Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Detection and characterization of merohedral twinning in two protein crystals: bacteriorhodopsin and p67^{phox}

Antoine Royant,[a,b]† Sylvestre Grizot,[a]†§ Richard Kahn,[a] Hassan Belrhali,[b]‡ Franck Fieschi,[a] Ehud M. Landau[c] and Eva Pebay-Peyroula[a]*

[a]Institut de Biologie Structurale, CEA–CNRS–Université Joseph Fourier UMR5075, 41 Rue Jules Horowitz, F-38027 Grenoble CEDEX 1, France, [b]European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, BP 220, F-38043 Grenoble CEDEX, France, and [c]Membrane Protein Laboratory, Department of Physiology and Biophysics and Sealy Centers for Structural Biology and Molecular Science, The University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-0641, USA

† These authors contributed equally to this work.
‡ Current address: EMBL, 6 Rue Jules Horowitz, BP 181, F-38042 Grenoble CEDEX 9, France.
§ Current address: NIDDK/LMB, 50 South Drive, Bethesda, MD 20892-8030, USA.

Correspondence e-mail: pebay@ibs.fr

Received 25 October 2001
Accepted 28 February 2002

The estimation of the twinning ratio in crystals of two proteins, bacteriorhodopsin and a truncated form of p67^{phox}, is described. Various approaches are combined to detect and determine accurately the twinning ratio. For each protein, three data sets from crystals exhibiting twinning ratios ranging from 0 to almost 50% are analysed. Self-rotation functions and R_{sym} values considering space groups of higher symmetry are indicative of twinning. Precise values of the twinning ratios are derived from Yeates' statistical approaches and Britton plots. The twinning ratios are also obtained by analysing the second-, third- and fourth-order moments of the intensity distribution. The twinning ratios obtained from the various approaches are compared. Second-, third- and fourth-order moments of the intensity distribution are useful indicators for evaluating the twinning precisely in order to correct the intensities or to screen rapidly for non-twinned crystals.

1. Introduction

Twinning has recently been observed in numerous protein crystal systems (Luecke *et al.*, 1998; Valegard *et al.*, 1998; Royant *et al.*, 2000; Frazão *et al.*, 1999; Ban *et al.*, 1999; Taylor *et al.*, 2000). Twinning is a crystal-growth defect that arises when two or more domains within a crystal have different orientations and may even vary among crystals grown in the same batch. If the lattices of the various domains overlay perfectly in three dimensions, the twinning is merohedral. As a consequence, the reciprocal lattices of these domains coincide, summing the intensities of the Bragg spots corresponding to different Miller indices. Hemihedral twinning is a special case of merohedral twinning that occurs when two crystal domains exist. In the case of a 50% hemihedral twinning, the crystal is composed of two domains of equal volume and it is impossible to retrieve the diffracting intensities for each of the individual domains. For twinning ratios smaller than 50% diffraction intensities can be corrected. It is known from small-molecule studies that some space groups favour twinning. This is particularly the case for hexagonal, trigonal or tetragonal space groups (for reviews, see Yeates, 1997; Chandra *et al.*, 1999). Because small-molecule crystals usually diffract to very high resolution, the ratio between the number of refined parameters and the number of experimental observations is large enough to allow a global refinement of both crystal domains at the same time.

The mixing of diffraction spots with different Miller indices affects the global intensity distribution and adds additional symmetry relationships, relative to crystallographic symme-

tries, among the Bragg reflections. The characterization of twinning (existence, type and ratio) is a crucial step when diffraction data from a new crystal form are analysed. Various twinning analyses for protein crystals have been developed in the last 20 y (Fisher & Sweet, 1980; Rees, 1980, 1982; Yeates, 1988, 1997; Gomis-Rüth *et al.*, 1995; Chandra *et al.*, 1999). Presently, although twinning poses an additional difficulty, it is not necessarily an obstacle to solving structures. The possibility of molecular replacement from twinned crystals has been discussed (Breyer *et al.*, 1999), various refinement programs such as *SHELX* (Sheldrick & Schneider, 1997) or *CNS* (Brünger *et al.*, 1998) consider twinning, and even multiple-wavelength anomalous diffraction methods have been used to solve a structure from twinned crystals (Yang *et al.*, 2000). Many analyses used to detect twinning exploit the additional symmetry by examining the $R_{\rm sym}$ values in higher space groups, self-rotation functions or deviations from Wilson statistics (Rees, 1980, 1982; Yeates, 1988, 1997). These criteria are usually used to detect non-crystallographic symmetry. Packing considerations might possibly be sufficient to exclude the presence of two or more molecules in the asymmetric unit related by non-crystallographic symmetry. Alternatively, one can follow the number of negative intensities obtained during the correction of the diffraction data from the twinning effect as a function of the twinning ratio (Fisher & Sweet, 1980). Moreover, when a first model of the structure is available, it is also possible to compare the observed Bragg intensities with calculated intensities (Redinbo & Yeates, 1993; Gomis-Rüth *et al.*, 1995).

