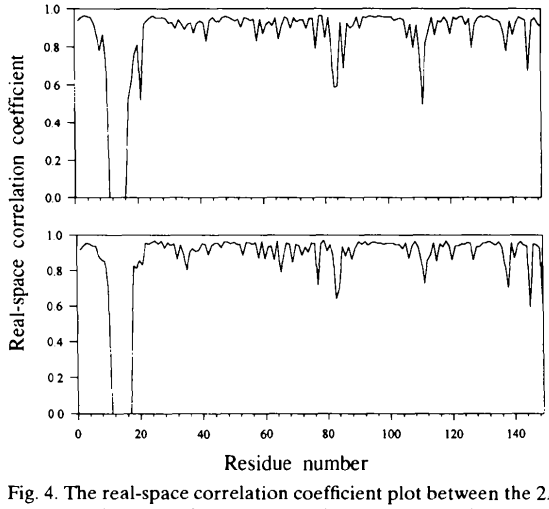


Fig. 4. The real-space correlation coefficient plot between the $2F_o - F_c$ map and the calculated electron density *versus* residue number for the two independently built monomers in one asymmetric unit.

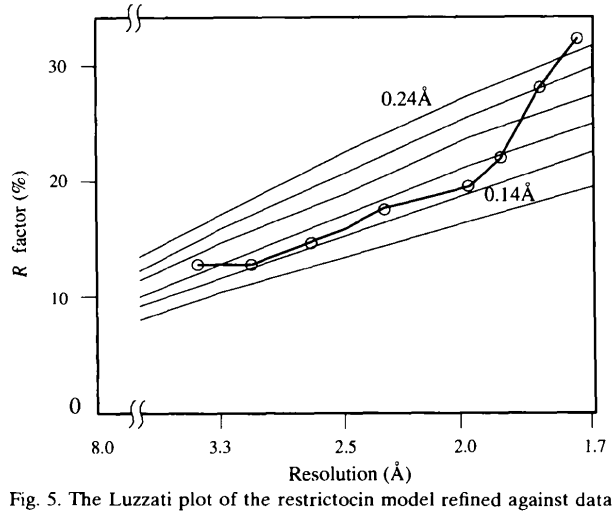


Fig. 5. The Luzzati plot of the restrictocin model refined against data set Laue62. Data points are located at the high-resolution side of each bin.

