Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Twinned crystals and anomalous phasing

Zbigniew Dauter

Synchrotron Radiation Research Section, National Cancer Institute, Brookhaven National Laboratory, Building 725A-X9, Upton, NY 11973, USA

Correspondence e-mail: dauter@bnl.gov

Received 3 March 2003
Accepted 23 September 2003

Merohedral or pseudomerohedral twinning of crystals cannot be identified from inspection of the diffraction patterns. Several methods for the identification of twinning and the estimation of the twin fraction are suitable for macromolecular crystals and all are based on the statistical properties of the measured diffraction intensities. If the crystal twin fraction is estimated and is not too close to 0.5, the diffraction data can be detwinned; that is, related to the individual crystal specimen. However, the detwinning procedure invariably introduces additional inaccuracies to the estimated intensities, which substantially increase when the twin fraction approaches 0.5. In some cases, a crystal structure can be solved with the original twinned data by standard techniques such as molecular replacement, multiple isomorphous replacement or multiwavelength anomalous diffraction. Test calculations on data collected from a twinned crystal of gpD, the bacteriophage λ capsid protein, show that the single-wavelength anomalous diffraction (SAD) method can be used to solve its structure even if the data set corresponds to a perfectly twinned crystal with a twin fraction of 0.5.

1. Introduction

Twinning occurs when a crystalline specimen consists of multiple domains which are mutually reoriented according to a specific transformation that does not belong to the symmetry operations of the crystal point group but is related in some way to the crystal lattice (Koch, 1992). The splitting or cracking of crystals often encountered by practicing crystallographers should not therefore be described as twinning. The size of the twin domains is large in comparison with the crystal cell dimensions, so that X-rays diffracted by the separate parts do not interfere and the total diffraction corresponds to the sum of intensities originating from individual domains. This is in contrast to disorder in a crystal, where there are differences in the location of the content of the neighbouring unit cells. In the case of disorder (either static or dynamic), diffracted X-rays represent the averaged (spatially or temporally) content of all unit cells.

Twins can be classified according to the mutual disposition of the individual domains. Various kinds of twins have different manifestations in reciprocal space, which has important consequences for the process of structure solution of twinned crystals. Donnay & Donnay (1974) differentiate between twins with twin-lattice symmetry, characterized by a single reciprocal lattice with all reflections perfectly overlapped, and twins with twin-lattice quasi-symmetry, with some

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

reflections originating from individual separated domains. According to the terminology introduced by Friedel (1928), they correspond to merohedral or non-merohedral twins, respectively.

Merohedral twins reflect the possibility of different non-equivalent orientations of crystals within a lattice of symmetry higher than the crystal point-group (class) symmetry. This is possible in the tetragonal, trigonal, hexagonal or cubic systems, with lattices always having the highest symmetry of the system (holohedry). If the crystal class is a subgroup of the lattice symmetry, *e.g.* with only a half (hemihedry) or a quarter (tetartohedry) of its symmetry elements, the unit cell of such a crystal transformed by a symmetry operation of the holohedry but not belonging to the crystal class will conform perfectly within the same lattice, but will not be equivalent to the initial orientation. In reciprocal space, such a transformation will lead to the exchange of indices between non-equivalent reflections having different intensities, as in Fig. 1. In general, the relation between different twin domains may involve rotations, mirror reflections or point inversions, and all are observed in small-molecule crystallography and mineralogy. Since proteins, carbohydrates and nucleic acids are chiral, in macromolecular crystallography the only possible twin operations are proper rotations and further discussion will be restricted to twins related by rotations. In Friedel's nomenclature, this corresponds to relations between holoaxial hemihedry (maximum rotational symmetry of the lattice) and holoaxial tetartohedry (a half of the previous symmetry operations or a quarter of the lattice holohedry) or holoaxial ogdohedry (a half of the previous elements). Table 1 lists space groups that can lead to merohedral twinning and specifies the possible twinning operations.

Table 1
Crystal space groups that can be oriented in multiple non-equivalent ways within the crystal lattice.

In each case, the twinning operation and reindexing matrix are given. Other selections of twinning operations are possible; the twofold rotations listed below are the simplest.

Space group	Twin operation	Reindexing	Matrix
$P4$, ($P4_1$, $P4_3$), $P4_2$, $I4$, $I4_1$	Twofold ‖ [110]	$hkl > kh{-}l$	010/100/00−1
$P3$, ($P3_1$, $P3_2$)	Twofold ‖ c	$hkl > {-}h{-}kl$	−100/0−10/001
	or twofold ‖ [110]	$hkl > kh{-}l$	010/100/00−1
	or twofold ‖ [−110]	$hkl > {-}k{-}h{-}l$	0−10/−100/00−1
$R3$	Twofold ‖ [110]	$hkl > kh{-}l$	010/100/00−1
$P321$, ($P3_121$, $P3_221$)	Twofold ‖ c	$hkl > {-}h{-}kl$	−100/0−10/001
$P312$, ($P3_112$, $P3_212$)	Twofold ‖ c	$hkl > {-}h{-}kl$	−100/0−10/001
$P6$, ($P6_1$, $P6_5$), ($P6_2$, $P6_4$), $P6_3$	Twofold ‖ [110]	$hkl > kh{-}l$	010/100/00−1
$P23$, $P2_13$, $I23$, $I2_13$, $F23$	Twofold ‖ [110]	$hkl > kh{-}l$	010/100/00−1

Merohedral twinning cannot be distinguished from the appearance of the diffraction pattern, since reflections from different twin components overlap perfectly as a consequence

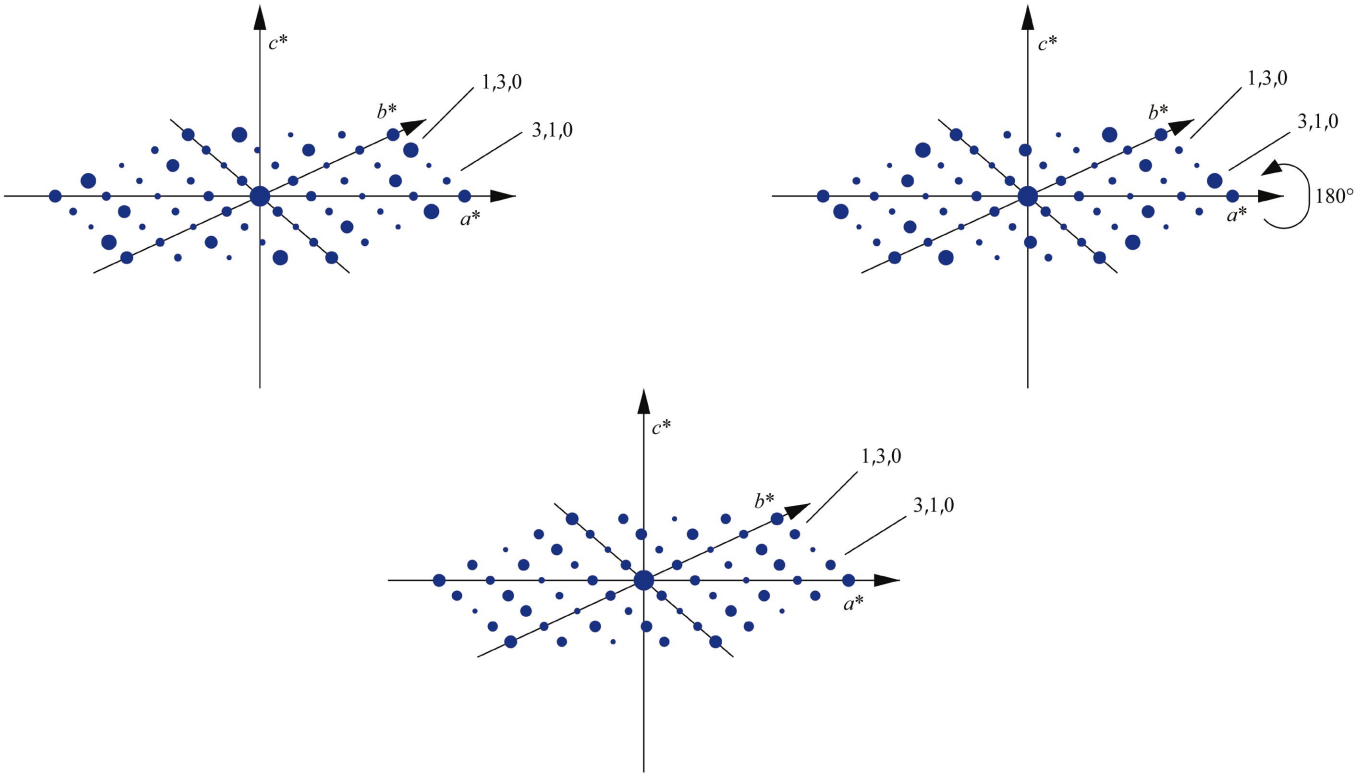

Figure 1
The schematic diffraction pattern of a hexagonal crystal in the $P6$ space group in two possible orientations within its hexagonal lattice of $P622$ (strictly $P6/mmm$) symmetry, related by the twofold rotation around the a^* axis. Since this twofold rotation does not belong to the crystal symmetry, reflections with the same indices have different intensities in the two orientations (top). In the perfect twin, both orientations superimpose, and the diffraction pattern acquires $P622$ symmetry (bottom).