Here, we report analyses of two protein crystals: bacteriorhodopsin (BR) and a truncated form of p67^{phox}, a cytosolic factor of the NADPH oxidase from neutrophils. Highly ordered three-dimensional microcrystals of bacteriorhodopsin, a light-driven proton pump located in the membrane of *Halobacterium salinarum*, were obtained for the first time by a novel crystallization approach using lipidic cubic phases (Landau & Rosenbusch, 1996). The space group is $P6_3$ and the structure was initially solved by molecular replacement to 2.5 Å resolution (Pebay-Peyroula *et al.*, 1997). The refinement was further extended to 2.35 Å, taking a twinning ratio of 18% into account (PDB code 1ap9). The resolution was subsequently improved owing to the availability of larger crystals and the structure was refined to 2.2 Å resolution using diffraction data from 35% twinned crystals (Pebay-Peyroula, unpublished results). Finally, a structure at 1.9 Å resolution was obtained from non-twinned crystals (Belrhali *et al.*, 1999). Using the same lipidic cubic phase crystallization method, the BR structure was refined by a different group using almost perfectly twinned crystals (Luecke *et al.*, 1998) to 2.3 Å resolution and to 1.55 Å resolution (Luecke *et al.*, 1999) from 24% twinned crystals. The two highest resolution structures refined with *CNS* (Belrhali *et al.*, 1999) and refined with *SHELX* (Luecke *et al.*, 1999) are highly similar. In the case of p67^{phox}, a truncated form of the protein crystallizes in space group $P3_1$. Since no similar structure was known, experimental phases were determined by the SAD (single-wavelength anomalous diffraction) method from a protein containing selenomethionines instead of methionines (Grizot *et al.*, 2001). Merohedral twinning was observed for both protein crystals, with twinning ratios ranging from 0 to 50%. Knowing the twinning ratio for each data set, we were able to correct the experimental data and extract the diffraction intensities produced by a single crystal domain. In the following, we report how twinning was detected qualitatively from self-rotation functions and $R_{\rm sym}$ values and how it was assessed quantitatively from Yeates statistics and Britton plots. In addition, we derive the second-, third- and fourth-order moments of the intensity distribution as a function of twinning ratio and discuss the possibility of quantitatively determining the twinning ratios from these values. Analysis of twinning ratios allowed efficient screening for non-twinned crystals, which was necessary to collect SAD data of high quality in the case of p67^{phox} (Grizot *et al.*, 2001). It also enabled screening for non-twinned crystals for the ground state and the K-state structure of BR (Belrhali *et al.*, 1999; Edman *et al.*, 1999) and to correct the data from twinning for the L-state of BR (Royant *et al.*, 2000).

2. Results and discussions

2.1. Generalities on merohedral twinning

For merohedral twinning, the symmetry operation relating domain 1 to domain 2 within a crystal is such that the Bragg spots with different Miller indices of both domains coincide exactly, giving rise to mixed intensities. Thus, Bragg spots associated with $\mathbf{h}_1$ from one domain coincide perfectly with those associated with $\mathbf{h}_2$ from the second domain, where $\mathbf{h}_2$ is related to $\mathbf{h}_1$ by a symmetry operation S. The intensities measured from the twinned crystal can be expressed as

$$I^T(\mathbf{h}_1) = (1-x)I^U(\mathbf{h}_1) + xI^U(\mathbf{h}_2), \quad (1)$$

$$I^T(\mathbf{h}_2) = xI^U(\mathbf{h}_1) + (1-x)I^U(\mathbf{h}_2), \quad (2)$$

where x is the volume fraction of the smallest domain, *i.e.* the twinning ratio, $\mathbf{h}_2 = S(\mathbf{h}_1)$ and I^U is the intensity diffracted by a non-twinned crystal. These equations are reduced into a single equation in the case of perfect (50%) twinning,

$$I^T(\mathbf{h}_1) = I^T(\mathbf{h}_2) = [I^U(\mathbf{h}_1) + I^U(\mathbf{h}_2)]/2, \quad (3)$$

introducing a perfect additional symmetry on intensities. As a consequence, there is no way to retrieve the intensities of the individual domains in the case of perfect twinning and the only possibility is to refine both structures at the same time as performed by the program *SHELX* (Sheldrick & Schneider, 1997). The number of parameters to refine is not increased because the symmetry between the two crystals is imposed during the refinement, but (3) divides the number of independent reflections by two. The ratio of the number of observations to the number of parameters is therefore reduced by a factor of two compared with the non-twinned case. For small-molecule crystals diffracting to very high resolution, this ratio still remains acceptable for refinement. For macromolecules, the ratio is usually limiting. This is especially critical when the solvent content is low and there-

fore the number of atoms in a unit cell is relatively high. If the twinning is not perfect, *i.e.* the twinning ratio is smaller than 50%, (1) and (2) can be inverted and the intensities for reflections $\mathbf{h}_1$ and $\mathbf{h}_2$ diffracted by one domain are directly extractable,

$$I^U(\mathbf{h}_1) = [(1-x)I^T(\mathbf{h}_1) - xI^T(\mathbf{h}_2)]/(1-2x), \quad (4)$$

$$I^U(\mathbf{h}_2) = -[xI^T(\mathbf{h}_1) - (1-x)I^T(\mathbf{h}_2)]/(1-2x). \quad (5)$$

However, the errors in the intensities are summed during this calculation and increase dramatically when x approaches 0.5,

$$\sigma^U(\mathbf{h}_1) = [(1-x)\sigma^T(\mathbf{h}_1) + x\sigma^T(\mathbf{h}_2)]/(1-2x), \quad (6)$$

$$\sigma^U(\mathbf{h}_2) = [x\sigma^T(\mathbf{h}_1) + (1-x)\sigma^T(\mathbf{h}_2)]/(1-2x). \quad (7)$$

These equations can be applied provided that the twinning ratio is known and the structure can be refined in a similar way as with non-twinned data. The ratio of observed reflections to refined parameters is the same as for a non-twinned crystal.