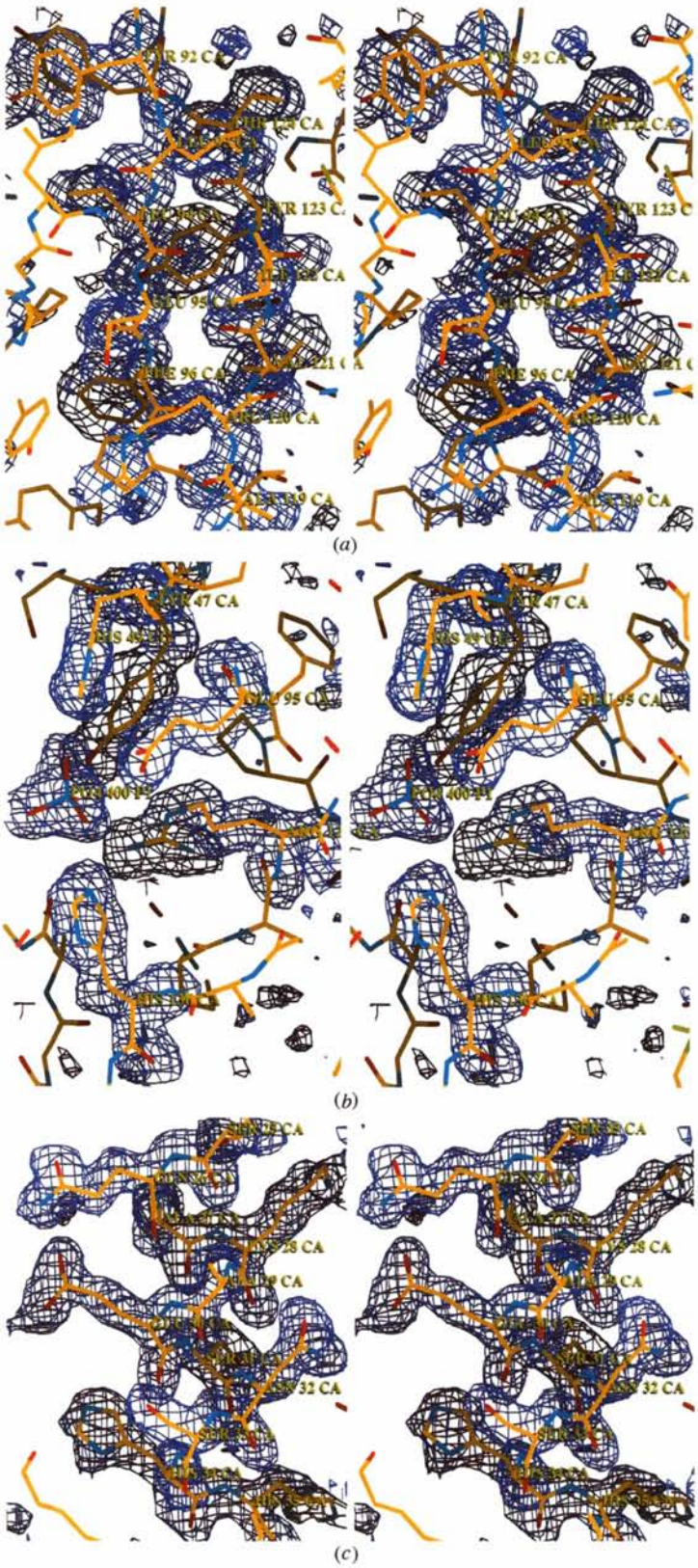


Fig. 6. Stereoview of omit $F_o - F_c$ maps for typical regions in restrictocin structure. (*a*) Two adjacent antiparallel β-strands; (*b*) the residues at the catalytic center, Tyr47 His49, Glu95, Arg120 and His136, where the tetrahedral electron density represents an inorganic phosphate group from the crystallization buffer; (*c*) a segment of a three-turn α-helix.

6c). The map quality is comparable to, and arguably better than, that of the best maps that can be obtained by conventional monochromatic techniques at this resolution. This observation forms the basis of our claim for the comparability of our present Laue data and excellent monochromatic data. R factors depend on the data redundancy and on the λ-curve, and have recently been shown to be unreliable measures of data quality (Diederichs & Karplus, 1997; Weiss & Hilgenfeld, 1997; Z. Ren, unpublished work).

3.3. *Comparison of monochromatic FAST and Laue data sets*

We also compared the structure-factor amplitudes derived from the monochromatic FAST data and those from the Laue data. Since the monochromatic data set was only collected to 2.0 Å resolution, our comparison was restricted to the resolution range 10.0–2.0 Å. The correlation coefficient between the structure-factor amplitudes of the Laue and monochromatic data sets is 0.987. This good agreement is also indicated by the overall R factor of 11% between the data sets (defined in Fig. 7 legend), and by the plot of R factors *versus* resolution (Fig. 7). The discrepancy between the Laue and monochromatic structure-factor amplitudes becomes bigger when the diffraction data get weaker at higher resolution. Consistent with this observation, the correlation plot shows larger scatter for weak reflections.

We have carried out further refinements, starting from the same model, against both the monochromatic and Laue data sets over the resolution range 10.0–2.0 Å using the same refinement protocol. The refinement statistics are shown in Table 3. Despite the fact that the Laue data set incorporates more reflections in the refinement, it shows much lower values of the working R factor and the free R factor than those for the monochromatic data set. In terms of stereochemistry, the model refined against the Laue data is also slightly better than that from the monochromatic data. This suggests that the discrepancy between these two data sets is largely contributed by errors in this monochromatic data set rather than in the Laue data set. Fig. 8 shows the $2F_o - F_c$ maps of residue Trp50 of restrictocin in four different contour levels, calculated with both the Laue data set and with the FAST data set. At a contour level of 1σ, both the Laue and monochromatic maps show the shape of the side chain, but with higher contour levels, the continuity of the density derived from the FAST data starts to break down. The density derived from the Laue data is better defined, and at a contour level of 2σ, a hole is observed in the density for the Trp side chain.