of the identity of the holohedral and merohedral lattice metrics. However, a similar situation may arise when the crystal metric fortuitously corresponds to the higher symmetry system. Such cases are termed pseudomerohedry. Again, the twinning operation is a symmetry element present in the high-symmetry group of the lattice but absent in the proper symmetry class of the crystal. Table 2 gives some examples of pseudomerohedral twinning possibilities. Because the pseudomerohedry results from a serendipitous coincidence of cell dimensions, the overlap of twin lattices may not be ideal and sometimes a broadening or split of reflection profiles can be observed, especially at high resolution. Otherwise,

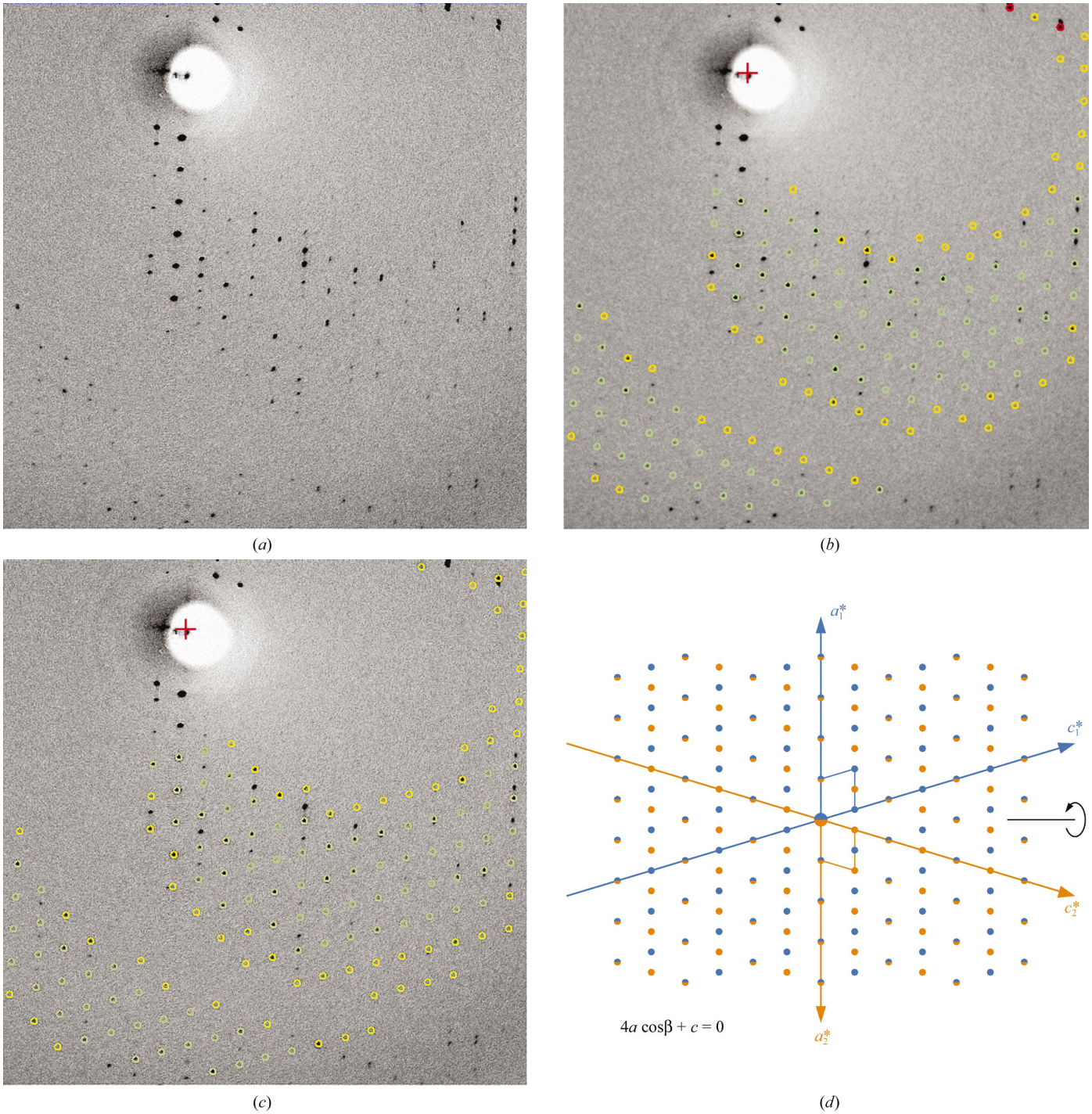

Figure 2
Examples of diffraction images from non-merohedrally twinned crystals. (*a*)–(*d*) A crystal with unit-cell dimensions fulfilling the condition: $4a\cos\beta + c = 0$, which leads to an overlap of reflections on every second line with $l = 2n$: (*a*) the diffraction pattern, (*b*) and (*c*) two possible indexing schemes, (*d*) a scheme explaining the underlying geometry. (*e*)–(*h*) A crystal with an overlap of every fourth line with $l = 4n$ and unit-cell dimensions fulfilling $8a\cos\beta + c = 0$.

pseudomerohedrally twinned crystals have identical characteristics to the classic merohedral twins.

Sometimes only a subset of reflections from two twin domains overlap, which is termed non-merohedral twinning. This type of twinning is characterized by abnormalities in the diffraction pattern, with some planes or lines of reflections in reciprocal lattice having different numbers of reflections to others, which cannot happen in a single lattice (except as a result of absences resulting from screw axes or glide planes). Examples of diffraction images from non-merohedrally twinned crystals are shown in Fig. 2.

If the lattices of twin components fit in less than three dimensions, diffraction images recorded from such epitaxic twins are usually characterized by streaks or overlap of

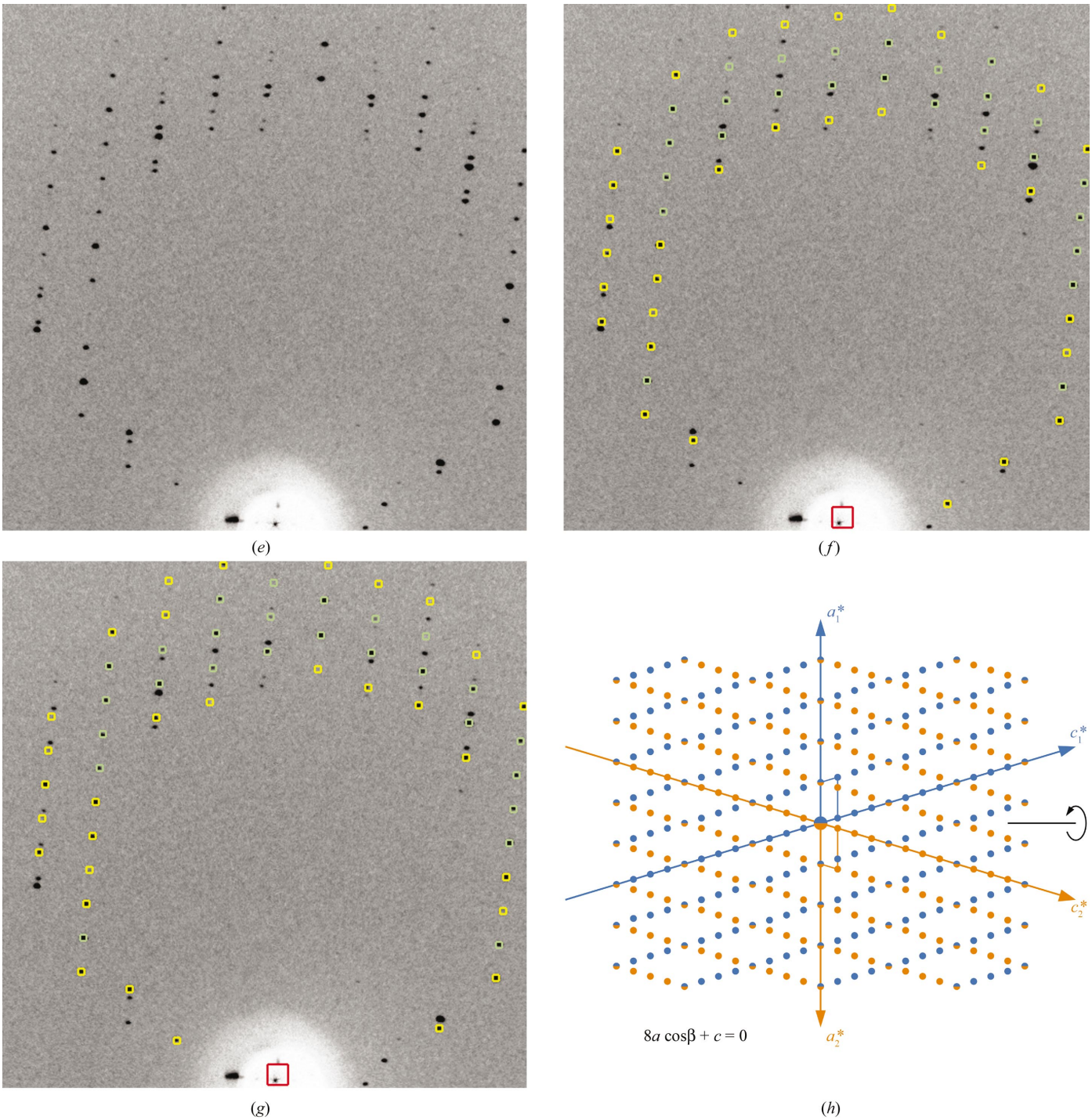

Figure 2 (continued)
Examples of diffraction images from non-merohedrally twinned crystals. (*e*) The diffraction image, (*f*) and (*g*) two possible interpretations, (*h*) the explanation.