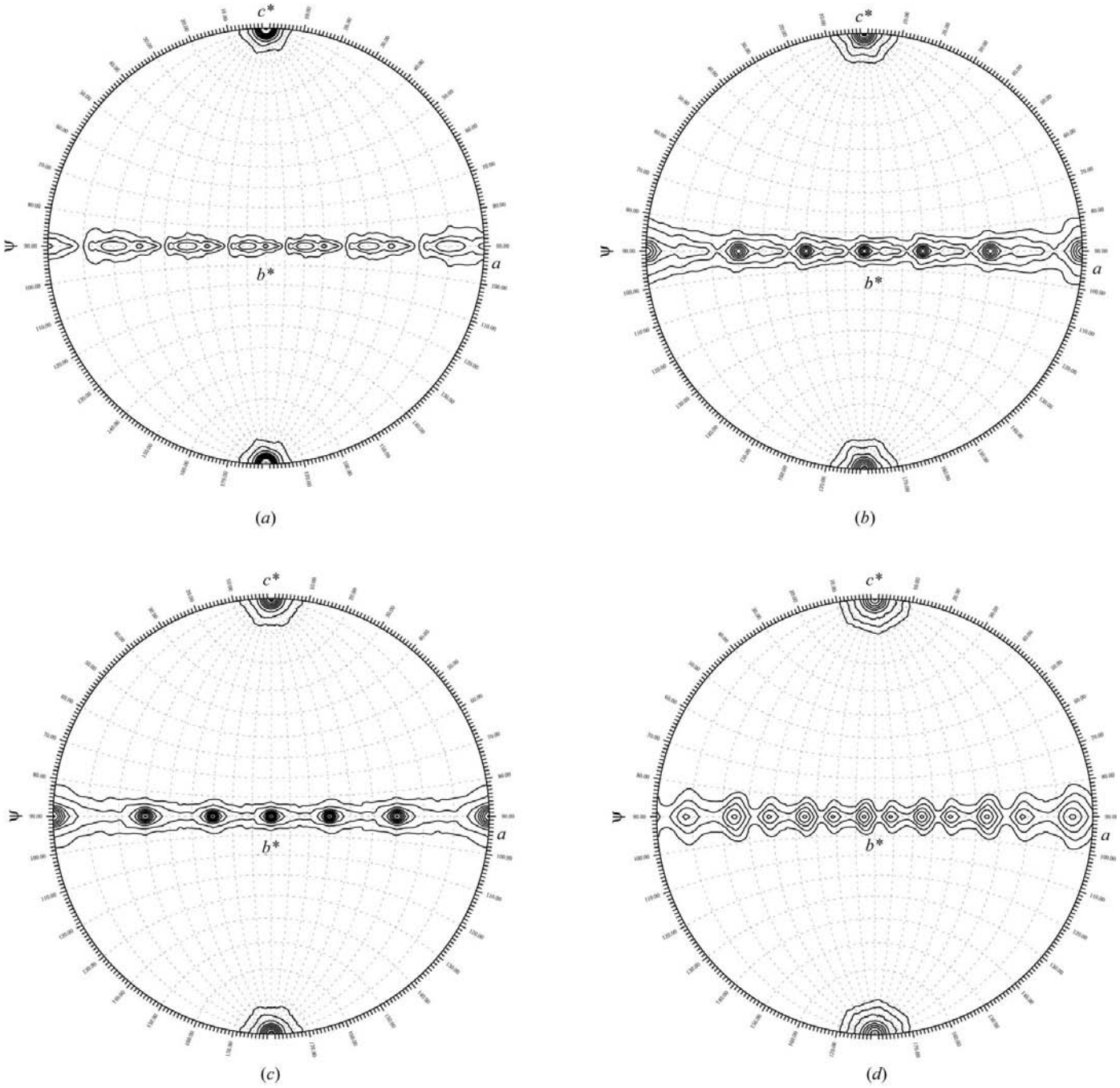

Figure 1
Self-rotation functions. (*a*), (*b*) and (*c*) represent the SRFs for data sets 1, 2 and 3 of BR, respectively. (*d*) is obtained from calculated data using the BR model determined from electron microscopy.

Additionally, the refinement can be performed with the programs commonly used in protein crystallography and therefore the results (*R* factors, electron-density maps *etc.*) are easy to compare with those obtained with other structures.

2.2. Twinning detection: self-rotation function and R_{sym} values

Twinning can be detected from a self-rotation function (SRF), similar to the determination of non-crystallographic symmetries. The tight packing observed in space group $P6_3$ for BR or $P3_1$ for p67^{phox} does not allow an additional internal crystallographic or non-crystallographic symmetry and additional peaks in the SRF can only be explained by the presence of two (or more) symmetry-related domains within the crystal. The relative height of the extra peak compared with the peak related to crystallographic symmetry is indicative of the importance of the twinning. Three data sets are analysed for BR (sets 1–3) and for p67^{phox} (sets 4–6) and the data-collection statistics are summarized in Table 1. Experimental details are reported in Belrhali *et al.* (1999) for BR and in Grizot *et al.* (2001) for p67^{phox}. The SRFs were calculated with the program *GLRF* (Tong & Rossmann, 1990).

Table 1
Data-collection statistics.

	BR			p67^{phox}		
	Data set 1	Data set 2	Data set 3	Data set 4	Data set 5	Data set 6
Space group	$P6_3$	$P6_3$	$P6_3$	$P3_1$	$P3_1$	$P3_1$
Unit-cell parameters (Å)	a = 60.8, c = 110.5	a = 61.1, c = 110.6	a = 61.0, c = 110.3	a = 67.7, c = 49.9	a = 67.5, c = 49.9	a = 67.6, c = 50.3
Resolution range (Å)	38.0–1.95	38.0–2.10	38.0–2.0	30.0–2.05	30.0–2.0	30.0–2.8
R_{sym} (%)	4.6	5.5	7.8	6.4	5.6	7.5
Unique reflections	17996	13653	15474	15011	16760	6319
Completeness (%)	99.5	99.6	98.5	99.7	99.0	99.9
Redundancy	5.4	4.9	6.6	2.3	3.6	3.7