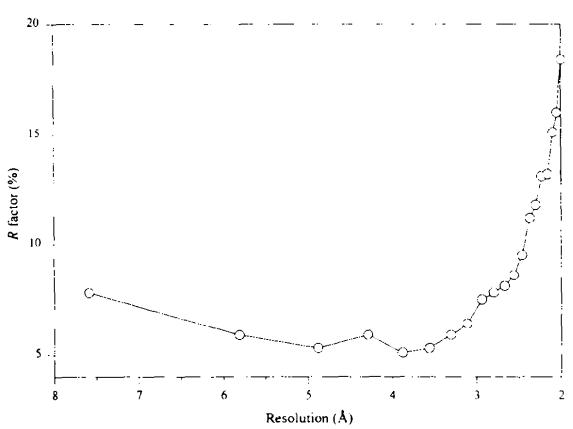


Fig. 7. R factor *versus* resolution shows the agreement between the structure-factor amplitudes derived from the monochromatic FAST data and Laue data in the resolution range 10.0–2.0 Å. Data points are located at the high-resolution side of each bin. R factor = $\Sigma_{hkl}|F_{\text{Laue}} - kF_{\text{mono}}|/\Sigma_{hkl}F_{\text{mono}}$, where k is a scale factor.

Table 3. *Comparisons of refinement statistics of the Laue and FAST data sets in the resolution range 10.0–2.0 Å*

	Laue (NSLS X26C)	Monochromatic (FAST)
Number of reflections	18566	17936
R factor (%)	15.65	17.44
Free R factor (%)	21.62	23.77
R.m.s. deviations		
Bond length (Å)	0.020	0.021
Bond angle (°)	2.09	2.14
Dihedral angle (°)	27.24	27.33

4. Discussion

4.1. *Laue data-collection strategies*

Careful planning of the Laue data-collection strategy is very important to achieve high accuracy and completeness. Accurate measurement of Laue spots at low Bragg angle will directly affect the energy deconvolution process, while Laue spots at high Bragg angle will contribute substantially to the extraction of the λ-curve at longer wavelengths. These spots are also critical to the energy overlap deconvolution process (Ren & Moffat, 1995*b*). To take account of both effects, we adopted a strategy of two exposure times and two different maximum Bragg angles obtained by shifting the image plates. The 40 mm image-plate shift was determined by the highest diffracting resolution of the restrictocin crystal, the crystal-to-detector distance and the known wavelength range of the beamline. The exposure time also has to be carefully chosen to take full advantage of the dynamic range of the detector, yet to avoid crystal heating artifacts due to X-ray absorption in the intense synchrotron white beam. The adiabatic heating rate of a crystal resulting from X-ray absorption at the X26C beamline at the NSLS is estimated to be as high as 200 K s^{-1} (Chen, 1994). If exposures longer than 30 ms are required to obtain well

exposed patterns, they must be delivered in sub-exposures separated by a cooling period.

Since a typical single Laue image contains many more spots than a conventional monochromatic image, the Laue method is known as an efficient data collection technique and may yield a data set with fewer images (Clifton *et al.*, 1991). But as in the monochromatic method, redundant measurements are very important to overall accuracy, and in the Laue case also to the success of wavelength normalization and energy-overlap deconvolution. Since the number of high-quality images that can be obtained before unacceptable radiation damage occurs is limited and unpredictable for a new experiment, we adopted a strategy to address the effect in restrictocin data collection. We first swept a large spindle angular range of 216° with a large angular interval of 8° to ensure the completeness of the data set, and then inserted further images to achieve higher redundancy if, as was indeed the case, the crystal still diffracts reasonably well. In this way, we maximized our chance to obtain a data set that was both complete and redundant even at the highest resolution.

4.2. *Common errors in Laue data collection and reduction*

Several common errors may occur during Laue data collection and reduction. First, Laue experiments may have been conducted with less attention to experimental design: very sparse spindle angles may be selected in order to complete a data set; too few images may have been taken, resulting in a less complete data set than expected; and sometimes, the Laue technique fails to achieve high completeness when as few frames as possible are taken instead of as many as possible. We suggest that an angular spacing of 8° or less is necessary for a λ-curve of the form used here for crystals that diffract to well beyond 2 Å resolution (see Fig. 6 and Table 1 of Ren & Moffat, 1995*b*). A finer angular spacing increases both data completeness and redundancy.

