

Ab initio structure determination of low-molecular-weight compounds using synchrotron radiation Laue diffraction

Raimond B. G. Ravelli,^{a*}† Mia L. Raves,^b Sjors H. W. Scheres,^a Arie Schouten^a and Jan Kroon^a

^aDepartment of Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands, and ^bDepartment of Structural Biology, Weizmann Institute of Science, 76100 Rehovot, Israel.
E-mail: ravelli@embl-grenoble.fr

(Received 20 March 1998; accepted 7 September 1998)

The potential of using polychromatic synchrotron radiation for *ab initio* structure determination of low-molecular-weight compounds is investigated. Three different organic compounds were studied. For each of the structures the cell volume was determined with the aid of a Zr attenuator that was placed in the direct beam to obtain a sharp λ_{\min} edge. Space-group determination did not prove to be complicated; it was aided by deconvolution of the multiple spots. Two low-temperature data sets were collected, and it appeared that spot streaking owing to the inherent increase of mosaicity due to freezing was not too detrimental to the application of the Laue method. Although both the R_{merge} values (around 12%) and the final R_1 values (around 10%) are higher than usually found in monochromatic experiments, all structures show reasonable geometry. Comparison of one of the structures with a reference structure obtained with monochromatic data shows an r.m.s. deviation between the 21 non-H-atom positions of 0.019 Å.

Keywords: Laue diffraction; *ab initio* structure determination; white radiation.

1. Introduction

The Laue diffraction method makes use of the full spectrum of X-rays produced by a synchrotron as opposed to the monochromatic technique. This enables the collection of a large amount of intensity measurements in very short exposure times. Two important applications that exploit these benefits have proved to be successful: structure determination on very small crystals (see, for example, Snell *et al.*, 1995; Kariuki & Harding, 1995), and time-resolved studies on proteins (Schlichting *et al.*, 1990; Bolduc *et al.*, 1995; Šrajer *et al.*, 1996; Genick *et al.*, 1997).

However, these advantages are nowadays challenged by different applications of monochromatic radiation. The extreme brilliance of third-generation synchrotrons permits a reduction in the size of the beam to the micrometre scale, and monochromatic structure determinations on extremely small protein crystals (a few micrometres thick) have been reported (Xiao *et al.*, 1995). Time-resolved studies on (enzymatic) reactions that occur on a time scale of minutes can also be followed using fast monochromatic techniques like the large-oscillation-angle Weissenberg method (Gouet *et al.*, 1996; Hajdu & Andersson, 1993). Intermediates that are formed during these reactions can be trapped either chemically (Verschuere *et al.*, 1993) or by

freezing techniques, which eliminates the necessity for rapid data collection. The advantage of monochromatic techniques is that they can circumvent problems associated with the Laue diffraction experiment, like sensitivity to mosaicity of the crystals, harmonic and spatial overlap, low-resolution hole, radiation damage, unfavourable signal-to-background ratio and labour-intensive data processing. Nevertheless, the Laue technique remains the only technique fast enough to collect structural information on a millisecond and even nanosecond time scale in order to follow reactions (Šrajer *et al.*, 1996; Genick *et al.*, 1997).

Here we present the use of the Laue technique in a field in which the disadvantages are less problematic: *ab initio* structure determination of low-molecular-weight compounds. The mosaicity of crystals of these compounds is more likely to be acceptable as far as spatial overlap is concerned. Even if it is high (as expressed in a radial smearing of the spots), the relatively small number of spots on each frame limits the risk of overlap. As with macromolecules, both the low-resolution hole and the harmonic overlap problem can, at least partly, be overcome by collecting a large number of frames on a crystal and using a reliable deconvolution algorithm. Radiation damage can be prevented by cooling the crystal at liquid-nitrogen temperature. The gain in intensity by using a synchrotron beam compared with an in-house rotating anode can cancel out the disadvantage of the unfavourable signal-to-back-

† Present address: European Molecular Biology Laboratory (EMBL), Grenoble Outstation, c/o ILL, BP 156, 38042 Grenoble CEDEX 9, France.

ground ratio inherent to the Laue technique. Finally, during the last few years, important progress has been made in both hardware (CCD and IP detectors) and software to simplify the data collection and processing.

Three features presented in this paper are relatively new and thus far not used in combination with each other in the Laue field: freezing crystals, deconvolution of multiples, and absolute scaling of the unit cell with the use of an attenuator. Some important steps in the data reduction are discussed in the next section. The sensitivity of direct methods to solve structures with the kind of systematic incompleteness that often occurs with Laue data is quantified using the $C(z)$ formalism of Giacovazzo *et al.* (1994). Although the quality of the data presented in this paper is still not ideal, it is shown that the Laue technique can be a fast and efficient alternative for traditional monochromatic structure determination.

The relevance of carrying out *ab initio* structure determinations of low-molecular-weight compounds using Laue data not only lies in the field of tiny crystal studies (Kariuki & Harding, 1995) or possibly high-throughput structure determinations. These well defined systems can also be considered as appropriate pilot models to improve on data collection and processing. This is expected to be beneficial to the applications of the Laue method that make full use of the intense white beam produced by a synchrotron: time-resolved (perturbation) studies, both on low-molecular-weight compounds (Cheetham *et al.*, 1995; H. Graafsma, personal communication) and proteins (Šrajer *et al.*, 1996; Genick *et al.*, 1997).

2. Selected theoretical aspects

A good overview of Laue data collection and reduction methods can be found in Helliwell, Habash *et al.* (1989), Ren & Moffat (1995a), Yang *et al.* (1997) and Ravelli (1998). Here, only a small number of steps are discussed that proved to be the most important for a reliable *ab initio* structure determination of the compounds investigated.

2.1. Profile fitting

Laue diffraction patterns are very sensitive to the mosaic spread of the crystal (see, for example, Bartunik & Borchert, 1989; Hajdu *et al.*, 1991). A high mosaicity will cause the diffraction spots to be smeared and, eventually, to disappear into the background. Freezing the crystal is likely to increase the mosaicity: it is therefore important to be able to account for streakiness during integration. An analytical profile-fitting method is able to fit even severely streaked spots (Ren & Moffat, 1995a). However, in the unfavourable case that freezing has slightly cracked the crystal, causing some spots to be doubled, it might be better to use an empirical profile-fitting algorithm (Rossmann, 1979) like that recently implemented in the program *LAUEGEN* (Campbell, 1995).