Table 2
Examples of pseudomerohedral twinning.

In the rhombohedral cell, R corresponds to the rhombohedral setting with $a = b = c$, $\alpha = \beta = \gamma$ and H to the hexagonal setting with $a = b$ and $\gamma = 120°$.

Crystal symmetry	Cell conditions	Apparent symmetry	Twinning operation	Reindexing
Monoclinic P	$a \simeq c$	Orthorhombic C	Twofold ‖ [101]	$hkl > l{-}kh$
Monoclinic P	$\beta \simeq 90°$	Orthorhombic P	Twofold ‖ a	$hkl > h{-}k{-}l$
Rhombohedral R	$\alpha \simeq 60°$	Cubic F	Twofold ‖ [011]	$hkl > -h, l - h, k - h$
			or twofold ‖ [101]	$hkl > l - k, - k, h - k$
			or twofold ‖ [110]	$hkl > k - l, h - l, -l$
Rhombohedral H	$c \simeq a6^{1/2}$	Cubic F	Twofold ‖ [012]	$hkl > -h, (2h + k + l)/3, (4h + 8k - l)/3$
			or twofold ‖ [−102]	$hkl > (h + 2k - l)/3, -k, (-8h, -4k, -l)/3$
			or twofold ‖ [1−12]	$hkl > (-h - 2k + l)/3, (-2h - k - l)/3, (4h - 4k - l)/3$
Rhombohedral R	$\alpha \simeq 90°$	Cubic P	Twofold ‖ a	$hkl > h{-}k{-}l$
			or twofold ‖ b	$hkl > -hk{-}l$
			or twofold ‖ c	$hkl > -h{-}kl$
Rhombohedral H	$c \simeq a(3/2)^{1/2}$	Cubic P	Twofold ‖ [101]	$hkl > (h + 2k + 2l)/3, -k, (4h + 2k - l)/3$
			or twofold ‖ [0−11]	$hkl > -h, (2h + k - 2l)/3, (-2h - 4k - l)/3$
			or twofold ‖ [−111]	$hkl > (-h - 2k - 2l)/3, (-2h - k + 2l)/3, (-2h + 2k - l)/3$
Rhombohedral R	$\alpha \simeq 109.5°$	Cubic I	Twofold ‖ [11−1]	$hkl > k, h, (-h - k - l)$
			or twofold ‖ [1−11]	$hkl > l, (-h - k - l), h$
			or twofold ‖ [−111]	$hkl > (-h - k - l), l, k$
Rhombohedral H	$c \simeq a(3/8)^{1/2}$	Cubic I	Twofold ‖ [−201]	$hkl > (h + 2k - 4l)/3, -k, (-2h - k - l)/3$
			or twofold ‖ [021]	$hkl > -h, (2h + k + 4l)/3, (h + 2k - l)/3$
			or twofold ‖ [2−21]	$hkl > (-h - 2k + 4l)/3, (-2h - k - 4l)/3, (h - k - l)/3$

reflection profiles in a particular direction. Such diffraction patterns are difficult to interpret and are similar to images from split or highly mosaic crystals.

Further discussion will be restricted to merohedral or pseudomerohedral twins of macromolecules, excluding mirror planes or center inversions as twin operations.

2. Effect of twinning on diffraction intensities

Merohedral or exact pseudomerohedral twinning cannot be recognized by inspection of diffraction images, since the geometric characteristics of the reciprocal lattice do not reveal that it may consist of two perfectly overlapped sets of reflections. In view of this, one of the criteria often given for identification of twins, that 'the metric symmetry of the lattice is higher than the Laue symmetry of the crystal', is not useful at the stage of the initial interpretation of diffraction images from the newly obtained crystals. This is true for any pair of subgroup/supergroup relations, *e.g.* a primitive tetragonal crystal can be indexed as orthorhombic, monoclinic or triclinic.

However, the true symmetry of the crystal depends on its content, not on the geometry of its lattice; therefore, the crystal symmetry should be inferred from the symmetry of the distribution of reflection intensities in reciprocal space. The metric of the reciprocal lattice can exclude some higher symmetry systems, but cannot identify which of the lower symmetry systems (subgroups) is correct. The triclinic crystal may by chance have all cell dimensions almost equal and all angles very close to 90°, yet despite having the cubic lattice metric, the intensities diffracted from this crystal will not merge in any symmetry higher than triclinic.

In the case of merohedral twins, the symmetry of the distribution of reflection intensities in reciprocal space may be misleading. If the original (untwinned) intensities of the individual reflections in a pair related by the twin operation are I_1 and I_2, the intensities measured from a crystal with the twin fraction α will be $J_1 = \alpha I_1 + (1 - \alpha)I_2$ and $J_2 = (1 - \alpha)I_1 + \alpha I_2$. For perfect twins with $\alpha = 0.5$, both intensities are equal and the intensities measured from such a crystal will merge perfectly in the symmetry higher than that of the real crystal point group. If the twin fraction is smaller than 0.5, the data will merge well in the correct point group, but in higher symmetry will give an $R_{\rm merge}$ value lower than 55%, expected for a completely wrong symmetry assignment. Such a behavior of $R_{\rm merge}$ is a simple and typical qualitative warning that a partial merohedral twinning may be present.

The above discussion refers to the distribution of reflection intensities in reciprocal space. Since X-rays diffracted from separate twin domains do not interfere, the intensities of reflections measured from a twinned crystal consist of the sum of the intensities of the individual twin domains, weighted by their relative volumes illuminated by the X-ray beam. If the separate domains had the same orientation, as in the single-crystal specimen, the diffracted X-rays would interfere and the intensity of the diffracted X-rays would correspond to the sum of the structure amplitudes, not the intensities. As a consequence, reflections measured from a twinned crystal display intensity statistics that differ from classic Wilson's theory. This can be explained simply as follows. A set of intensities from a single crystal is characterized by a certain low percentage of very weak reflections and a certain small amount of very strong reflections. In the twinned crystal, reflection intensities are added up in pairs related by the twin law. The probability that the twin operation connects two very weak or two very strong reflections is, therefore, rather low and, as a consequence, in the data set measured from a twinned crystal the amount of extreme, very low or very strong intensities is still

Table 3
Various statistical formulae for untwinned and twinned crystals.

(*a*) Intensity distributions.

Cumulative intensity distributions	
Non-centrosymmetric reflections	
Untwinned	${}_1N(z) = 1 - \exp(-z)$
50% twinned	${}_1N(z, 0.5) = 1 - (1 + 2z)\exp(-2z)$
α twinned	${}_1N(z, \alpha) = (\alpha[\exp(-z/\alpha) - 1] - (1 - \alpha)\{\exp[-z/(1 - \alpha)] - 1\})/(1 - 2\alpha)$
Centrosymmetric reflections	
Untwinned	${}_{-1}N(z) = \mathrm{erf}(z/2)^{1/2}$
50% twinned	${}_{-1}N(z, 0.5) = 1 - \exp(-z)$
Cumulative $S(H)$ distributions	
Non-centrosymmetric reflections	$S(H) = H/(1 - 2\alpha)$
Centrosymmetric reflections	$S(H) = \cos^{-1}[H/(2\alpha - 1)/\pi]$

(*b*) Moments of I and H.

	$\langle I^2\rangle/\langle I\rangle^2$	$\langle F\rangle^2/\langle I\rangle$ (Wilson ratio)	$\langle H\rangle$	$\langle H^2\rangle$
Non-centrosymmetric			$\frac{1}{2} - \alpha$	$(1 - 2\alpha)^2/3$
Untwinned	2.0	0.785 (= $\pi/4$)	0.5	0.333
50% twinned	1.5	0.885	0.0	0.0
Centrosymmetric			$2(1 - 2\alpha)/\pi$	$(1 - 2\alpha)^2/2$
Untwinned	3.0	0.637 (= $2/\pi$)	0.637	0.5
50% twinned	2.0	0.785	0.0	0.0

(*c*) Moments of L and cumulative $N(|L|)$ distributions.

	$\langle\|L\|\rangle$	$\langle L^2\rangle$	$N(L)$	$N(\|L\|)$
Non-centrosymmetric				
Untwinned	1/2	1/3	$(L + 1)/2$	$\|L\|$
Twinned	3/8	1/5	$(L + 1)^2(2 - L)/4$	$\|L\|(3 - L^2)/2$
Centrosymmetric untwinned	$2/\pi$	1/2	$\arccos(-L)/\pi$	$(2/\pi)\arcsin(\|L\|)$

smaller than in the case of a single crystal. The intensities from twinned crystals have therefore a sharper distribution around the mean value than that resulting from Wilson statistics.

As pointed out by Rees (1980), the structural disorder leading to additional local symmetry has an effect opposite to twinning, with the intensity distribution flatter than in Wilson's distribution.

The theory of intensity statistics for merohedrally twinned crystals was worked out by Stanley (1955, 1972) and Rees (1980, 1982). For very convenient practical comparisons, one may use the cumulative intensity distribution $N(z)$, giving the fraction of all reflections with intensities below the limit z (Howells *et al.*, 1950). The formulae for untwinned and perfectly twinned centric and non-centric reflections are shown in Table 3. As can be seen, the distribution for a perfectly twinned centrosymmetric crystal is the same as for an untwinned non-centrosymmetric crystal. Therefore, in the calculation of intensity statistics, the centrosymmetric reflections have to be treated separately, since in the high-symmetry cubic, hexagonal or tetragonal space groups, particularly at low resolution, the fraction of centrosymmetric reflections is significant.