A probabilistic approach to energy-overlap deconvolution has been described (Bourenkov *et al.*, 1996), which should theoretically result in more accurate deconvoluted components. Since this approach does not require redundant measurements, it may be applied to more reflections, and, therefore, further increase the completeness at low resolution. However, the characteristic 'low-resolution-hole' of Laue data results from under-sampling at low resolution more than from energy overlaps (Cruickshank *et al.*, 1987). No numerical approach can retrieve data that are not measured in any way, unless further prior knowledge (such as structural information) is supplied. A finer angular spacing is essential for more complete low-resolution Laue data [compare Fig. 2(*a*) with Fig. 5 of Bourenkov *et al.* (1996), and the omit maps at 1.7 Å resolution in Fig. 6 with the density maps calculated from model phases at 1.4 Å resolution in Fig. 8 of Bourenkov *et al.* (1996)].

Second, the maximum Bragg angle acceptance is often too limited in a Laue experiment, either because of limited detector size or because of concerns of spatial overlaps. Spots that diffract at high Bragg angles are critical to recovery of the λ-curve at long wavelengths. Inaccuracy in part of the λ-curve will directly affect data accuracy at those wavelengths, and further affect

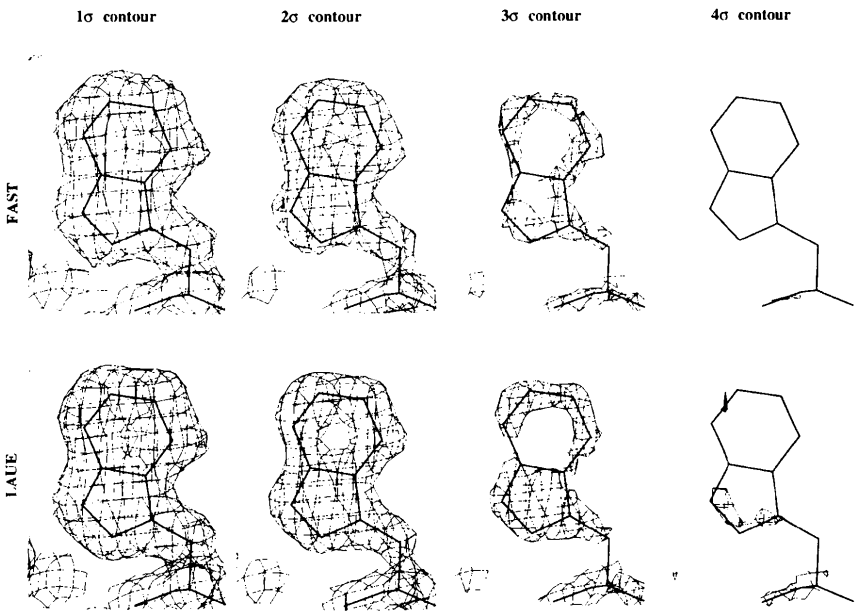


Fig. 8. Comparisons between the $2F_o - F_c$ maps of residue Trp50 contoured at 1σ, 2σ, 3σ and 4σ levels. The maps in the upper panel are derived from the monochromatic FAST data and those in the lower are from Laue data. The maps are calculated using reflections in the resolution range 10.0–2.0 Å.

the quality of the harmonic deconvolution at all wavelengths. Third, the soft limits d_{min}, λ_{min}, and λ_{max} are often set to values that underpredict the true diffraction limits during Laue data processing. This error may arise if the data-reduction software cannot handle spatial overlaps properly, since underprediction will apparently relieve the symptoms of spatial overlap by ignoring those spots that either diffract weakly at high resolution or lie in the two wings of the X-ray spectrum, so that the majority of spots can be integrated. This approach is actually a tradeoff when spatial overlaps cannot be modeled correctly. Underprediction of soft limits however poses problems when high accuracy and high completeness of data set are desired. Intensities of unpredicted spots may be combined with those of adjacent, predicted spots; spot profiles may be skewed by the presence of adjacent, unpredicted spots; and fewer reflections have to share the total intensity measured in an energy-overlapped Laue spot. We prefer to deliberately overpredict the diffraction limit and the wavelength range, and then to choose the appropriate soft limits based on examination of statistical plots (Ren & Moffat, 1995*a*).