2.2. Determination of the unit-cell parameters on an absolute scale

The determination of the unit-cell and crystal-orientation parameters is crucial for the indexing of a single-crystal X-ray diffraction pattern. These parameters can be used to construct two matrices; $[A]$ for a standard setting of the (reciprocal) unit cell, and $[C]$ to rotate the unit cell to its actual setting. The two matrices relate the indices of each reflection $\mathbf{h}(h, k, l)$ to the reciprocal-lattice vector $\mathbf{x}(x, y, z)$ by (Helliwell, Habash *et al.*, 1989; Ravelli *et al.*, 1996)

$$\mathbf{x} = [C][A]\mathbf{h}. \quad (1)$$

The spot positions (X_f, Y_f) for a reflection \mathbf{h} in a Laue diffraction pattern can now be calculated, assuming a flat detector perpendicular to the outgoing beam at a distance d_f from the crystal, and a beam that hits the crystal coming from the negative Z axis, by

$$\begin{aligned} X_f &= d_f[xz/(z^2 - x^2 - y^2)], \\ Y_f &= d_f[yz/(z^2 - x^2 - y^2)]. \end{aligned} \quad (2)$$

These equations clearly show the scaling problem in the Laue diffraction experiment: the matrix $[A]$ can be multiplied by an arbitrary constant s without affecting the predicted spot positions. There is only one restriction on \mathbf{x} : it should lie in an accessible part of reciprocal space, bounded by the λ_{\min} , λ_{\max} and d_{\min} spheres. However, by scaling these spheres by s , these boundary restrictions will remain exactly identical.

Fortunately, an approximate value of λ_{\min} is usually known for each beamline set-up, and can be used to determine the value of s within roughly 20% (Kariuki & Harding, 1995), leading to an initial estimate of the cell volume. Furthermore, in most Laue diffraction experiments the cell parameters and volume are already known from monochromatic experiments, and can be used directly to solve the scaling problem. For *ab initio* structure determination using only Laue data, however, a start-off point is needed to determine the cell volume more accurately. It has been suggested (Carr *et al.*, 1992; Kariuki & Harding, 1995) and demonstrated for a known structure (Carr *et al.*, 1993) that an attenuator with an absorption edge (just) above the expected minimum wavelength of the X-ray beam can be used to obtain a sharp wavelength cut-off. The Pt absorption edges from a Pt-coated focusing mirror (Ren & Moffat, 1995a) or the $K\alpha$ or $K\beta$ radiation produced by a (rotating) anode in a laboratory Laue experiment (Ravelli *et al.*, 1996), however, can also be used. Here, we report on the use of a thin foil of Zr (thickness 25 μm) placed in the direct beam.

2.3. Deconvolution of multiples

Several deconvolution algorithms have been described in the literature (Hao *et al.*, 1993, 1995; Ren & Moffat, 1995b; Bourenkov *et al.*, 1996; Xie & Hao, 1997). Deconvolution of multiples is not only very important for the completeness, especially in the low-resolution shells, but

Table 1

Details of the data collection.

Compound names are given in the text. The same crystal was used for the two data sets DMANM and DMANMInt.

Compound code	DMANM	DMANMInt	tBDD	1DA
Structural formula	C ₄₀ H ₄₆ N ₄ O ₁₄	C ₄₀ H ₄₆ N ₄ O ₁₄	C ₁₀ H ₁₄ O ₂	C ₁₈ H ₁₉ N ₃ O ₂
Crystal size (mm)	0.6 × 0.5 × 0.1	0.6 × 0.5 × 0.1	0.8 × 0.5 × 0.5	0.3 × 0.3 × 0.4
Temperature (K)	295	105	105	295
Synchrotron mode	25 bunch	25 bunch	25 bunch	Single bunch
Current (mA)	145	145	170	80
Beam	Focused	Focused	Focused	Unfocused
Distance (mm)	52	52	52	37
Exposure time (s)	4.0	8.0	2.0	0.5
Number of frames	11	11	10	7
Spindle angles (°)	0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100	0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100	5, 10, 15, 20, 25, 30, 35, 40, 45, 50	0, 15, 22.5, 30.0, 37.5, 45, 52.5

also for the assignment of screw axes and glide planes by identification of systematic absences.

The deconvolution algorithm that is incorporated in *LAUEVIEW* (Ren & Moffat, 1995*b*) solves the equation

$$I_j(\mathbf{h}_0) = \sum_{n=n_{\min}}^{n_{\max}} F^2(n\mathbf{h}_0)/f_{\text{general}_j}(n\mathbf{h}_0), \quad (3)$$

where $I_j(\mathbf{h}_0)$ is the total intensity of the multiple spot, $f_{\text{general}_j}(n\mathbf{h}_0)$ is the general scale factor of the n th harmonic $n\mathbf{h}_0$ for the multiple spot j , and $F(n\mathbf{h}_0)$ represents the structure-factor amplitude. Equation (3) can be solved for $F^2(n\mathbf{h}_0)$ by traditional least-squares methods if the number of measurements of $I_j(\mathbf{h}_0)$ (including singles and symmetry-related measurements) is equal to or greater than the number of unknown values of $F(n\mathbf{h}_0)$. This can be accomplished by increasing the number of frames collected on one crystal.

2.4. Direct methods

It has been discussed (Harding, 1988) that direct methods might lose their power to solve low-molecular-weight structures in the case of the Laue diffraction experiment. Not only does the overall completeness tend to be lower compared with a monochromatic experiment, but especially low-resolution reflections, important in forming triplets, suffer from either the low-resolution hole or the harmonic overlap problem (not all multiples can be deconvoluted; see §2.3).