The higher moments of the intensity distribution provide another global distinguishing criterion (Stanley, 1972). They have been routinely used for confirming the presence or absence of the center of symmetry in small structure crystals and are also distinctive for twinned crystals, as indicated in Table 3.

3. Detection of (pseudo)merohedral twinning

Several tests based on various properties of the intensity distributions have been proposed and used in practice in small-molecule and macromolecular crystallography. They will be illustrated here with data collected from the pseudo-merohedrally twinned $P2_1$ crystal of gpD, the capsid-stabilizing protein of bacteriophage λ (Yang, Dauter *et al.*, 2000; Yang, Forrer *et al.*, 2000), and from merohedrally twinned $P4_3$ crystals of interleukin-1β (Rudolph *et al.*, 2003).

3.1. Packing considerations

The presence of twinning may sometimes be concluded from crystal-packing considerations. Even if the diffraction data merge well in the high-symmetry space group, there may not be enough space in the unit cell to accommodate the required number of molecules. An unrealistically low Matthews number, V_M, and solvent content (Matthews, 1968) in high-symmetry crystals suggest possible twinning. IL-1β crystallizes with unit-cell parameters $a = b = 53.9$, $c = 77.4$ Å. In space group $P4_3$ with four molecules in the unit cell, it gives a Matthews coefficient V_M of 2.9 Å^3 Da^{-1} and a solvent content of 58% (Rudolph *et al.*, 2003). Even if this crystal was perfectly twinned, the space group $P4_322$ could not be accepted, because eight molecules in the unit cell would require the V_M value to be 1.45 Å^3 Da^{-1}, well below the acceptable limit of 1.75 Å^3 Da^{-1}.

3.2. Moments of intensities distribution

As mentioned above, one of the consequences of twinning is the smaller fraction of very weak as well as very strong intensities in the entire population of reflections. This is analogous to the difference between diffraction patterns of centrosymmetric and non-centrosymmetric crystals. Experienced small-structure crystallographers are able to differentiate between the presence or absence of the center of symmetry by looking at a single precession or Weissenberg photograph, from the amount of reflections that are unobservable by eye (*i.e.* very weak). In these terms, the diffraction intensities from twinned crystals appear hypo-noncentric, with less variation of their values than the average. This is understandable, since twinning results in the averaging of unrelated intensities in pairs. In a quantitative representation, the variance (and higher moments) of the intensity distribution is smaller for twinned crystals than for single specimens. The Wilson (1949) ratio $\langle F\rangle^2/\langle I\rangle$ is expected to be 0.785 for normal non-centrosymmetric reflections and 0.885 for such reflections from twinned crystals, as shown in Table 3 (Stanley, 1972). Fig. 3 shows these values calculated in resolution ranges from gpD and IL-1β data and, as a control, the obviously untwinned accurate data from lysozyme (Dauter *et al.*, 1999). The points shown in the figure are considerably

scattered and although on average they suggest that the crystal of lysozyme was not twinned and that those of gpD and IL-1β were considerably twinned, it is difficult to satisfactorily estimate the degree of twinning from these graphs.

3.3. Britton plot of 'negative intensities'

As pointed out above, the intensities measured from a twinned crystal are combined from two twin domains with weights proportional to their volumes,

$$I_1 = \alpha J_1 + (1-\alpha)J_2,$$
$$I_2 = (1-\alpha)J_1 + \alpha J_2,$$

where I_1 and I_2 are the measured intensities of a pair of reflections related by the twinning operation, J_1 and J_2 are the true intensities of these reflections as originating from untwinned crystal and α is the twinning ratio, equal to the volume fraction of the smaller twin component.

If α is smaller than 0.5, this system of equations can be solved for J_1 and J_2,

$$J_1 = [(1-\alpha)I_2 - \alpha I_1]/(1-2\alpha),$$
$$J_2 = [(1-\alpha)I_1 - \alpha I_2]/(1-2\alpha).$$

These equations are the basis of the detwinning procedure, which is possible only for $\alpha < 0.5$. Detwinning of perfectly twinned data is impossible, since for $\alpha = 0.5$ the system of detwinning equations becomes indeterminate.

Data collected from merohedrally twinned crystals can very conveniently be tested and detwinned using the Yeates & Fam twinning server at UCLA (http://www.doe-mbi.ucla.edu/Services/Twinning/).

Britton (1972) pointed out, and earlier Zalkin *et al.* (1964) implicitly mentioned, that the assumption of too large a value of α results in an estimation of negative true intensities. Based on this principle, Fisher & Sweet (1980) proposed a practical method for the estimation of the twin fraction α. The 'Britton plot', giving the number of negative intensity estimations as a function of the assumed value of α in the detwinning procedure, has two linear asymptotes, one for $\alpha < \alpha_{\rm opt}$ and another for $\alpha > \alpha_{\rm opt}$. The point at which these two lines cross gives the estimated value of the twin fraction. The Britton plot for gpD and IL-1β is shown in Fig. 4. It suggests twin fractions of 0.34 for gpD and 0.32 for IL-1β.

3.4. Murray-Rust F_1/F_2 plot

The twinning equations can be rearranged to give the ratio J_2/J_1,

$$J_2/J_1 = [I_2 - \alpha(I_1 + I_2)]/[I_1 - \alpha(I_1 + I_2)].$$

Since both J_1 and J_2 are positive, neither numerator nor denominator should be negative, as pointed out by Britton (1972), and these two conditions can be expressed as $\alpha < I_1/(I_1 + I_2)$ and $\alpha < I_2/(I_1 + I_2)$ or, as one inequality,

$$\alpha/(1-\alpha) < I_2/I_1 < (1-\alpha)/\alpha,$$

or in terms of structure amplitudes,

$$[\alpha/(1-\alpha)]^{1/2} < F_2/F_1 < [(1-\alpha)/\alpha]^{1/2}.$$

This property was proposed by Murray-Rust (1973) for estimation of the twin fraction. All points in the graph of F_1 plotted against F_2 should lie between the two limiting straight lines corresponding to $[\alpha/(1-\alpha)]^{1/2}$ and $[(1-\alpha)/\alpha]^{1/2}$. The slope of the lines bounding the points on the graph can be used to estimate the twin fraction. The Murray-Rust plots for gpD and IL-1β are shown in Fig. 5. The twin fractions

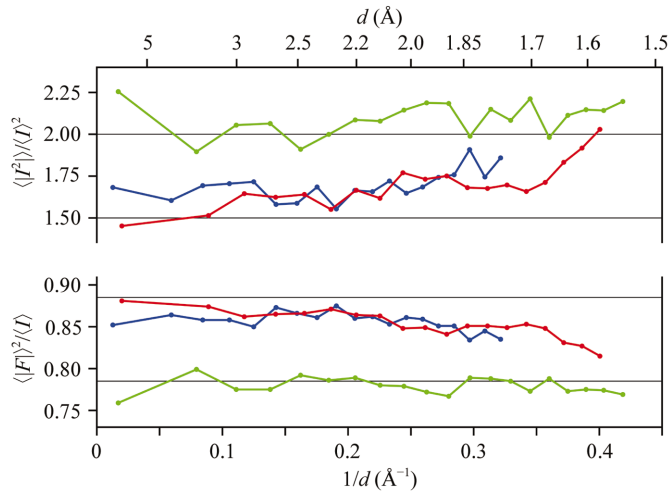

Figure 3
The Wilson ratio $\langle F\rangle^2/\langle I\rangle$ (bottom) and the average intensity ratio $\langle I^2\rangle/\langle I\rangle^2$ (top) calculated in resolution ranges for twinned data from crystals of gpD with $\alpha = 0.36$ (blue lines) and IL-1β with $\alpha = 0.34$ (red lines). As a control, the analogous values for the untwinned accurately measured data from lysozyme in space group $P4_32_12$ (Dauter *et al.*, 1999) are shown in green. Only non-centrosymmetric reflections were included. The theoretically expected values for the Wilson ratio $\langle F\rangle^2/\langle I\rangle$ are 0.785 (untwinned) and 0.885 (perfectly twinned) and for the $\langle I^2\rangle/\langle I\rangle^2$ ratio are 2.0 (untwinned) and 1.5 (perfectly twinned) and are marked as thin black lines. These tests are suggestive of twinning, but the graphs are too noisy for quantitative estimation of the twinning fraction.

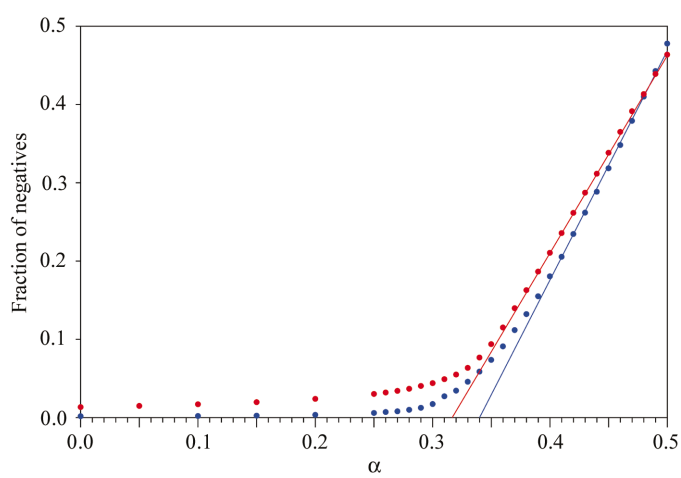

Figure 4
The Britton plot of the number of the negative intensities resulting from the detwinning procedure *versus* the assumed twinning fraction for gpD (blue) and IL-1β (red). The asymptotic straight lines extrapolate the right part of the graph to the estimated twinning ratio.

suggested by these plots are about 0.32 for gpD and 0.34 for IL-1β.