2.2.1. Bacteriorhodopsin. The self-rotation functions are shown in Fig. 1. They indicate an additional twofold symmetry axis parallel to the *a* axis for data sets 2 and 3 (Figs. 1*b* and 1*c*). This corresponds to the existence of two domains for which the Bragg spots [*hkl*] of domain 1 superimpose with spots [$kh\bar{l}$] of domain 2. The relative ratio of the crystallographic twofold peak to the peak induced by the twinning varies from 35 to 97% for the three data sets, indicating a broad range of twinning (Figs. 1*a*, 1*b* and 1*c*). Surprisingly, the SRF on simulated data, based on the electron-microscopy model (Grigorieff *et al.*, 1996) positioned in the three-dimensional crystals, shows a twofold symmetry axis in addition to the crystallographic *c* axis (Fig. 1*d*). However, this peak is of low intensity and located in the *ab* plane, a few degrees off the *a* axis. In the absence of any twinning, this additional peak probably results from an inherent symmetry of BR induced by the helices which are perpendicular to the *ab* plane. From simulated data with increasing twinning ratios, the extra peak increases in intensity and shifts progressively toward the *a* axis (data not shown). Initially, we used the intensity ratio between the crystallographic and this extra peak from the calculated SRF as a calibration in order to determine the twinning ratio of experimental data. However, the results obtained using the SRF were shown to be biased by strong diffraction spots and/or by poor completeness.

2.2.2. p67^{phox}. In addition to the crystallographic threefold axis, the SRF indicates a twofold axis parallel to the *c* axis (data not shown). The Bragg spots [*hkl*] of domain 1 superimpose with spots [$\bar{h}\bar{k}l$] of domain 2. The additional twofold symmetry peak displayed on experimental SRFs ranges from 0 to 100% of the height of the crystallographic threefold peak.

Table 2
Twinning ratio (%) estimated from Yeates statistics and Britton plots.

	BR			p67^{phox}		
	Data set 1	Data set 2	Data set 3	Data set 4	Data set 5	Data set 6
Britton plot	0	25	43	5	21	43
Yeates statistics	5	25	43	9	21	44
$\langle H \rangle$	4	24	42	8	21	43
$\langle H^2 \rangle$	4	24	41	7	21	42

In the absence of any model, the twinning ratio cannot be estimated from the SRFs, but it is clear that the twinning ranges from almost 0% for data set 4 to almost 50% for data set 6, while data set 5 displays an intermediate twinning ratio.

The calculation of R_{sym} in a higher symmetry space group is also indicative of twinning. For BR, the values of R_{sym} in space group *P*622 are 17, 12 and 5.3% for data sets 1, 2 and 3 respectively, while for p67^{phox} the R_{sym} values calculated in space group *P*6 are 26, 24 and 8.1% for data sets 4, 5 and 6, respectively. These values are comparable to the R_{sym} values in the correct space group (Table 1) for a twinning ratio higher than 40%. However, for an intermediate twinning value the R_{sym} in the higher space group does not necessarily indicate twinning, as seen for data set 5.

2.3. Twinning ratio determination: Yeates statistics and Britton plot

After scaling and merging the intensities (Collaborative Computational Project, Number 4, 1994), a small number of negative intensities resulting from the background correction are still present. If the twinning is overestimated, correction of the diffraction intensities according to (4) and (5) leads to an increased number of negative intensities, as shown in the Britton plots in Fig. 2 (Fisher & Sweet, 1980). The value for which the number of negative intensities increases drastically gives an estimation of the twinning ratio. The values estimated for data sets 1 to 6 are given in Table 2. Yeates (1988, 1997) developed a statistical approach based on the cumulative distribution of *H*, where *H* is defined as

$$H = [I^T(\mathbf{h_1}) - I^T(\mathbf{h}_2)]/[I^T(\mathbf{h}_1) + I^T(\mathbf{h}_2)]. \quad (8)$$

For acentric reflections, the theoretical cumulative distribution is

$$S(H) = H/(1 - 2x). \quad (9)$$

Graphic representation of the experimental $S(H)$ compared with theoretical distributions obtained for different twinning ratios is shown in Fig. 3. For BR it shows 5, 25 and 43% for data sets 1, 2 and 3, respectively, and for p67^{phox} it shows 9, 21 and 44% for data sets 4, 5 and 6, respectively. The twinning ratio can also be derived from the mean value or from the mean-square value of H which, for acentric reflections, are given theoretically as

$$\langle H \rangle = 0.5 - x \quad (10)$$

and

$$\langle H^2 \rangle = (1 - 2x)^2/3. \quad (11)$$

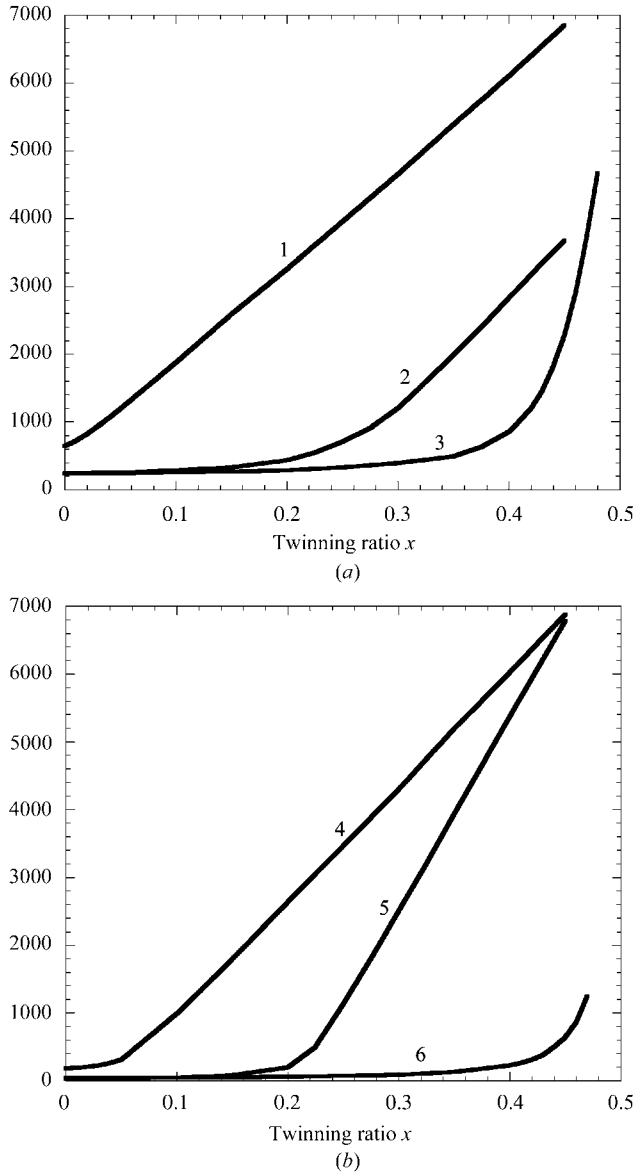

Figure 2
Britton plots. The number of negative intensities as a function of twinning ratio for (*a*) data sets 1, 2 and 3 of BR and (*b*) data sets 4, 5 and 6 of p67^{phox}.