To quantify the sensitivity of direct methods to the incompleteness of an individual Laue data set, the theoretical approach of Giacovazzo *et al.* (1994) can be used. It utilizes the reliability parameter $\alpha_{\mathbf{h}}$ of the Cochran distribution (Cochran, 1955). In the absence of any information on the phases, the expectation value of α , $\langle\alpha\rangle$, and its variance $\sigma_{\alpha_{\mathbf{h}}}^2$ can be calculated (Casarano *et al.*, 1984) by

$$\begin{aligned} \langle\alpha_{\mathbf{h}}\rangle &= \sum_{j=1}^r G_j D_1(G_j), \\ \sigma_{\alpha_{\mathbf{h}}}^2 &\simeq (1/2) \sum_{j=1}^r G_j^2 [1 + D_2(G_j) - 2D_1^2(G_j)], \end{aligned} \quad (4)$$

where $G_j = 2|E_{\mathbf{h}}E_{\mathbf{k}}E_{-\mathbf{h}-\mathbf{k}}|/N^{1/2}$, $E_{\mathbf{h}}$ is a normalized structure factor, N is the number of atoms in the unit cell, r is the number of triplets selected for reflection \mathbf{h} , and $D_j = I_j/I_0$ with I_j being a modified Bessel function of order j . It has been shown by Giacovazzo *et al.* (1994), that the parameter $z_{\mathbf{h}} = \langle\alpha_{\mathbf{h}}\rangle/\sigma_{\alpha_{\mathbf{h}}}$ can be considered as a signal-to-noise ratio. Diffraction data for which z values are greater than 2 or 3 for most strong reflections constitute a data set which is suitable for the successful application of direct methods.

Although the individual values of G are expected to be large for low-molecular-weight molecules in general, r might become critically low in the case of Laue structure determination, resulting in smaller values for z . By calculating the cumulative distribution function $C(z)$ for each Laue data set, one can check if problems are to be expected when trying to solve the structure.

3. Data collection

Four data sets were collected on three different compounds: bis[1,8(dimethylamino)naphthalene]mellitate dihydrate (DMANM), *trans*-bicyclo[4.4.0]decane-3,8-dione (tBDD) and 5-(*N,N*-dimethylamino)-3,3-dimethyl-2-(4-nitrophenyl)-3*H*-indole (1DA). Two of the data sets were collected at a temperature of 105 K. The compound codes and structural formulae are listed in Table 1. All Laue data were collected at beamline X26C at the National Synchrotron Light Source (NSLS) in Brookhaven, USA, using a BioCARS Laue bench camera, home-built freezing equipment, a 0.2 mm collimator and Fuji HR-III_N imaging plates, 20.1 × 25.2 cm in size. The Fuji plates were digitized off-line with a BAS2000 scanner at 100 μm raster size. Three of the four data sets were collected with the white beam focused by a Pt-coated cylindrical mirror; one data set, 1DA, was collected without focusing. A 25 μm Zr foil was placed in the direct beam between the shutter and the collimator to obtain a sharp λ_{\min} cut-off in the white X-ray spectrum.

The exposure times varied between 0.5 and 8 s, the crystal-to-detector distances between 37 and 52 mm (see Table 1). Although the purpose was to collect a minimum of ten frames for each data set, a number of frames had to be rejected for reasons such as scanning problems, crystal cracking, bad image plates *etc.* For each data set, a total of ~ 1.5 h was spent on crystal selection and mounting, collecting and scanning test frames, actual data collection and frame scanning. A carousel could not be used due to the extremely short crystal-to-detector distance, so that the image plates had to be changed manually after each exposure. Table 1 lists the final number of frames used for each data set, and the spindle angles at which the frames were collected.

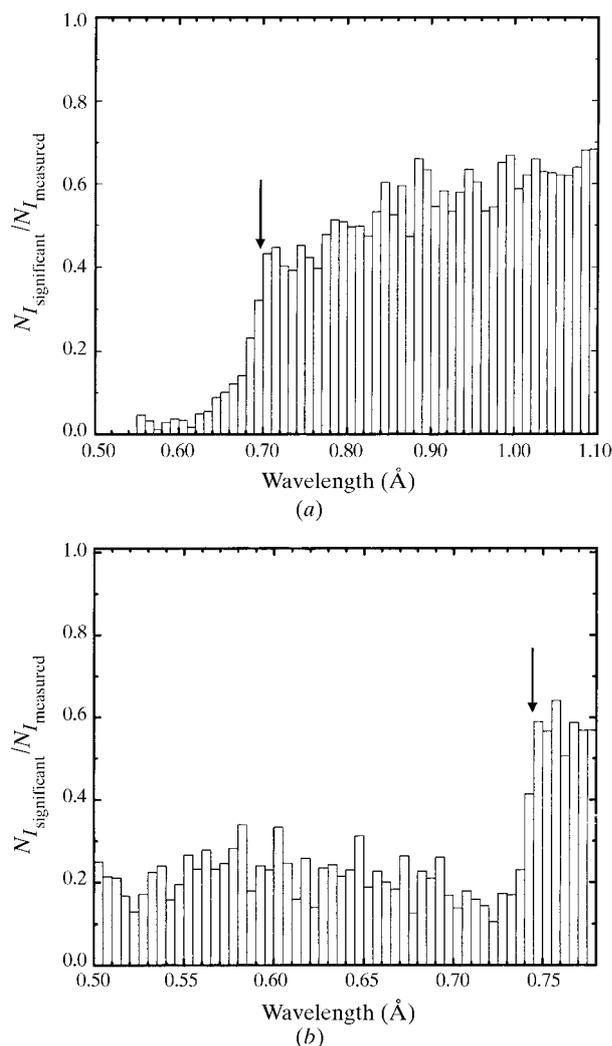


Figure 1
Histograms of the ratio of the number of significant intensities to the number of measured intensities (non-overlapping singles) as a function of the wavelength. The wavelength is on a scale obtained from an initial estimate (based on λ_{\min} of the X-ray beam) of the cell volume (Table 2). The absorption edge of Zr, indicated by an arrow, is located at the sharp decline in the histograms. All the frames within one data set are used. (a) Histogram for DMANM. (b) Histogram for 1DA.

4. Data processing

All test frames were indexed immediately after collection and scanning using the program *LAUECELL* (Ravelli *et al.*, 1996), which is able to determine the crystal orientation as well as the relative cell parameters. In the case of serious indexing problems, the program *LAUEGEN* (Campbell, 1995) was used first to determine the direct-beam position more accurately, after which *LAUECELL* was able to find the correct cell. Cell and orientation parameters were imported into *LAUEGEN* (Campbell, 1995). The *LAUEGEN* program was used to check whether the predicted pattern could be refined to acceptable values of the root-mean-square deviation between predicted and measured spot positions, whether the λ_{\min} Zr edge could be observed, and whether there were enough high-resolution spots. The latter two checks can be performed by the histogram analysis developed by Hao *et al.* (1995). A different crystal was mounted in case of any indication of problems.