3.5. Rees $N(z)$ plot

Howells *et al.* (1950) first introduced the cumulative intensity distribution $N(z)$ plot to differentiate between centrosymmetric and non-centrosymmetric crystals. The argument z is the fraction of the local average intensity, calculated in narrow resolution ranges, and $N(z)$ is the fraction of reflections with intensities below this level. Various intensity distribution functions show different $N(z)$ curves, particularly for low values of z. As noted above, the fraction of weak intensities is lower for non-centrosymmetric crystals than for centrosymmetric crystals and is lower still for twinned crystals, which shows up clearly on the $N(z)$ plot. This criterion was proposed and theoretically worked out by Stanley (1955, 1972) and elaborated by Rees (1980), who gave the general formula for the cumulative intensity distribution for a non-centrosymmetric case with the twin fraction α, as indicated in Table 3. As Stanley (1955) pointed out, the $N(z)$ curve for twinned crystals is characteristic 'in having an opposite initial curvature'; it has a sigmoidal shape in contrast to an exponential character for normal crystals. He also stated that this test 'would be tedious to apply', which is no longer valid after 50 years of progress in computing technology.

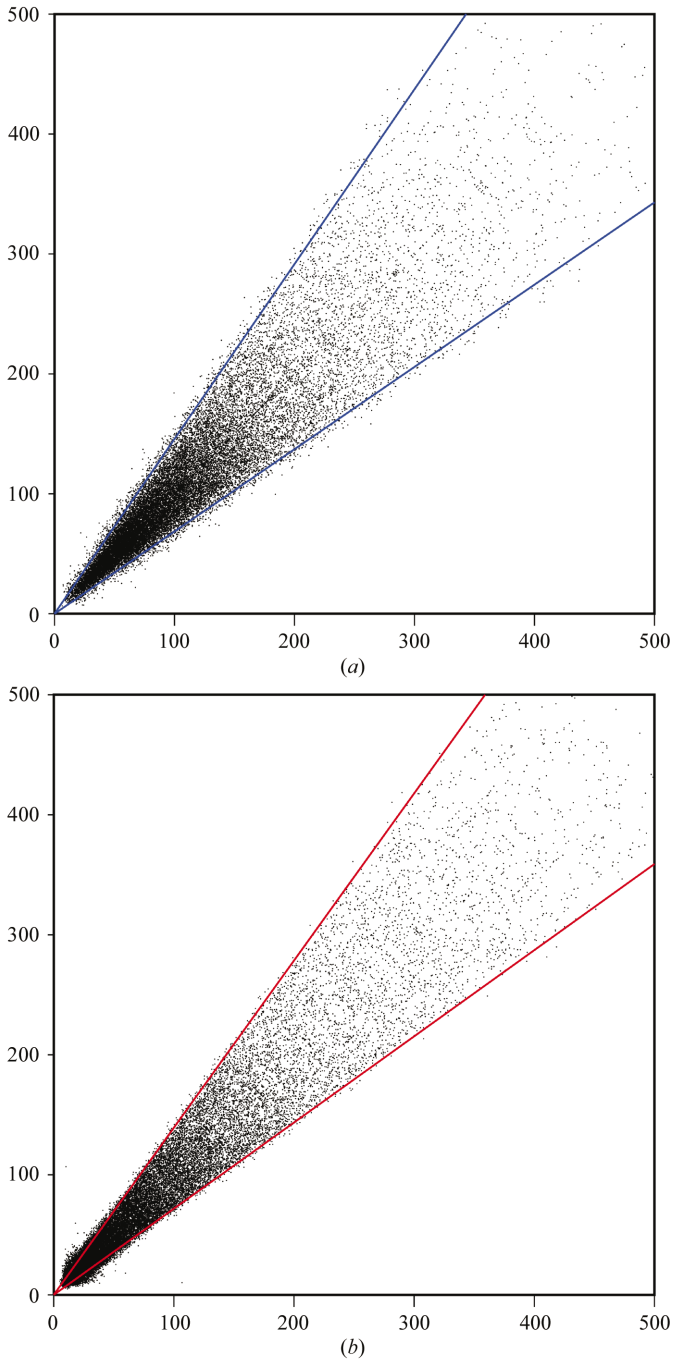

Figure 5
The Murray-Rust plot of pairs of twin-related structure factors F_1 against F_2 for (*a*) gpD and (*b*) IL-1β. The slopes of the boundaries of regions populated by points correspond to the twinning fractions 0.32 and 0.34 for gpD and IL-1β, respectively.

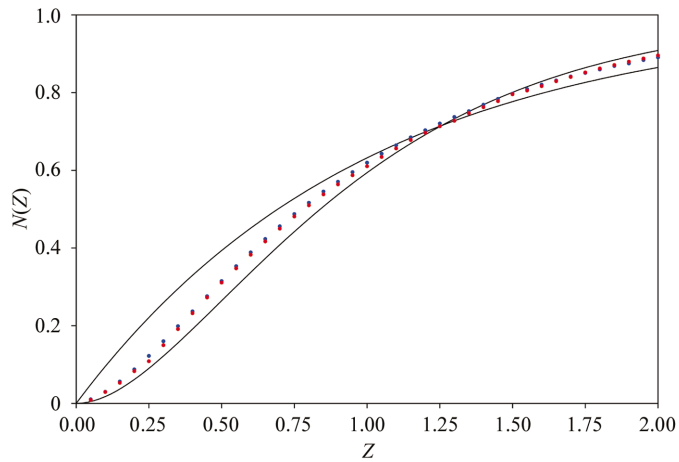

Figure 6
The cumulative intensity distribution $N(z)$ plots for gpD (blue dots) and IL-β (red dots). The black lines correspond to non-twinned (exponential shape) and perfectly twinned (sigmoidal shape) non-centrosymmetric data.

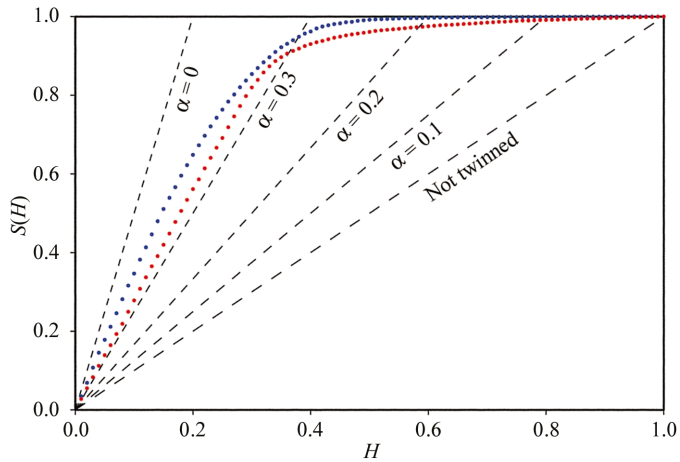

Figure 7
The Yeates $S(H)$ plots indicate the twinning fractions of 0.36 for gpD (blue dots) and 0.32 for IL-1β (red dots). Dashed lines illustrate the theoretically expected slopes $(1 - 2\alpha)^{-1}$ for various twinning fractions.

Table 4
Macromolecular structures of twinned crystals available in the literature.

Structure	Method	Space group	Twin fraction	Reference
Plastocyanin	MR	$P3_2$	0.5	Redinbo & Yeates (1993)
FMD virus O_1 G67	MR	$I23$	0.5	Lea & Stuart (1995)
Procarboxypeptidase A complex	MR	$R3$	0.37	Gomis-Rüth *et al.* (1995)
Fish deoxyhemoglobin	MR	$P2_1$	0.28	Ito *et al.* (1995)
N-terminal half of lactoferrin	MR	$P3_1$	0.40	Breyer *et al.* (1999)
Signal transduction protein P_{II}	MR	$P6_3$	0.17	Carr *et al.* (1996)
Bacteriorhodopsin	MR	$P6_3$	0.46	Luecke *et al.* (1998)
α-Lactalbumin	MR	$P3$	0.49	Chandra *et al.* (1999)
Xenobiotic acetyltransferase	MR	$R3$	0.20	Chandra *et al.* (1999)
R-phycoerythrin	MR	$R3$	0.48	Contreras-Martel *et al.* (2001)
Cocaine hydrolytic antibody	MR	$P2_1$	0.43	Larsen *et al.* (2002)
B-phycoerythrin	MIR	$R3$	0.01–0.06	Fisher & Sweet (1980)
Rat mast cell protease II	SIR + MR	$P3_1$	0.03–0.24	Reynolds *et al.* (1985)
Hydroxylamine oxidoreductase	MIR	$P6_3$	0.08–0.36	Igarashi *et al.* (1997)
RNA heptamer	SIRAS	$P1$	0.5	Mueller *et al.* (1999)
50S ribosomal subunit	MIRAS	$P2_1$	0.27	Ban *et al.* (1999)
Peroxiredoxin 5	SIR	$P2_1$	0.24–0.46	Declercq & Evrard (2001)
Deacetoxycephalosporin C synthase	MIR	$R3$	0.06–0.27	Terwisscha van Scheltinga *et al.* (2001)
gpD	MAD	$P2_1$	0.36	Yang, Dauter *et al.* (2000), Yang, Forrer *et al.* (2000)
Interleukin-1β	MAD	$P4_3$	0.40	Rudolph *et al.* (2003)

Fig. 6 illustrates the $N(z)$ distributions for gpD and IL-1β, together with the theoretical curves for untwinned and perfectly (α = 0.5) twinned crystals. This test qualitatively characterizes twinned crystals, but cannot provide an accurate quantitative estimation of the degree of twinning.