5. Concluding remarks

Table 1 lists the data-collection parameters for monochromatic and Laue data sets. The most significant difference between them is the exposure time: the total exposure time for the Laue data set is 2.4×10^5 shorter than for the FAST monochromatic data set. Two factors account for this huge difference. First, the synchrotron X-ray beam is much more brilliant than that from our in-house rotating anode X-ray generator. Second, the Laue diffraction geometry takes full use of the polychromatic X-ray beam in a wide spectrum while the monochromatic method only picks up X-rays at one wavelength (Moffat, 1997).

In this paper, we have demonstrated that with a series of new algorithms (Ren & Moffat, 1995*a*,*b*), a complete set of Laue data can be obtained with an accuracy highly comparable to that of an excellent monochromatic data set; and the Laue data can be successfully applied to conventional crystallographic studies, such as high-resolution structure refinement. In the present case, the Laue data set is superior to the monochromatic set over the same resolution range. Of course, this may reflect random and systematic errors from all kinds of sources in the process of data collection and reduction, such as the high dark current background and small active area of the TV-based FAST area detector, poor scaling algorithms, *etc.* Structure refinement of restrictocin based on this Laue data set further demonstrates its quality.†

We thank Dr William Kenealy, originally of the University of Wisconsin at Madison and now at J. Whittier Biologics Inc. of Madison, Wisconsin, who purified the restrictocin. We thank Drs Sergio Martinez and Janet Smith of Purdue University for providing the restrictocin and helpful advice on its crystallization. This work is supported by grants from NIH to KM.

† Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1AGZ and R1AGZSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: HA0161).

References

Bartunik, H. D. & Borchert, T. (1989). *Acta Cryst.* A**45**, 718–726.

Bourenkov, G. P., Popov, A. N. & Bartunik, H. D. (1996). *Acta Cryst.* A**52**, 797–811.

Bourgeois, D., Longhi, S., Wulff, M. & Cambillau, C. (1997). *J. Appl. Cryst.* **30**, 153–163.

Bradbrook, G., Deacon, A., Habash, J., Helliwell, J. R., Helliwell, M., Nieh, Y. P., Snell, E. H., Trapani, S., Thompson, A. W., Campbell, J. W., Allinson, N. M., Moon, K., Ursby, T. & Wulff, M. (1995). *SPIE J.* **2521**, 161–177.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Campbell, J. W. (1993). CCLRC Daresbury Laboratory, Warrington, England.

Campbell, J. W. & Hao, Q. (1993). *Acta Cryst.* A**49**, 889–893.

Cassetta, A., Deacon, A., Emmerich, C., Habash, J., Helliwell, J. R., McSweeney, S., Snell, E., Thompson, A. W. & Weisgerber, S. (1993). *Proc. R. Soc. London Ser. A*, **442**, 177–192.

Chen, Y. (1994). PhD thesis, pp. 106–187, Cornell University, Ithaca, New York, USA.

Clifton, I. J., Elder, M. & Hajdu, J. (1991). *J. Appl. Cryst.* **24**, 267–277.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cruickshank, D. W. J., Helliwell, J. R. & Johnson, L. N. (1992). Editors. *Time-resolved Macromolecular Crystallography*. London: The Royal Society/Oxford Science Publications.

Cruickshank, D. W. J., Helliwell, J. R. & Moffat, K. (1987). *Acta Cryst.* A**43**, 656–674.

Cruickshank, D. W. J., Helliwell, J. R. & Moffat, K. (1991). *Acta Cryst.* A**47**, 352–373.

Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.

Duke, E., Wakatsuki, S., Hadfield, A. & Johnson, L. N. (1994). *Protein Sci.* **4**, 1178–1196.