Processing of the full data sets was started by indexing using *LAUECELL*. Cell volumes were estimated by using 0.55 and 1.8 Å for λ_{\min} and λ_{\max} , respectively, and by comparing the number of predicted spots with the number of observed spots as a function of cell volume and resolution. The Bravais lattice of the crystal was chosen after running the program *LEPAGE* (LePage, 1982). Full refinement and integration were performed independently with *LAUEGEN* (Campbell, 1995) and *LAUEVIEW* (Ren & Moffat, 1995a). The refinement gave comparable results for both programs, in the range 0.03–0.08 mm for the r.m.s. deviation between the observed and predicted spot positions for a maximum number of spots. Integration using the empirical-profile fitting method implemented in *LAUEGEN* not only proved to be much faster than the analytical profile-fitting method used in *LAUEVIEW*; in some unfavourable cases of split spot profiles at high values of 2θ it also appeared to be more accurate. The histogram analysis of *LAUEGEN* was used to derive the absolute scale of the unit-cell parameters, by inspecting the ratio between the number of spots with significant intensity and the number of predicted spots, *versus* the wavelength. The histograms of the individual frames were averaged for each data set; Fig. 1 shows two examples of histograms thus obtained. Values of the starting cell volume, the wavelength found for the Zr edge at the sharp decline in the histograms, scaled cell parameters and final cell volumes are listed in Table 2.

All frames were re-integrated using the scaled cell dimensions and resolution limits. A value of 0.40 Å was taken for λ_{\min} in the case of an unfocused beam, and 0.60 Å in the case of a focused beam. It was necessary to use λ_{\min} values lower than the wavelength of the Zr edge (λ_{edge}) to account for the few spots that still had significant intensity below the absorption edge (see Fig. 1). Since only a few spots showed overlap, a resolution-dependent filter could be applied after the integration (performed by

Table 2

Crystallographic data I: (scaled) cell dimensions.

Estimated standard deviations are given in brackets. The absorption edge of Zr, used to calculate the scale factors, is at 0.6889 Å. Compound names are given in the text.

Compound code	DMANM	DMANMInt	tBDD	1DA
Initial cell volume (Å ³)	2000	2000	600	2000
$\lambda_{\text{edge}}^{\text{observed}}$ (Å)	0.700 (10)	0.700 (10)	0.765 (15)	0.745 (5)
Scale factor on volume	0.95 (2)	0.95 (2)	0.73 (4)	0.791 (16)
Scaled cell parameters				
<i>a</i> (Å)	9.46 (14)	9.45 (14)	5.24 (10)	8.98 (6)
<i>b</i> (Å)	12.56 (18)	12.59 (18)	10.6 (2)	11.20 (8)
<i>c</i> (Å)	17.0 (2)	16.9 (2)	7.99 (16)	16.36 (11)
α (°)	106.4 (2)	105.70 (7)	90	90
β (°)	97.71 (3)	97.74 (5)	96.758 (10)	105.967 (16)
γ (°)	93.797 (5)	93.70 (5)	90	90
Scaled cell volume (Å ³)	1910 (80)	1910 (80)	440 (25)	1580 (30)
Bravais lattice	<i>P</i> triclinic	<i>P</i> triclinic	<i>P</i> monoclinic	<i>P</i> monoclinic

LAUEGEN; Campbell, 1995): all spots that were predicted (based on a known starting λ curve and Wilson statistics) to have an intensity lower than a certain threshold were discarded. The singles were merged based on the highest-symmetry Laue group belonging to the Bravais lattice that was chosen in *LAUECELL*. The values of R_{merge} and the overall appearance of the wavelength-normalization curve were carefully inspected during the scaling to check if the correct Laue group had been selected.

The scaling of the data was performed using the program *LAUEVIEW* (Ren & Moffat, 1995a). Low-order Chebyshev polynomials (third or fourth order) were fitted against the singles diffracting at a wavelength between 0.72 and 1.6 Å in order to minimize R_{merge} . The small number of singles at higher wavelengths made it necessary to refine the wavelength-normalization curve extremely carefully. Only the parts at lower wavelengths (below ~ 1.1 Å) could be refined with higher-order Chebyshev terms (8) without showing overfitting. The refinement of the λ curve was alternated with refinement of isotropic scale and temperature factors. The combined wavelength-normalization curves are shown in Fig. 2. Only details in the curves that are visible between 0.8 and 1.2 Å are likely to be realistic: details in the curves beyond 1.2 Å are artifacts caused by appending different parts of the curves. In the case of the unfocused beam, it was possible to extend the λ curve below the absorption edge of Zr.

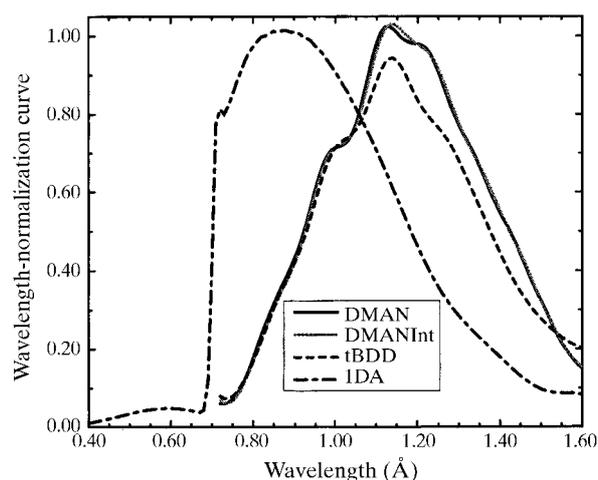
The wavelength-normalization curves were used to deconvolute the multiples (Ren & Moffat, 1995b), which were scaled and included in the final data sets. All R_{merge} factors between the deconvoluted multiples and the singles are good, between 2.7 and 7.8%, which is not surprising since the singles are actively used in the energy-deconvolution procedure (Ren & Moffat, 1995b). Table 3 gives an overview of the statistics of the final data sets.

5. Structure determination

The program *PLATON* (Spek, 1990) was used to determine the different possible space groups for each

compound. Glide planes were identified for tBDD and 1DA, but no data was available to discriminate between screw and rotation axes. *SIR92* (Altomare *et al.*, 1994) was used for each data set to determine the $C(z)$ distribution (as defined in §2.4). These distributions (shown in Fig. 3) show that all data sets have values of z larger than 3 for most of their strong reflections, indicating that the (systematic) incompleteness of these data sets should not impose serious limitations on the application of direct methods. If only the singles are used in the $C(z)$ distribution, the values are less promising.