3.6. Yeates *S*(*H*) plot

This test, proposed by Yeates (1988, 1997), is based on the behavior of the ratio of the difference to the sum of intensities of reflections related by the twin operation, $H = |I_1 - I_2|/(I_1 + I_2)$. The dependence of the cumulative distribution of this parameter, $S(H)$, on the twin fraction is very simple (see Table 3) and is linear in H for non-centrosymmetric crystals. The slope of the $S(H)$ plot is $1/(1 - 2\alpha)$ and depends on the twinning factor more sensitively than the $N(z)$ plot. Moreover, omission of pairs of weak reflection from the calculations does not bias the results, but actually leads to improved accuracy of the estimation of the twin fraction. This way of estimating a twin fraction is usually more robust and accurate than other tests. Fig. 7 illustrates the $S(H)$ plots for gpD and IL-1β. Apart from the cumulative distribution $S(H)$, the average values $\langle H\rangle$ and $\langle H^2\rangle$ are characteristic for particular twin fractions, as noted in Table 3. These average values are 0.168 and 0.0408 for gpD and 0.154 and 0.0333 for IL-1β, which gives an estimation of the twin fraction as 0.332 and 0.325 for gpD and as 0.346 and 0.342 for IL-1β, respectively.

3.7. Local intensity differences: *L*-function

Recently, Padilla & Yeates (2003) introduced a very robust statistic, especially suited to crystals displaying abnormalities in their diffraction patterns, resulting for example from pseudo-centering or anisotropy. In such cases, the intensities of some reflections are systematically different. If diffraction extends to higher resolution in one direction of reciprocal space than in another, the intensity statistics for such anisotropically diffracting crystals will be abnormal. Similarly, if there are multiple molecules in the asymmetric unit related by a parallel translation by a vector close to a fraction of the unit cell, some classes of reflections, depending on the parity of their indices, will have abnormally weak intensities, shifting the overall intensity statistics towards the centrosymmetric character, in the direction opposite to that of twinning. In such cases, the twinning tests described above will not be decisive. However, the newly proposed L-function, based on the local intensity distribution, is very robust and unambiguous.

The quantity L is the ratio of the difference to the sum of a pair of reflections $L = (I_1 - I_2)/(I_1 + I_2)$, but the two reflections I_1 and I_2, in contrast to the H-function, are not related by the twin law. Instead, they are selected in the vicinity of each other in the reciprocal space. For pseudo-symmetric crystals their indices may for example differ by 2 or 4, comparing neighboring reflections from the same parity group.

The average values of the moments of L are characteristic for twinned and untwinned crystals, as well as its cumulative distribution, $N(|L|)$, are given in Table 3.

3.8. Additional remarks

In the detwinning of diffraction data and the use of the twinning tests described above, one should be aware of three important points. Firstly, detwinning is possible only if α is not too close to 0.5. Fisher & Sweet (1980) estimated that the relation between uncertainties of the detwinned, $\sigma(J)$, and measured, $\sigma(I)$, intensities is

$$\sigma(J) = \sigma(I)[(1 - 2\alpha + 2\alpha^2)^{1/2}/(1 - 2\alpha)]$$

and for α approaching 0.5 the detwinned intensities become very inaccurate. Of course, for perfect twins, with α = 0.5, detwinning is not possible at all.

Secondly, several twinning tests (those of Britton, Murray-Rust and Yeates) are based on comparison of the twin-related reflection intensities. This obviously requires that the diffraction data are merged in the proper low-symmetry point group. However, if the data were collected from a crystal that is not twinned, the above tests would indicate perfect twinning, since the reflections connected by the potential twinning relation (but in fact the valid symmetry operation) have equal intensities. As mentioned before, the R_{merge} value does not discriminate between perfectly twinned and non-twinned crystals. The only useful tests in such cases are those based purely on

the distribution of measured intensities; that is, Wilson ratios or $N(z)$ cumulative intensity distribution, where it does not matter whether the data have been merged in the crystal or the lattice symmetry. It is a good practice to always check the $N(z)$ distribution of the processed data, particularly in the hexagonal, tetragonal or cubic space groups.

Thirdly, care should be exercised in detwinning the anomalous diffraction data in order to avoid mixing the positive and negative Friedel mates. Since macromolecules are chiral, the twinning operations can only be pure rotations, always relating Friedel mates of the same character. Therefore, the positive and negative Friedel mates can be detwinned separately. In practice, however, attention should be directed to the definition of the asymmetric unit in this procedure. For example, for the monoclinic crystal of gpD, the asymmetric unit defined as in the *CCP*4 (Collaborative Computational Project, Number 4, 1994) programs is $-h_{\max} \leq h \leq h_{\max}$, $0 \leq k \leq k_{\max}$, $0 \leq l \leq l_{\max}$ and the twinning operation is the twofold axis diagonal between the a and c directions, relating reflections h, k, l and $l, -k, h$. However, reflections with negative k do not belong to the same asymmetric unit and have to be converted by the mirror symmetry to l, k, h (if h is positive) or to $-l, k, -h$ (if h is negative), which requires exchanging the Friedel mates. For example, the measurement $I^{+}(2, 3, 4)$ should be paired with $I^{-}(4, 3, 2)$ and $I^{+}(-3, 4, 5)$ with $I^{-}(-5, 4, 3)$. Reflections $h0l$ are centrosymmetric and reflections hkh and $-hkh$ are somewhat special, since the twinning operation relates them with their own Friedel mates.

4. Solution of twinned crystal structures

4.1. Phasing

Not many macromolecular structures of twinned crystals are available in the literature. Basic information on these structures is collected in Table 4. However, there is probably a number of available structures that may be undetected twins (Yeates & Fam, 1999).

If the twin fraction α is smaller than 0.5, it is possible to first detwin the data and then use such data for structure solution. It is not clear over which range of α the detwinning procedure is beneficial. If α is small, it may not be necessary to apply detwinning in order to obtain the structure solution; if α is too close to 0.5, the detwinning procedure may introduce errors exceeding the effect of twinning itself. The benefit gained from the detwinning procedure significantly depends on the accuracy of the originally measured intensities. Often, the structures can be solved from the original, not the detwinned, data. In fact, the structure of gpD was solved by MAD before detecting that the crystal used had a twin fraction of 0.36 (Yang, Dauter *et al.*, 2000).

If the molecular-replacement calculations are performed on the original data, multiple solutions can be expected in the rotation function, reflecting different orientations of the individual twin domains with corresponding weights. In the translation function the smaller twin domain contributes additional noise, but the principal solution should correspond to the main domain.

The MIR approach relies on the small differences between the native and derivative data; if the twin fractions of derivative and native crystals are high or significantly different, the phasing procedure can be severely impaired. According to the analysis of Yeates & Rees (1987), four derivatives are necessary in principle to uniquely estimate the protein phase, instead of two as in the classic MIR approach.

The identification of heavy-atom sites may be not trivial. It may not be possible to fully interpret the isomorphous difference (as well as the anomalous difference) Patterson for highly twinned crystals in any of the potential space groups. Since the X-rays diffracted by the individual twin domains do not interfere, the Patterson synthesis calculated with such data corresponds to the superposition of two Patterson maps transformed by the twinning operation. Certain Harker sections will reflect the higher symmetry, but the potential Harker sections originating from the twinning operations (rotations) will be featureless, even with measured intensities merged in the higher apparent but wrong symmetry. For example, for a perfectly twinned crystal in the space group $P4$, the data can be merged in $P422$. In the Patterson synthesis the Harker section $w = 0$ will have $4mm$ symmetry, as expected for the 422 class, but the Harker sections in the (100) and (110) planes will not show any peaks resulting from the potential existence of dyads in the ab plane.

The MAD and SAD approaches may be expected to be less severely impaired by twinning than MIR, since diffraction data are usually collected from a single specimen with a constant twin fraction.

4.2. Refinement

Once the preliminary model of the structure is available, its parameters and the twin fraction can be refined, taking into account the twinning relations between structure amplitudes. This is possible by the method proposed by Jameson (1982) and Pratt *et al.* (1971) or the similar iterative approach (Redinbo & Yeates, 1993; Yeates, 1997), where the calculated squared amplitudes and the derivatives are accumulated from the individual twin contributions after appropriate transformation. This approach is implemented in *SHELXL* (Sheldrick & Schneider, 1997; Herbst-Irmer & Sheldrick, 1998) and *CNS* (Brünger *et al.*, 1999). It has the additional advantage of producing on output the detwinned structure factors for standard Fourier map calculations. Almost all twinned macromolecular crystal structures have been refined in this way.