Genick, U. K., Borgstahl, E. O., Ng, K., Ren, Z., Pradervand, C., Burke, P. M., Srajer, V., Teng, T.-Y., Schildkamp, W., McRee, D., Moffat, K. & Getzoff, E. D. (1997). *Science*, **275**, 1471–1475.

Getzoff, E. D., Jones, K. W., McRee, D., Moffat, K., Ng, K., Rivers, M. L., Schildkamp, W., Singer, P. T., Spanne, P., Sweet, R. M., Teng, T.-Y. & Westbrook, E. M. (1993). *Nucl. Instrum. Methods Phys. Res. B*, **79**, 249–255.

Hajdu, J. & Andersson, I. (1993). *Ann. Rev. Biophys. Biomol. Struct.* **22**, 467–498.

Hajdu, J. & Johnson, L. (1990). *Biochemistry*, **29**, 1669–1678.

Hao, Q., Campbell, J. W., Harding, M. M. & Helliwell, J. R. (1993). *Acta Cryst.* A**49**, 528–531.
Helliwell, J. R. (1992). *Philos. Trans. R. Soc. London Ser. A*, **340**, 221–232.
Helliwell, J. R., Habash, J., Cruickshank, D. W. J., Harding, M. M., Greenhough, T. J., Campbell, J. W., Clifton, I. J., Elder, M., Machin, P. A., Papiz, M. Z. & Zurek, S. (1989). *J. Appl. Cryst.* **22**, 483–497.
Jones, T. A., Zhou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Lamy, B., Davies, J. & Schindler, D. (1992). In *Genetically Engineered Toxins*, edited by A. E. Frankel, pp. 237–258. New York: Marcel Dekker.
McRee, D. E. (1992). *J. Mol. Graphics*, **10**, 44–46.
Martinez, S. & Smith, J. (1991). *J. Mol. Biol.* **218**, 489–492.
Moffat, K. (1998). *Methods Enzymol.* **277**, 433–447.
Moffat, K. & Helliwell, J. R. (1989). *Topics Curr. Chem.* **151**, 61–74.
Ren, Z. & Moffat, K. (1994). *J. Synchrotron Rad.* **1**, 78–92.
Ren, Z. & Moffat, K. (1995*a*). *J. Appl. Cryst.* **28**, 461–481.
Ren, Z. & Moffat, K. (1995*b*). *J. Appl. Cryst.* **28**, 482–493.
Ren, Z., Ng, K., Borgstahl, G. E. O., Getzoff, E. D. & Moffat, K. (1996). *J. Appl. Cryst.* **29**, 246–260.
Schlichting, A., Almo, S., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E., Petsko, G. & Goody, R. (1990). *Science*, **345**, 309–315.
Shrive, A. K., Clifton, I. J., Hajdu, J. & Greenhough, T. J. (1990). *J. Appl. Cryst.* **23**, 169–174.
Singer, P., Smälas, A., Carty, R., Mangel, W. & Sweet, R. (1993). *Science*, **259**, 669–673.
Srajer, V., Teng, T.-Y., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M. & Moffat, K. (1996). *Science*, **274**, 1726–1729.
Stoddard, B., Koenigs, P., Porter, N., Petratos, K., Petsko, G. & Ringe, D. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 5503–5507.
Szebenyi, D. M. E., Bilderback, D., LeGrand, A., Moffat, K., Schildkamp, W. & Teng, T.-Y. (1988). *Trans. Am. Crystallogr. Assoc.* **24**, 167–172.
Teng, T.-Y., Srajer, V. & Moffat, K. (1997). *Biochemistry*, **36**, 12087–12100.
Wakatsuki, S. (1993). In *Data Collection and Processing*, edited by L. Sawyer, N. W. Isaacs & S. Bailey. Warrington: Daresbury Laboratory.
Weiss, M. S. & Hilgenfeld, R. (1997). *J. Appl. Cryst.* **30**, 203–205.
Vellieux, F., Hajdu, J., Verlinde, L., Groenduk, H., Read, R., Greenhough, T., Campbell, J., Kalk, K., Littlechild, J., Watson, H. & Hol, W. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 2335–2359.
Yang, X. & Moffat, K. (1996). *Structure*, **4**, 837–852.