The twinning ratios derived from the Britton plots, from the graphic representation, as well as from the mean values and mean-square values of H of the Yeates statistics are summarized in Table 2. They are consistent for the three data sets for both proteins within a few percent. The determination of the twinning ratio using different approaches prevents the obtaining of values that could be biased by artefacts such as incomplete data.

2.4. Intensity distribution: second-, third- and fourth-order moments

Diffraction data obey the Wilson statistics and these were shown to be affected by twinning. Stanley (1972) derived the intensity distributions for x = 0 (no twinning) and x = 1/2 (perfect twinning). The moments of the distribution are usually calculated from the second to the fourth order with the

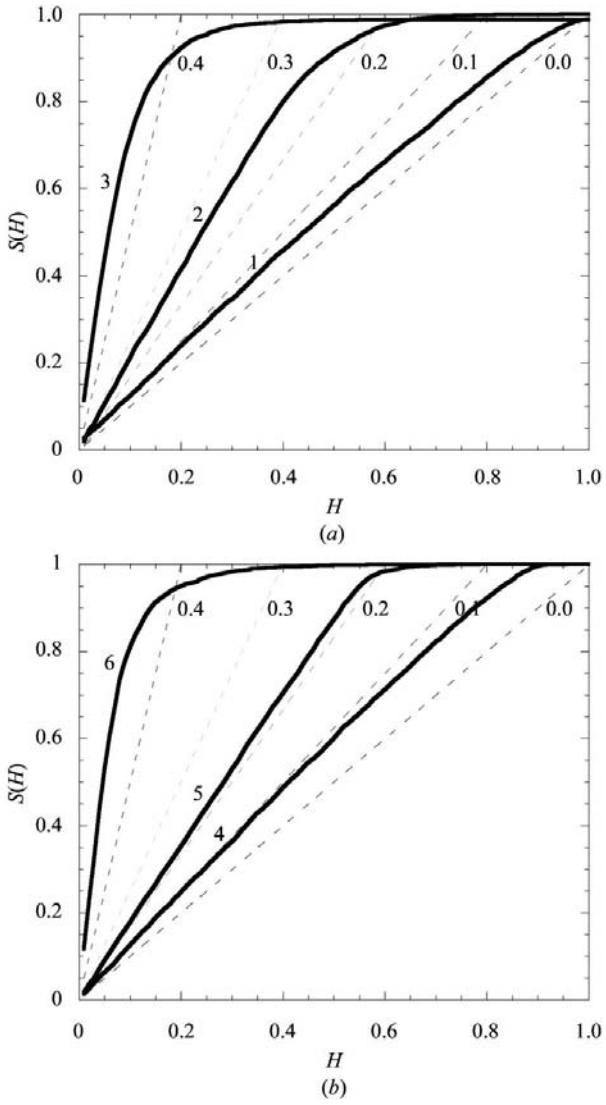

Figure 3
Cumulative distributions $S(H)$. The dashed lines represent the theoretical cumulative distribution of H for twinning ratios of 0, 0.1, 0.2, 0.3 and 0.4, respectively. (*a*) shows the distributions for data sets 1, 2 and 3 of BR. (*b*) shows the distributions for data sets 4, 5 and 6 of p67^{phox}.

usual analysis programs (*TRUNCATE*; Collaborative Computational Project, Number 4, 1994), and the experimental values can be compared with the expected theoretical values. As described earlier, only the second, third and fourth moments are of interest, as higher order moments are too sensitive to experimental errors (Giacovazzo, 1980*a*). Twinning will result in shrinking the distribution. The distributions of centric and acentric reflections are examined separately. In our case, because of the small number of centric reflections, only the acentric reflections were examined. The theoretical moment of order n of the intensity distribution as a function of the twinning ratio is (*Appendix A*)

$$M_n(x) = n![(1-x)^{n+1} - x^{n+1}]/(1-2x). \qquad (12)$$

The second-, third- and fourth-order moments are shown in Fig. 4. For p67^{phox}, M_2, M_3 and M_4 (Table 3) indicate twinning ratios of 2–4, 18–19% and 38–40% for data sets 4, 5 and 6, respectively. These values are consistent with those obtained from the Britton plot or the Yeates statistics (Table 2). After correcting from twinning, the values agree with a normal distribution for data sets 4 and 5. The deviations observed for data set 6 denote the difficulty of correcting data that are highly twinned. In the case of BR, the moments are larger than the theoretical ones. In order to exclude the influence of data treatment, particularly the intensity distribution, which might be affected by the way that weak reflections are integrated, BR data were integrated both with *DENZO* (Otwinowski & Minor, 1997) and *XDS* (Kabsch, 1993). Although these programs deal with the negative reflections in a different manner, the same distributions were obtained from both programs. In certain cases, pseudosymmetries or non-crystallographic symmetries were also reported to affect intensity distributions of some reciprocal zones and to induce centrosymmetric type distributions for non-centrosymmetric reflections (Giacovazzo, 1980*b*). Indeed, self-rotation functions on BR data from non-twinned crystals, in addition to the crystallographic symmetry-induced peaks, exhibit peaks that represent a twofold axis in the *ab* plane a few degrees off the *a* axis (Fig. 1*a*). This pseudo-symmetry can be explained by an intrinsic symmetry of the molecule at low and medium resolution. Another contribution to the deviation of M_n *versus* the normal values may arise from anisotropic diffraction. Although BR diffraction is not strongly anisotropic, this hypothesis can be tested by simulating anisotropy on calculated data. We were not able to ascertain the values of the two contributions, but we used the relative M_n values prior and after twinning correction to quantify the twinning ratio. The moments obtained after correcting the data from the twinning are still higher than the values for the theoretical distribution, but the corrected M_2 and M_3 are comparable for the three data sets. Knowing the deviations of M_2 and M_3 from the theore-

Table 3
Intensity distribution: second-, third- and fourth-order moments.