The structures were solved by direct methods revealing most of the non-H atoms [using both *SHELXS97* (Sheldrick, 1997) and *SIR92* (Altomare *et al.*, 1994)], followed by difference Fourier syntheses. All possible space groups were checked: $P2/n$ and $P2_1/n$ for tBDD, $P1$ and $P1$ for DMANM and DMANMInt, and $P2/c$ and $P2_1/c$ for 1DA. The correct space groups were corroborated by the struc-

**Figure 2**

Wavelength-normalization curves. An unfocused beam was used for 1DA with the synchrotron running in single-bunch mode. A focused beam was used for DMANM, DMANMInt and tBDD, while the synchrotron was running in 25-bunch mode.

Table 3

Crystallographic data II: data set statistics.

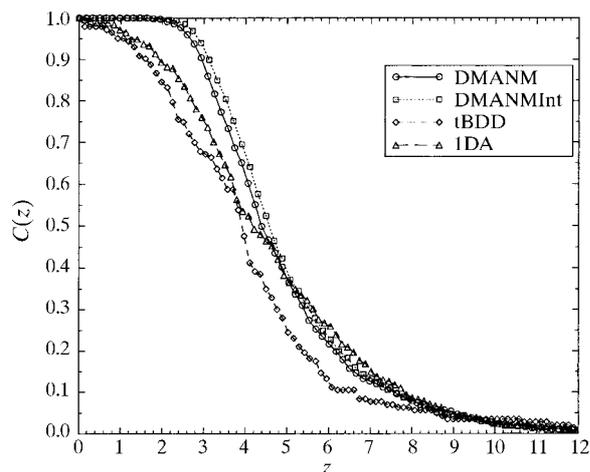
Compound names are given in the text.

Compound code	DMANM	DMANMInt	tBDD	1DA
d_{\min} (Å)	0.85	0.75	0.75	0.70
Singles				
Number of measurements	10181	11877	2352	9571
Number of unique reflections	4581	5920	628	3228
Redundancy	2.2	2.0	3.7	3.0
$R_{\text{merge}}^{\dagger}$	12.3	11.6	13.2	11.5
Multiples				
Number of multiples	762	686	188	751
Total				
Number of unique reflections	4799	6168	709	3478
Completeness (%)				
$\infty-3d_{\min}$	30.0	34.0	47.1	39.3
$3d_{\min}-2d_{\min}$ ($\infty-2d_{\min}$)	73.5 (60.5)	76.4 (63.8)	85.6 (72.9)	70.5 (60.8)
$2d_{\min}-d_{\min}$ ($\infty-d_{\min}$)	75.6 (73.7)	65.2 (65.0)	59.7 (61.4)	69.7 (68.6)

$\dagger R_{\text{merge}}$ is defined as $\sum_{\mathbf{h}} \sum_i (|F_{\mathbf{h}_i}^2 - \langle F_{\mathbf{h}}^2 \rangle|) / \sum_{\mathbf{h}} \sum_i F_{\mathbf{h}_i}^2$, with $\langle F_{\mathbf{h}}^2 \rangle = (1/N) \sum_{i=1}^N F_{\mathbf{h}_i}^2$, where the summation is performed over all identical or equivalent reflections.

ture determination and refinement; they are listed in Table 4.

Refinement on F^2 with all unique reflections was carried out by full-matrix least-squares techniques (*SHELXL93*; Sheldrick, 1993). H atoms not belonging to hydroxyl and methyl groups were introduced on calculated positions and included in the refinement riding on their carrier atoms with isotropic thermal parameters related to U_{eq} of the carrier atoms. Only those hydroxyl and methyl H atoms that could be located from inspection of the electron-density difference maps were included in the refinement with geometry constraints. All non-H atoms were refined with anisotropic thermal parameters. Weights were introduced in the final refinement cycles. Numerical details of the structure determination are given in Table 4. Thermal-motion ellipsoid plots are shown in Fig. 4.

**Figure 3**

Cumulative distribution function of $z = \langle \alpha \rangle / \sigma_{\alpha}$, as defined by Giacovazzo *et al.* (1994). For all our Laue data sets, z values larger than 3 are found for most of the strong reflections, indicating that the systematic incompleteness of these data sets would not impose serious limitations on application of direct methods.

The structure of 1DA was the only one that could be compared with that of a published monochromatic study (van Walree *et al.*, 1997), carried out independently on the same batch of crystals. The cell volume determined by the traditional monochromatic method was $1550.0 (2) \text{ \AA}^3$ (measured at 150 K), which verified the correctness of the scaling of the unit cell obtained from the Laue data [volume $1580 (30) \text{ \AA}^3$, measured at room temperature]. For tBDD, cell parameters were published previously (Miesen *et al.*, 1994). As in the case of 1DA, the volume obtained from the Laue data [$440 (25) \text{ \AA}^3$] is in good agreement with the volume determined by monochromatic experiment (442.4 \AA^3).

6. Discussion and conclusions

In this paper it is shown that *ab initio* structure determination of low-molecular-weight compounds using Laue data is feasible. The step that makes the structure determination independent from monochromatic experiment is the scaling of the unit cell to an absolute volume. The use of a Zr foil gives a sharp edge in the wavelength-normalization curve. This edge can be used to scale the cell volume with an accuracy of a few percent. This method of determining the absolute cell volume has been demonstrated before (Carr *et al.*, 1992, 1993; Kariuki & Harding, 1995), but not applied to unknown structures thus far.

For the compounds presented here, it proved to be possible to solve the structure by means of direct methods. Although our data sets are far from complete, enough strong triplets could be found. The same observation was made by Helliwell, Gomez de Anderez *et al.* (1989), who solved an organic structure with Laue data using the Sayre tangent formula. In connection with the discussion by Harding (1988) about the risk that direct methods lose their power to solve the structure of low-molecular-weight compounds using only Laue data, we used the $C(z)$ distri-

Table 4

Details of the structure determinations of the low-molecular-weight compounds.

The number of unique reflections used in the refinement is often lower than the total number of reflections in each data set, as given in Table 3. This is due to rejection of systematically absent and inconsistent equivalent reflections.