5. SAD calculations on gpD

The SeMet variant of the capsid-stabilizing protein of bacteriophage λ, gpD, crystallizes in space group $P2_1$ with unit-cell parameters $a = 45.51$, $b = 68.52$, $c = 45.52$ Å, $\beta = 104.4°$, which are very similar to those of the native crystals. Because the a and c parameters are virtually identical, this cell can be

Table 5
Identification of anomalous sites and phasing statistics for gpD.

Data	gpD^O	gpD^D	gpD^T
Location of Se sites with *SHELXD*			
Percentage of good solutions	98	99	97
Height of peak No. 1	1.00	1.00	1.00
Height of peak No. 6	0.62	0.57	0.55
Height of peak No. 7	0.14	0.16	0.20
Phasing with *SHELXE*			
Phase difference (°)	53.1	56.7	65.1
CC total	0.69	0.63	0.49
CC main chain	0.76	0.71	0.57
CC side chains	0.64	0.57	0.46
Residues built by *APR*/*wARP* (out of 285)	277	272	6

indexed as $C222_1$ with a = 56.02, b = 72.13, c = 69.03 Å, where the orthorhombic a and b axes are parallel to the diagonals between the monoclinic a and c directions. Each monomer has two SeMet residues and the asymmetric unit contains three monomers arranged according to the non-crystallographic threefold axis. The structure has been solved from the 1.7 Å SeMAD data and refined anisotropically against the 1.1 Å native data (Yang, Forrer *et al.*, 2000) to an R factor of 9.9% with *SHELXL*. Both SeMet and native crystals were pseudomerohedrally twinned (Figs. 3, 4, 5*a* and 6) with a twin fraction which refined with *SHELXL* to 0.36. The details of structure solution and some MAD tests using different programs and protocols have been published separately (Yang, Dauter *et al.*, 2000).

Here, other phasing tests are presented based on the single-wavelength anomalous diffraction (SAD) approach. Only the 1.7 Å peak-wavelength SeMet data set was used for these calculations. The original data were modified in two ways. Firstly, they were detwinned according to the principles explained above. Secondly, the intensities related by the twin operation (a rotation around the twofold axis diagonal between a and c axes) were averaged, yielding data corresponding to a perfectly twinned crystal with α = 0.5. These three data sets will be referred to as gpD^O (original), gpD^D (detwinned) and gpD^T (twinned).

Fig. 8 shows the $v = \frac{1}{2}$ Harker sections of the anomalous difference Patterson syntheses calculated with the three sets of data. All six self-vectors are clearly visible in all three cases. As expected, the gpD^D synthesis is marginally clearer than gpd^O, with most of the peaks corresponding to the minor twin domain suppressed. Again, as expected, the gpD^T synthesis has orthorhombic symmetry, since contributions of both domains were deliberately equalized.

The location of the selenium sites was also attempted by the direct-methods approach with *SHELXD* (Schneider & Sheldrick, 2002) based on Bijvoet differences. This step was highly

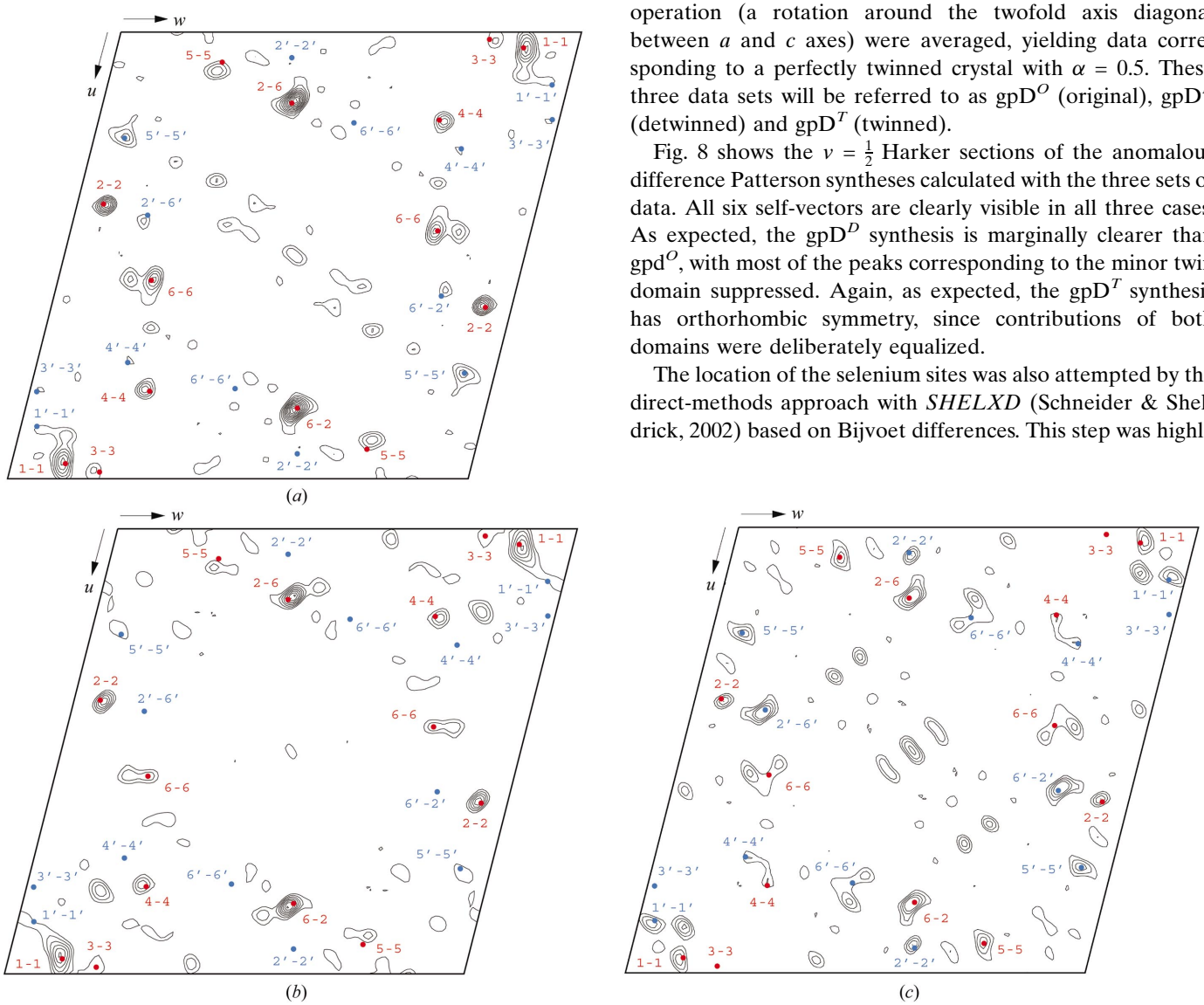

Figure 8
The Harker section $v = \frac{1}{2}$ of the anomalous difference Patterson for a SeMet gpD crystal, calculated with (*a*) the original gpD^O data, (*b*) the detwinned gpD^D data and (*c*) the fully twinned gpD^T data. The self-vectors corresponding to six Se atoms in the major twin domain are labeled in red and the vectors expected for the minor domain are marked in blue.

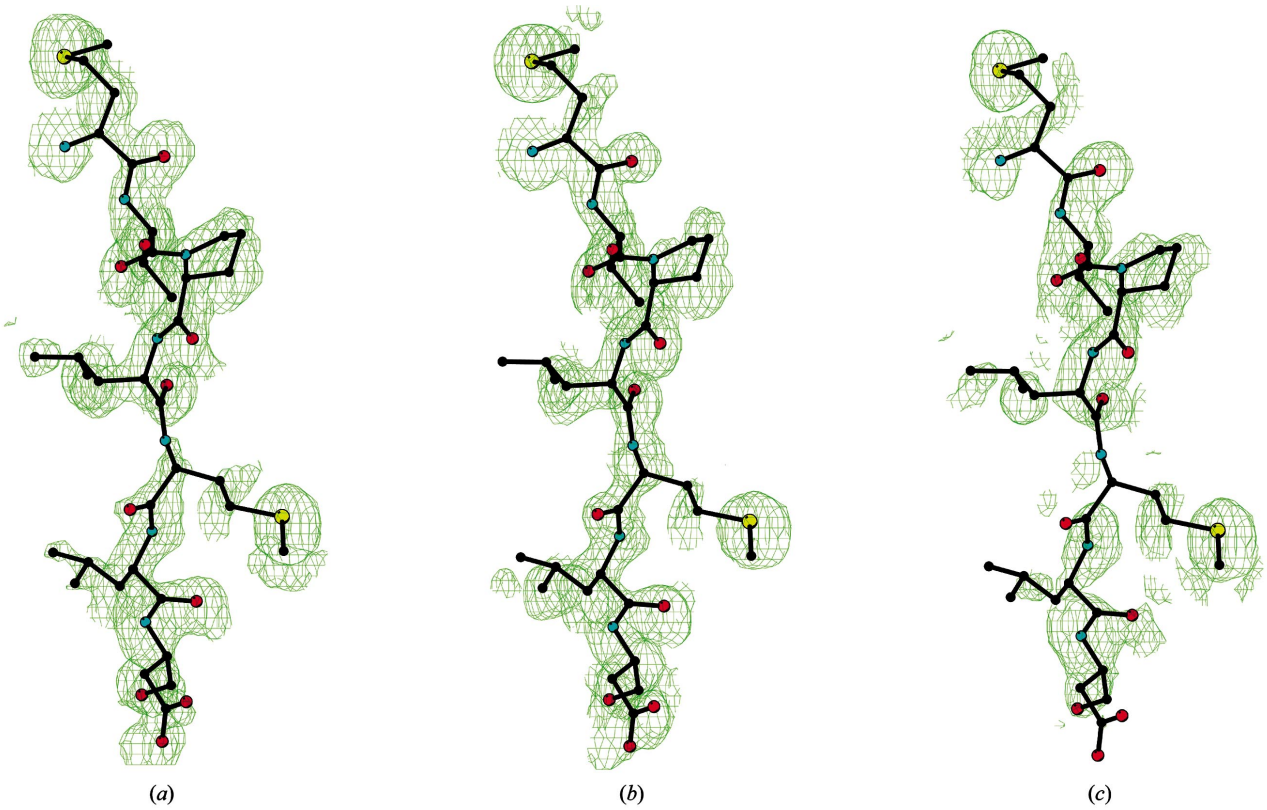

Figure 9
A region of the electron-density map at the 1σ level calculated with *SHELXE* phases for (*a*) the original gpD^O data, (*b*) the detwinned gpD^D data and (*c*) the fully twinned gpD^T data.

successful, with almost all multisolution trials leading to proper solutions of selenium sites, with a clear contrast between the last correct (sixth) and first spurious (seventh) peak (Table 5*a*). The results obtained from the perfectly twinned gpD^T data were almost as clear as from the other two sets.