The second-, third- and fourth-order moments of the intensity distribution are calculated for acentric reflections (program *TRUNCATE*; Collaborative Computational Project, Number 4, 1994). The values obtained after correcting the data from 4, 24 and 42% twinning for sets 1, 2 and 3 of BR, respectively, and 4, 21 and 44% twinning for sets 4, 5 and 6 of p67^{phox}, respectively, are given in parentheses. The theoretical values for M_2, M_3 and M_4 are 2, 6 and 24, respectively.

	BR			p67^{phox}		
	Data set 1	Data set 2	Data set 3	Data set 4	Data set 5	Data set 6
M_2	2.2 (2.3)	2.0 (2.3)	1.8 (2.1)	2.0 (2.0)	1.7 (2.0)	1.5 (1.9)
M_3	7.8 (8.4)	6.2 (8.7)	4.5 (6.6)	5.8 (6.3)	4.2 (5.1)	3.2 (5.9)
M_4	37.9 (41.8)	29.3 (47.4)	15.3 (28.1)	23.6 (26.5)	13.7 (27.3)	8.7 (37.2)

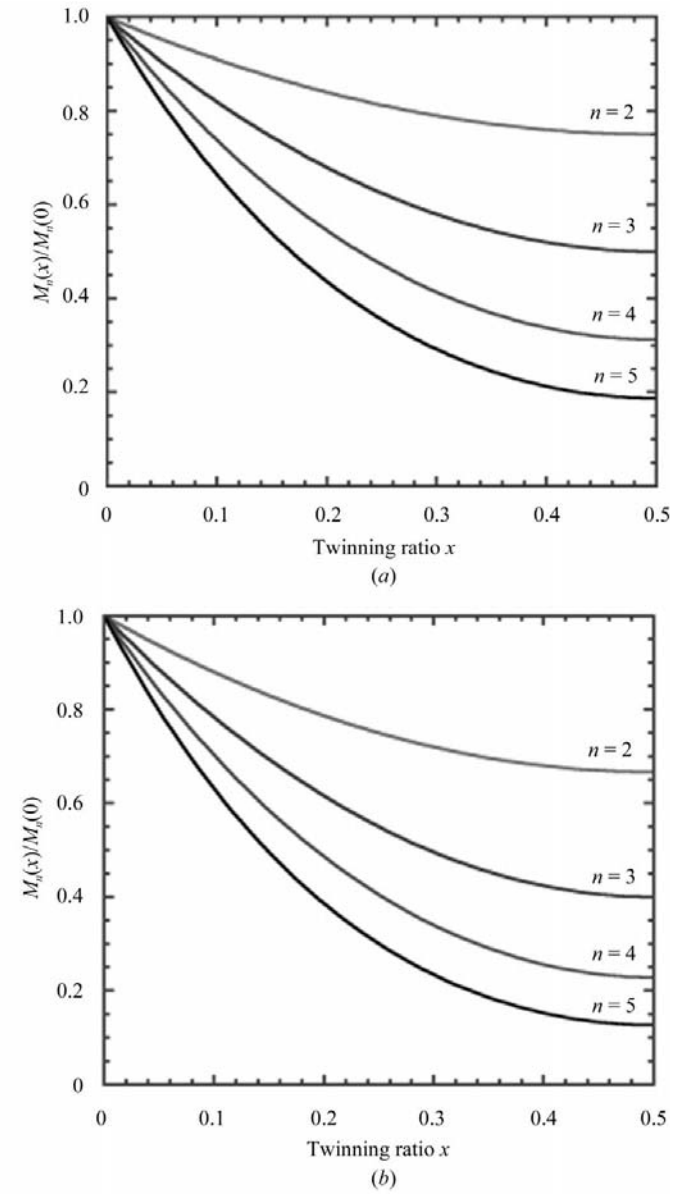

Figure 4
Normalized theoretical second-, third-, fourth- and fifth-order moment of intensity distribution as a function of twinning ratio. (*a*) Acentric, (*b*) centric. The values of Table 3 can be compared with the curves by multiplying M_n by $n!$.

tical values, we were able to determine almost on-line the twinning ratio and to select crystals with low twinning ratios.

3. Conclusions

We combined several methods in order to determine the presence and quantify the ratio of twinning in crystals of BR and p67^{phox}. In addition to known methods such as Yeates statistics or Wilson plots, we derived the M_n as a function of twinning ratios and show how the values deviate from the theoretical values. We analysed the experimental M_n before and after twinning corrections and found that in the case of BR crystals the absolute values of the moments depend not only on twinning but also on pseudo-symmetries, non-crystallographic symmetries or anisotropic diffraction. BR data from almost perfectly twinned crystals have a M_2 value larger than 1.5. As twinning and pseudo-symmetry induce deviations in opposite directions, M_2 alone, without comparison, is not sufficient to detect twinning. In conclusion, the various methods are consistent and deviations for one of them could highlight pseudo- or non-crystallographic symmetry and should be analysed carefully. As the twinning ratios may vary from 0 to 0.5 among crystals grown under the same conditions, twinning should be analysed for various data sets. Inspection of second-, third- and fourth-order moments of the intensity distributions is a very rapid and efficient way of screening for non-twinned (or slightly twinned) crystals almost on-line during data collection.