Compound code	DMANM	DMANMInt	tBDD	1DA
Space group	$P\bar{1}$	$P\bar{1}$	$P2_1/n$	$P2_1/c$
Number of molecules per unit cell	2	2	2	4
$F(000)$	852	852	180	656
Number of unique reflections	4780	6168	686	3411
with $F^{\text{obs}} > 4\sigma(F^{\text{obs}})$	4032	5374	629	2852
Number of parameters	553	545	55	211
Weighting scheme†				
w_1	0.0870	0.1704	0.1665	0.1546
w_2	1.32	1.56	0.19	0.13
Final R_1 ‡ [$F^{\text{obs}} > 4\sigma(F^{\text{obs}})$]	0.0906	0.1060	0.1012	0.1096
R_1 (all data)	0.1129	0.1233	0.1209	0.1358
wR_2	0.2091	0.2671	0.2573	0.2774
S	1.041	1.071	1.133	1.130
Minimum residual density ($\text{e } \text{Å}^{-3}$)	-0.31	-0.46	-0.55	-0.63
Maximum residual density ($\text{e } \text{Å}^{-3}$)	0.29	0.47	0.49	0.58

† Weighting scheme is defined as $w^{-1} \equiv \sigma^2(F^{\text{obs}})^2 + (w_1P)^2 + w_2P$, where the parameter $P \equiv \{\max(F_h^{\text{obs}}, 0) + 2F_h^{\text{calc}^2}\}/3$. ‡ $R_1 = \sum_h (|F_h^{\text{obs}} - |F_h^{\text{calc}}|) / \sum_h |F_h^{\text{obs}}|$. $wR_2 = [\sum_h \{w(F_h^{\text{obs}^2} - F_h^{\text{calc}^2})^2\} / \sum_h \{w(F_h^{\text{obs}^2})^2\}]^{1/2}$. The goodness of fit, S , is based on F^2 : $S = [\sum_h \{w(F_h^{\text{obs}^2} - F_h^{\text{calc}^2})^2\} / (n - p)]^{1/2}$, where n is the number of reflections and p is the total number of parameters refined.

bution devised by Giacovazzo *et al.* (1994) to evaluate this risk.

The potential of using Laue data for structure refinement has been demonstrated before, both for low-molecular-weight compounds (*e.g.* Wood *et al.*, 1983; Gomez de Anderes *et al.*, 1989; Dodd *et al.*, 1994; Kariuki & Harding, 1995) and for proteins (Bourgeois *et al.*, 1997; Yang *et al.*, 1997; Ravelli *et al.*, 1998). The Laue data presented in this paper are of modest quality in terms of R_{merge} and completeness. A partial explanation for the high R factors is related to the long exposure times for DMANM, DMANMInt and tBDD given a focused beam and the synchrotron running in 25 bunch mode. These times are indicative of some problems with the beamline at the time of the measurement of these three data sets. Nevertheless, the structures obtained show good geometry;† all temperature factors appear to be physically reasonable, and show the expected differences between the room-temperature and the liquid-nitrogen temperature data sets. Comparison of one structure (1DA) with the published structure determined in a monochromatic experiment showed a low r.m.s. deviation for all non-H-atom positions: 0.019 Å. For all structures a significant decrease in the R factors was observed upon introduction of the non-hydroxyl and non-methyl H atoms. Furthermore, a number of hydroxyl and methyl H atoms showed up in the electron-density difference maps.

The Laue technique could be used more routinely for structure determination of low-molecular-weight compounds, provided that a number of improvements are made:

(i) Our data showed a severe θ cut-off. This has serious implications on the sampling of the reflections at longer wavelengths at which the highest percentage of singles can be found (Cruickshank *et al.*, 1987; Ravelli, 1998), causing problems in the accuracy of the determination of the wavelength-normalization curve. An inaccurate λ curve affects the deconvolution of all multiples, which constitute up to 17% of our data. Refinement of the λ curve was unstable at higher wavelengths, and could only be performed using very low-order Chebyshev polynomials in that region. Since the crystal-to-detector distance was very short, a θ cut-off could only be prevented by eliminating the low-energy radiation from the X-ray spectrum. This could have been obtained by absorption of all longer wavelengths using thick Cu or Al attenuators.

(ii) The data completeness can be improved by collecting a larger number of frames for each compound. For unknown unit cells, a sampling of reciprocal space of 180° is advisable, and an angular width of 6° is needed to overlap the accessible part of reciprocal space from different exposures at lower resolutions [the maximum angular width depends on the X-ray spectrum and the desired resolution at which the accessible regions should overlap (see Moffat, 1997; Yang *et al.*, 1997; Ravelli, 1998)]. More frames also increase the redundancy of the single measurements, which stabilizes the refinement of the λ curve and the frame scaling factors.

(iii) Absorption correction can be important, especially in the case of Laue diffraction experiments on inorganic and organometallic compounds. Since the crystal is stationary in the Laue method, we may assume that the path length through the crystal and the air between crystal and detector varies smoothly, for different beams collected on one frame, as a function of the position of the spot.

† Deposited in the Cambridge Structural Database.

Maginn *et al.* (1993) demonstrated that the refinement of an absorption surface for each frame can give a significant reduction both in R_{merge} and R factor. For our data sets, we tried to refine an absorption surface by using the approach of Ren & Moffat (1995a): refinement of anisotropic scale

factors. This led to overfitting; although the R_{merge} values could be reduced by 2 to 3%, the structures started to show non-realistic temperature movements.