The identified anomalous scatterers sites were input to *SHELXE* (Sheldrick, 2003) for the estimation of protein phases with a simultaneous density modification. The phase sets obtained were input to automatic model building with *ARP*/*wARP* 5.1 (Perrakis *et al.*, 1999) in conjunction with *REFMAC* (Murshudov *et al.*, 1999). The results of phasing are summarized in Table 5(*b*) and a region of the resulting electron-density maps is illustrated in Fig. 9. The quality of phasing against the gpD^O and gpD^D data was practically equivalent, with an average error of 53–57°. The correlation coefficient (CC) between the *SHELXE* and the final refined electron-density maps was 63–69% overall and 71–73% for the main chain (Figs. 9*a* and 9*b*). *ARP*/*wARP* automatically built more than 90% of amino acids into both the gpD^O and gpD^D maps. The protein phase estimation was less successful with the fully twinned gpD^T data, where the resulting average phase error was 65° and the map was much poorer, with a CC of 49% overall (57% for main chain). This map was not clear enough for *ARP*/*wARP*, but a visual inspection showed that large portions of this map were interpretable (Fig. 9*c*). The stepwise model building and refinement of the partial model and the twin fraction with *SHELXL* eventually led to a complete structure.

6. Conclusions

Several simple tests exist and can be used to check if the diffraction data originate from twinned crystals. Such tests are strongly recommended as a 'default' for crystals of high symmetry or when unit-cell dimensions are amenable to potential pseudomerohedral twinning. Early detection of this phenomenon allows the crystallographer to avoid 'unexpected' obstacles in various stages of the structure solution and model refinement. Even if the initial model can be built with the original data, the existence of twinning, if not properly accounted for, will seriously impair the process of model refinement and the appearance of Fourier maps.

The degree of the adverse effect of twinning on the process of crystal structure solution is critically dependent on the accuracy of measured diffraction intensities and, of course, on the twin fraction. As a consequence of the resulting inaccuracies, the potentially constructive detwinning of the original data may in fact lead to more harm than benefit, considerably increasing the average error level of the estimated intensities. If the twin fraction is small, it may be not worthwhile performing the data detwinning; if it is close to 0.5,

detwinning will introduce very large errors. The range of twin fraction where the data detwinning is beneficial may be different for different cases. However, it is always beneficial to account for twinning in model refinement and map calculation.

As shown with the example of the gpD data, crystal twinning does not absolutely prevent the use of the anomalous signal for a successful structure solution. It may be possible even for a perfectly twinned crystal with accurately measured data and strong enough anomalous scattering signal.

The author wishes to express his gratitude to Markus Rudolph for the kind provision of the interleukin-1β data set and permission to use it as a test example at the CCP4 Study Weekend (York, January 2003) and in this paper

References

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B. & Steitz, T. A. (2000). *Nature (London)*, **400**, 841–847.

Breyer, W. A., Kingston, R.L., Anderson, B. F. & Baker, E. N. (1999). *Acta Cryst.* D**55**, 129–138.

Britton, D. (1972). *Acta Cryst.* A**28**, 296–297.

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., Rice, L. M., Simonson, T. & Warren, G. L. (1999). *Acta Cryst.* D**54**, 905–921.

Carr, P. D., Cheah, E., Suffolk, P. M., Vasudevan, S. G., Dixon, N. E. & Ollis, D. L. (1996). *Acta Cryst.* D**52**, 93–104.

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.

Contreras-Martel, C., Martinez-Oyanedel, J., Bunster, M., Legrand, P., Piras, C., Vernede, X. & Fontecilla-Camps, J. C. (2001). *Acta Cryst.* D**57**, 52–60.

Dauter, Z., Dauter, M., de La Fortelle, E., Bricogne, G. & Sheldrick, G. M. (1999). *J. Mol. Biol.* **289**, 83–92.

Declercq, J.-P. & Evrard, C. (2001). *Acta Cryst.* D**57**, 1829–1835.

Donnay, G. & Donnay, J. D. H. (1974). *Can. Miner.* **12**, 422–425.

Fisher, R. G. & Sweet, R. M. (1980). *Acta Cryst.* A**36**, 755–760.

Friedel, G. (1928). *Leçons de Cristallographie.* Paris: Berger-Levrault.

Gomis-Rüth, F. X., Fita, I., Kiefersauer, R., Huber, R., Aviles, F. X. & Navaza, J. (1995). *Acta Cryst.* D**51**, 819–823.

Herbst-Irmer, R. & Sheldrick, G. M. (1998). *Acta Cryst.* B**54**, 443–449.

Howells, E. R., Philips, D. C. & Rogers, D. (1950). *Acta Cryst.* **3**, 210–214.

Igarashi, N., Moriyama, H., Mikami, T. & Tanaka, N. (1997). *J. Appl. Cryst.* **30**, 362–367.

Ito, N., Komiyama, N. H. & Fermi, G. (1995). *J. Mol. Biol.* **250**, 648–658.

Jameson, G. B. (1982). *Acta Cryst.* A**38**, 817–820.

Koch, E. (1992). *International Tables for Crystallography*, Vol. *C*, edited by A. J. C. Wilson, pp. 10–14. Dordrecht: Kluwer Academic Publishers.

Larsen, N. A., Heine, A., de Prada, P., Radwan, E.-R., Yeates, T. O., Landry, D. W. & Wilson, I. A. (2002). *Acta Cryst.* D**58**, 2055–2059.

Lea, S. & Stuart, D. (1995). *Acta Cryst.* D**51**, 160–167.

Luecke, H., Richter, H.-T. & Lanyi, J. K. (1998). *Science*, **280**, 1934–1937.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Mueller, U., Muller, Y. A., Herbst-Irmer, R., Sprinzl, M. & Heinemann, U. (1999). *Acta Cryst.* D**55**, 1405–1413.

Murray-Rust, P. (1973). *Acta Cryst.* B**29**, 2559–2566.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.

Padilla, J. E. & Yeates, T. O. (2003). *Acta Cryst.* D**59**, 1124–1130.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Pratt, C. S., Coyle, B. A. & Ibers, J. A. (1971). *J. Chem. Soc. A*, pp. 2146–2151.

Redinbo, M. R. & Yeates, T. O. (1993). *Acta Cryst.* D**49**, 375–380.

Rees, D. (1980). *Acta Cryst.* A**36**, 578–581.

Rees, D. (1982). *Acta Cryst.* A**38**, 201–207.

Reynolds, R. A., Remington, S. J., Weaver, L. H., Fisher, R. G., Anderson, W. F., Ammon, H. L. & Matthews, B. W. (1985). *Acta Cryst.* B**41**, 139–147.

Rudolph, M. G., Kelker, M. S., Schneider, T. R., Yeates, T. O., Oseroff, V., Heidary, D. K., Jennings, P. A. & Wilson, I. A. (2003). *Acta Cryst.* D**59**, 290–298.

Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.

Sheldrick, G. M. (2003). *Z. Kristallogr.* **217**, 644–650.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Stanley, E. (1955). *Acta Cryst.* **8**, 351–352.

Stanley, E. (1972). *J. Appl. Cryst.* **5**, 191–194.

Terwisscha van Scheltinga, A. C., Valegård, K., Ramaswamy, S., Hajdu, J. & Andersson, I. (2001). *Acta Cryst.* D**57**, 1776–1785.

Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.

Yang, F., Dauter, Z. & Wlodawer, A. (2000). *Acta Cryst.* D**56**, 959–964.

Yang, F., Forrer, P., Dauter, Z., Conway, J. F., Cheng, N., Cerritelli, M. E., Steven, A. C., Plückthun, A. & Wlodawer, A. (2000). *Nature Struct. Biol.* **7**, 230–237.

Yeates, T. O. (1988). *Acta Cryst.* A**44**, 142–144.

Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.

Yeates, T. O. & Fam, B. C. (1999). *Structure*, **7**, R25–R29.

Yeates, T. O. & Rees, D. (1987). *Acta Cryst.* A**43**, 30–36.

Zalkin, A., Forrester, J. D. & Templeton, D. H. (1964). *Acta Cryst.* **17**, 1408–1412.