APPENDIX *A*

A1. Derivation of the moments of the intensity distribution for twinned crystals

Let us consider a twinned crystal composed of a fraction x_1 of one domain and a fraction x_2 of a second domain ($x_1 + x_2 = 1$). For a given reflection $\mathbf{h}$, the intensity from the twinned crystal, $I^T(\mathbf{h})$, is the linear combination $x_1 I^U(\mathbf{h}_1) + x_2 I^U(\mathbf{h}_2)$, where $I^U(\mathbf{h}_j)$ denotes the intensity of reflection $\mathbf{h}_j$ from the non-twinned crystal and the relationship between $\mathbf{h}$, $\mathbf{h}_1$ and $\mathbf{h}_2$ is known from the twin law.

For simplicity, the relationship $I^T(\mathbf{h}) = x_1 I^U(\mathbf{h}_1) + x_2 I^U(\mathbf{h}_2)$ will be noted $I = x_1 I_1 + x_2 I_2$.

The moment of order n, $\langle I^n \rangle = \int_R I^n p(I)\,\mathrm{d}I$, of the probability distribution $p(I)$ can be derived from the joint probability distribution $p(I_1, I_2)$ of I_1 and I_2, since

$$p(I) = \int_{R^2} p(I_1, I_2)\delta(I - x_1 I_1 - x_2 I_2)\,\mathrm{d}I_1\,\mathrm{d}I_2,$$

where $\delta(x)$ is the Dirac distribution of x and thus

$$\begin{aligned}\langle I^n \rangle &= \int_{R^3} I^n p(I_1, I_2)\delta(I - x_1 I_1 - x_2 I_2)\,\mathrm{d}I_1\,\mathrm{d}I_2\,\mathrm{d}I \\ &= \int_{R^2} (x_1 I_1 + x_2 I_2)^n p(I_1, I_2)\,\mathrm{d}I_1\,\mathrm{d}I_2.\end{aligned}$$

If the probability distributions $p_1(I_1)$ and $p_2(I_2)$ of I_1 and I_2 are independent, then $p(I_1, I_2) = P_1(I_1)P_2(I_2)$ and

$$\begin{aligned}\langle I^n \rangle &= \int_{R^2} (x_1 I_1 + x_2 I_2)^n p_1(I_1)p_2(I_2)\,\mathrm{d}I_1\,\mathrm{d}I_2 \\ &= n! \sum_{p=0}^{n} \frac{x_1^p x_2^{n-p} \langle I_1^p \rangle \langle I_2^{n-p} \rangle}{p!(n-p)!}.\end{aligned}$$

A2. Acentric reflections

The Wilson distribution of the intensity, I_j, for an acentric reflection is

$$p(I_j) = \frac{H(I_j)}{\langle I \rangle} \exp\left(-\frac{I_j}{\langle I \rangle}\right),$$

where $H(x)$ is the Heaviside distribution of x defined as

$$\begin{cases} x < 0 \Leftrightarrow H(x) = 0 \\ x > 0 \Leftrightarrow H(x) = 1 \end{cases}$$

and $\langle I \rangle$ is the mean intensity in a given resolution shell.

From this probability distribution, it can readily be shown that

$$\langle I_j^q \rangle = \langle I \rangle^q \int_0^\infty t^q \exp(-t)\,\mathrm{d}t = q! \langle I \rangle^q.$$

Therefore, if we consider the probability distributions of I_1 and I_2 as independent,

$$\langle I^n \rangle = n! \langle I \rangle^n \sum_{p=0}^{n} x_1^p x_2^{n-p} = n! \langle I \rangle^n \frac{x_2^{n+1} - x_1^{n+1}}{x_2 - x_1}.$$

Using x to denote the twin fraction of one of the crystals and taking into account the relationship $x_1 + x_2 = 1$ one derives

$$M_n(x) = \frac{\langle I^n \rangle}{\langle I \rangle^n} = n! \frac{(1-x)^{n+1} - x^{n+1}}{1 - 2x}.$$

For $x = 0$ or $x = 1$ (non-twinned crystal), the standard relationship for acentric reflections $M_n(0) = M_n(1) = n!$ is obtained.

For $x \rightarrow \frac{1}{2}$ (perfectly twinned crystal), one obtains

$$\begin{aligned}M_n\left(\tfrac{1}{2}\right) &= n! \lim_{\varepsilon \rightarrow 0} \frac{\left(\frac{1}{2} + \varepsilon\right)^{n+1} - \left(\frac{1}{2} - \varepsilon\right)^{n+1}}{2\varepsilon} \\ &= n!\,(x^{n+1})'_{(x = \frac{1}{2})} = \frac{(n+1)!}{2^n}.\end{aligned}$$

The variation *versus* x of $M_n(x)/M_n(0)$ for the first values of n is shown in Fig. 4, from which it can be seen that the normalized moments $M_n(x)$ of the intensity distribution of acentric reflections from a twinned crystal are always smaller than the corresponding normalized moments for an non-twinned crystal.

A3. Centric reflections

For a centric reflection, the Wilson distribution of the intensity I_j is

$$p(I_j) = \frac{H(I_j)}{(2\pi\varepsilon \langle I/\varepsilon \rangle I_j)^{1/2}} \exp\left(-\frac{I_j/\varepsilon}{2\langle I/\varepsilon \rangle}\right),$$

where ε is the multiplicity of the reflection and $\langle I/\varepsilon \rangle$ is the mean value of I/ε in a given resolution shell.