(iv) The scaling of the unit-cell parameters requires a sharp and clear edge in the X-ray spectrum. Obviously, a

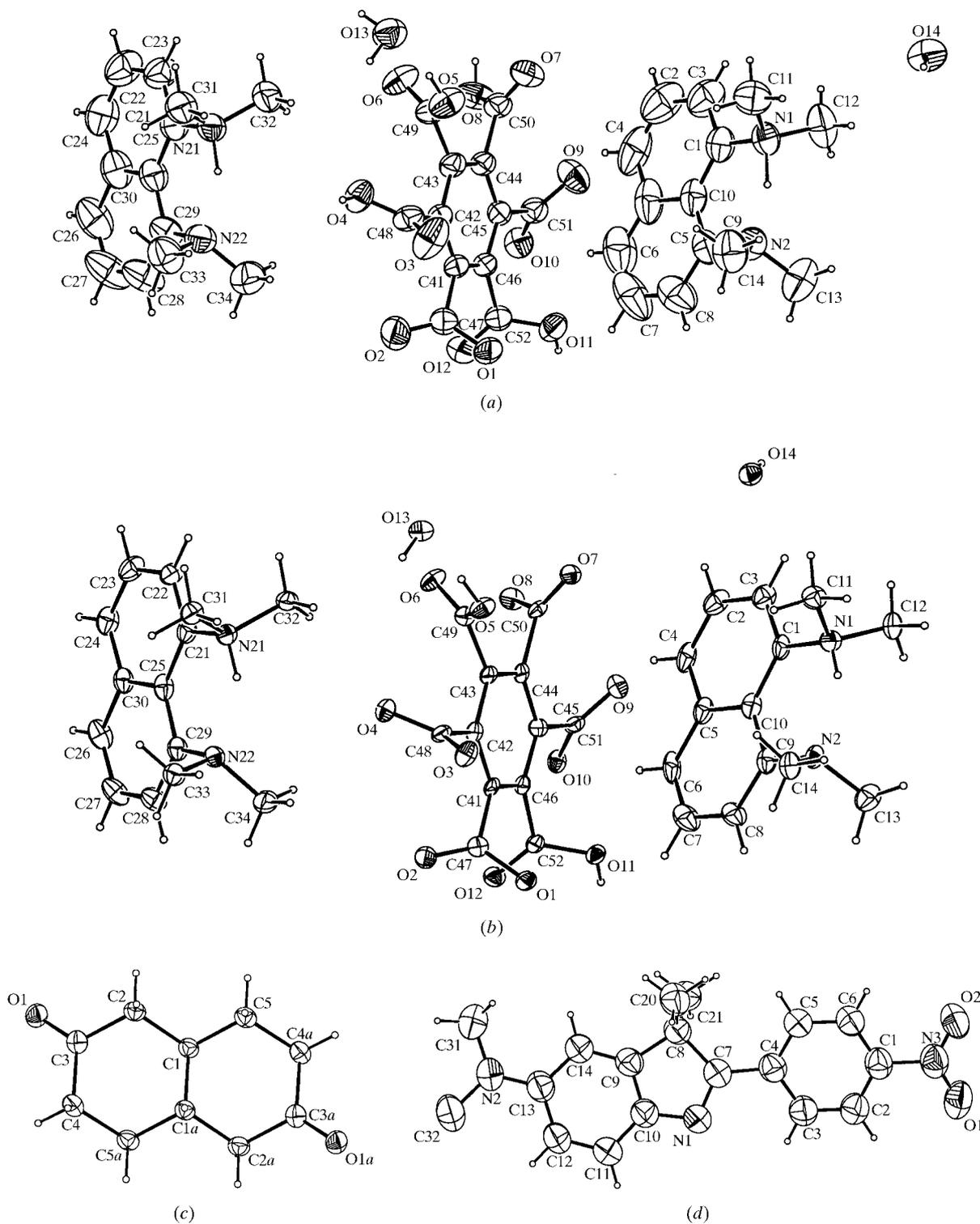


Figure 4 Thermal-motion ellipsoid plots (50% probability) with the adopted atom labelling of the X-ray structures of (a) DMANM, (b) DMANInt, (c) tBDD and (d) 1DA.

clear edge well above the minimum wavelength of the X-ray spectrum, as was obtained in the λ curve of 1DA (Fig. 2), is to be preferred over the edge in the other three curves, where it is located just above the minimum wavelength of the incident X-rays. The tail in the λ curve of 1DA at lower wavelengths, however, caused some difficulties with the deconvolution, which requires the full λ curve to be known accurately. Combination of different attenuators should prevent this problem. Alternatively, absorption edges of Pt, present in some coatings of focusing mirrors, may be used to scale the unit-cell parameters, provided that the measured data are sufficiently accurate and redundant.

(v) The integration software for Laue data collected on crystals of low-molecular-weight compounds is still amenable to improvements. The small number of spots per frames, the extremely low background for some spots and the large dynamical range of the spot intensity constitute a challenge for spot-profile fitting methods. The introduction of an extinction coefficient during the structure refinement showed a reduction of R factors for all our data sets, but only for unrealistically high values of this coefficient. This indicates either a systematic overestimation of the high-resolution reflections or a systematic underestimation of the low-resolution reflections. The latter may be caused by the deconvolution algorithm, or by an improper integration of intense spots (D. Bourgeois, personal communication).

We believe that, once the conditions listed above are fulfilled and structures can be obtained as easily and accurately as in the case of the monochromatic experiment using in-house laboratory equipment and standard software, one will be in a better position to exploit more widely the most important application of the Laue method: time-resolved (perturbation) studies.

The investigations were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO) and by the European Union. We would like to thank Grace Shea and Ling Peng for assistance in data collection, Dritan Siliqi for help in obtaining Fig. 3, and Joel Sussman for supporting these experiments and giving us the opportunity to collect data at X26C.

References

- Altomare, A., Cascarano, G., Giacovazzo, C., Guagliardi, A., Burla, M. C., Polidori, G. & Camalli, M. (1994). *J. Appl. Cryst.* **27**, 435–436.
- Bartunik, H. D. & Borchert, T. (1989). *Acta Cryst.* **A45**, 718–726.
- Bolduc, J. M., Dyer, D. H., Scott, W. G., Singer, P., Sweet, R. M., Koshland, D. E. & Stoddard, B. L. (1995). *Science*, **268**, 1312–1318.
- Bourenkov, G. P., Popov, A. N. & Bartunik, H. D. (1996). *Acta Cryst.* **A52**, 797–811.
- Bourgeois, D., Longhi, S., Wulff, M. & Cambillau, C. (1997). *J. Appl. Cryst.* **30**, 153–163.
- Campbell, J. W. (1995). *J. Appl. Cryst.* **28**, 228–236.
- Carr, P. D., Cruickshank, D. W. J. & Harding, M. M. (1992). *J. Appl. Cryst.* **25**, 294–308.
- Carr, P. D., Dodd, I. M. & Harding, M. M. (1993). *J. Appl. Cryst.* **26**, 384–387.
- Cascarano, G., Giacovazzo, C., Burla, M. C., Nunzi, A. & Polidori, G. (1984). *Acta Cryst.* **A40**, 278–283.
- Cheetham, G. M. T., Carr, P. D., Dodd, I. M., Kariuki, B. M. & Harding, M. M. (1995). *J. Synchrotron Rad.* **2**, 300–307.
- Cochran, W. (1955). *Acta Cryst.* **8**, 473–478.
- Cruickshank, D. W. J., Helliwell, J. R. & Moffat, K. (1987). *Acta Cryst.* **A43**, 656–674.
- Dodd, I. M., Hao, Q., Harding, M. M. & Prince, S. M. (1994). *Acta Cryst.* **B50**, 441–447.
- Genick, U. K., Borgstahl, G. E. O., Ng, K., Ren, Z., Pradervand, C., Burke, P. M., Šrajcar, V., Teng, T.-Y., Schildkamp, W., McRee, D. E., Moffat, K. & Getzoff, E. D. (1997). *Science*, **275**, 1471–1475.
- Giacovazzo, C., Guagliardi, A., Ravelli, R. B. G. & Siliqi, D. (1994). *Z. Kristallogr.* **209**, 136–142.
- Gomez de Anderez, D., Helliwell, H., Habash, J., Dodson, E. J., Helliwell, J. R., Bailey, P. D. & Gammon, R. E. (1989). *Acta Cryst.* **B45**, 482–488.
- Gouet, P., Jouve, H. M., Williams, P. A., Andersson, I., Andreoletti, P., Nussaume, L. & Hajdu, J. (1996). *Nat. Struct. Biol.* **3**, 951–956.
- Hajdu, J., Almo, S. C., Farber, G. K., Prater, J. K., Petsko, G., Wakatsuki, S., Clifton, I. & Fülöp, V. (1991). *Crystallographic Computing 5*, edited by D. Moras, A. Podjarny & J. Thierry, ch. 3, pp. 29–49. Oxford University Press.
- Hajdu, J. & Andersson, I. (1993). *Annu. Rev. Biophys. Biomol. Struct.* **22**, 467–498.
- Hao, Q., Campbell, J. W., Harding, M. M. & Helliwell, J. R. (1993). *Acta Cryst.* **A49**, 528–531.
- Hao, Q., Harding, M. M. & Campbell, J. W. (1995). *J. Appl. Cryst.* **28**, 447–450.
- Harding, M. M. (1988). *Chemical Crystallography with Pulsed Neutrons and Synchrotron X-rays*, edited by M. Carrondo & G. Jeffrey, pp. 537–561. Dordrecht: Riedel.
- Harding, M. M., Maginn, S. J., Campbell, J. W., Clifton, I. & Machin, P. (1988). *Acta Cryst.* **B44**, 142–146.
- 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.
- Helliwell, M., Gomez de Anderez, D., Habasch, J., Helliwell, J. R. & Vernon, J. (1989). *Acta Cryst.* **B45**, 591–596.
- Kariuki, B. M. & Harding, M. M. (1995). *J. Synchrotron Rad.* **2**, 185–189.
- Le Page, Y. (1982). *J. Appl. Cryst.* **15**, 255–259.
- Maginn, S. J., Harding, M. M. & Campbell, J. W. (1993). *Acta Cryst.* **B49**, 520–524.
- Miesen, F. W. A. M., Wollersheim, A. P. P., Meskers, S. C. J., Dekkers, H. P. J. M. & Meijer, E. W. (1994). *J. Am. Chem. Soc.* **116**, 5129–5133.
- Moffat, K. (1997). *Methods Enzymol.* **277**, 433–447.
- Ravelli, R. B. G. (1998). PhD thesis, Utrecht University, The Netherlands.
- Ravelli, R. B. G., Hezemans, A. M. F., Krabbendam, H. & Kroon, J. (1996). *J. Appl. Cryst.* **29**, 270–278.
- Ravelli, R. B. G., Raves, M. L., Ren, Z., Bourgeois, D., Roth, M., Kroon, J., Silman, I. & Sussman, J. (1998). *Acta Cryst.* **D54**, 1361–1368.
- Ren, Z. & Moffat, K. (1995a). *J. Appl. Cryst.* **28**, 461–481.
- Ren, Z. & Moffat, K. (1995b). *J. Appl. Cryst.* **28**, 482–493.
- Rossmann, M. G. (1979). *J. Appl. Cryst.* **12**, 225–238.
- Schlichting, I., Almo, S. C., Rapp, G., Wilson, K., Petratos, K., Lentfer, A., Wittinghofer, A., Kabsch, W., Pai, E. F., Petsko, G. A. & Goody, R. S. (1990). *Nature (London)*, **345**, 309–315.

- Sheldrick, G. M. (1993). *SHELXL93. Program for Crystal Structure Refinement*. University of Göttingen, Germany.
- Sheldrick, G. M. (1997). *SHELXS97. Program for Crystal Structure Determination*. University of Göttingen, Germany.
- Snell, E., Habash, J., Helliwell, M., Helliwell, J. R., Raftery, J., Kaucic, V. & Campbell, J. W. (1995). *J. Synchrotron Rad.* **2**, 22–26.
- Spek, A. L. (1990). *Acta Cryst.* **A46**, C-34.
- Šrajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M. & Moffat, K. (1996). *Science*, **274**, 1726–1729.
- Verschueren, K. H. G., Seljée, F., Rozeboom, H. J., Kalk, K. H. & Dijkstra, B. (1993). *Nature (London)*, **363**, 693–698.
- Walree, C. A. van, Maarsman, A. W., Marsman, A. W., Flipse, M. C., Smeets, W. J. J., Spek, A. L. & Jenneskens, L. W. (1997). *J. Chem. Soc. Perkin Trans. 2*, pp. 809–819.
- Wood, I. G., Thompson, P. & Matthewman, J. C. (1983). *Acta Cryst.* **B39**, 543–547.
- Xiao, B., Smerdon, S. J., Jones, D. H., Dodson, G. G., Sonejl, Y., Aitken, A. & Gamblin, S. J. (1995). *Nature (London)*, **376**, 188–191.
- Xie, Y. & Hao, Q. (1997). *Acta Cryst.* **A53**, 643–648.
- Yang, X., Ren, Z. & Moffat, K. (1997). *Acta Cryst.* **D54**, 367–377.