From this probability distribution, it can readily be shown that

$$\begin{aligned}\langle (I_j/\varepsilon)^q \rangle &= \langle I/\varepsilon \rangle^q \frac{2^q}{\pi^{1/2}} \int_0^\infty t^{q-1/2} \exp(-t)\, \mathrm{d}t \\ &= \frac{2^q}{\pi^{1/2}} \Gamma(q+\tfrac{1}{2}) \langle I/\varepsilon \rangle^q \\ &= (2q-1)!! \langle I/\varepsilon \rangle^q \\ &= \frac{(2q)!}{2^q q!} \langle I/\varepsilon \rangle^q.\end{aligned}$$

Therefore, if we consider the probability distributions of I_1 and I_2 as independent,

$$\langle (I/\varepsilon)^n \rangle = \frac{n! \langle I/\varepsilon \rangle^n}{2^n} \sum_{p=0}^{n} \frac{(2p)![2(n-p)]!}{(p!)^2[(n-p)!]^2} x_1^p x_2^{n-p}.$$

Using x to denote the twin fraction for one crystal and taking into account the relationship $x_1 + x_2 = 1$, one derives

$$\begin{aligned}M_n(x) &= \langle (I/\varepsilon)^n \rangle / \langle I/\varepsilon \rangle^n \\ &= \frac{n!}{2^n} \sum_{p=0}^{n} \frac{(2p)![2(n-p)]!}{(p!)^2[(n-p)!]^2} x^p (1-x)^{n-p}.\end{aligned}$$

For $x = 0$ or $x = 1$, one obtains the standard relationship for centric reflections from a non-twinned crystal,

$$M_n(0) = M_n(1) = (2n-1)!! = \frac{(2n)!}{2^n n!}.$$

For $x = \frac{1}{2}$ one obtains

$$M_n(\tfrac{1}{2}) = \frac{n!}{2^{2n}} \sum_{p=0}^{n} \frac{(2p)![2(n-p)]!}{(p!)^2[(n-p)!]^2} = n!,$$

which shows that the intensity distribution of centric reflections for a perfectly twinned crystal is similar to the intensity distribution of acentric reflections for an non-twinned crystal.

The variation of $M_n(x)/M_n(0)$ *versus* x for the first values of n is shown in Fig. 4.

We thank R. Henderson, K. Edman, J. Hajdu, R. Neutze and J. Navaza for helpful discussions and L. Nauton for technical assistance. Diffraction experiments were performed at ESRF (Grenoble) on beamlines ID14, ID29 and BM30A.

References

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B. & Steitz, T. A. (1999). *Nature (London)*, **400**, 841–847.
Belrhali, H., Nollert, P., Royant, A., Menzel, C., Rosenbusch, J. P., Landau, E. M. & Pebay-Peyroula, E. (1999). *Structure Fold. Des.* **7**, 909–917.
Breyer, W. A., Kingston, R. L., Anderson, B. F. & Baker, E. N. (1999). *Acta Cryst.* D**55**, 129–138.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chandra, N., Acharya, K. R. & Moody, P. C. E. (1999). *Acta Cryst.* D**55**, 1750–1758.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Edman, K., Nollert, P., Royant, A., Belrhali, H., Pebay-Peyroula, E., Hajdu, J., Neutze, R. & Landau, E. M. (1999). *Nature (London)*, **401**, 822–826.
Fisher, R. G. & Sweet, R. M. (1980). *Acta Cryst.* A**36**, 755–760.
Frazão, C., Sieker, L., Coelho, R., Morais, J., Pacheco, I., Chen, L., LeGall, J., Dauter, Z., Wilson, K. & Carrondo, M. A. (1999). *Acta Cryst.* D**55**, 1465–1467.
Giacovazzo, C. (1980*a*). *Direct Methods in Crystallography*, pp. 28–32. London: Academic Press.
Giacovazzo, C. (1980*b*). *Direct Methods in Crystallography*, pp. 32–38. London: Academic Press.
Gomis-Rüth, F.-X., Fita, I., Kiefersauer, R., Huber, R., Avilés, F. X. & Navaza, J. (1995). *Acta Cryst.* D**51**, 819–823.
Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. & Henderson, R. (1996). *J. Mol. Biol.* **259**, 393–421.
Grizot, S., Fieschi, F., Dagher, M.-C. & Pebay-Peyroula, E. (2001). *J. Biol. Chem.* **276**, 21627–21631.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Landau, E. M. & Rosenbusch, J. P. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 14532–14535.
Luecke, H., Richter, H.-T. & Lanyi, J. K. (1998). *Science*, **280**, 1934–1937.
Luecke, H., Schobert, B., Richter, H.-T., Cartailler, J.-P. & Lanyi, J. (1999). *J. Mol. Biol.* **291**, 899–911.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P. & Landau, E. M. (1997). *Science*, **277**, 1676–1681.
Redinbo, M. R. & Yeates, T. O. (1993). *Acta Cryst.* D**49**, 375–380.
Rees, D. C. (1980). *Acta Cryst.* A**36**, 578–581.
Rees, D. C. (1982). *Acta Cryst.* A**38**, 201–207.
Royant, A., Edman, K., Ursby, T., Pebay-Peyroula, E., Landau, E. M. & Neutze, R. (2000). *Nature (London)*, **406**, 645–648.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Stanley, E. (1972). *J. Appl. Cryst.* **5**, 191–194.
Taylor, H. O., O'Reilly, M., Leslie, A. G. & Rhodes, D. (2000). *J. Mol. Biol.* **303**, 693–707.
Tong, L. & Rossmann, M. G. (1990). *Acta Cryst.* A**46**, 783–792.
Valegard, K., van Scheltinga, A. C., Lloyd, M. D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H. J., Baldwin, J. E., Schofield, C. J., Hajdu, J. & Andersson, I. (1998). *Nature (London)*, **394**, 805–809.
Yang, F., Dauter, Z. & Wlodawer, A. (2000). *Acta Cryst.* D**56**, 959–964.
Yeates, T. O. (1988). *Acta Cryst.* A**44**, 143–144.
Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.