Structure determination of a cocaine hydrolytic antibody from a pseudomerohedrally twinned crystal

Nicholas A. Larsen,[a] Andreas Heine,[a] Paloma de Prada,[b] El-Rashdy Redwan,[a] Todd O. Yeates,[c] Donald W. Landry[b] and Ian A. Wilson[a]*

[a]The Department of Molecular Biology and Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California, USA, [b]The Department of Medicine, Columbia University, College of Physicians and Surgeons, New York City, New York, USA, and [c]Molecular Biology Institute, University of California Los Angeles, Los Angeles, California, USA

Correspondence e-mail: wilson@scripps.edu

Few examples of pseudomerohedrally twinned macromolecular crystals have been described in the literature. This unusual phenomenon arises when a fortuitous unit-cell geometry makes it possible for twinning to occur in a space group that ordinarily does not allow twinning. Here, the crystallization, structure determination and refinement of the cocaine hydrolytic antibody 15A10 at 2.35 Å resolution are described. The crystal belongs to space group $P2_1$, with two molecules in the asymmetric unit and unit-cell parameters $a = 37.5$, $b = 108.4$, $c = 111.3$ Å and β fortuitously near 90°; the refined twinning fraction is $\alpha = 0.43$. Interestingly, the noncrystallographic symmetry (NCS) and twin operators are nearly parallel, which appears to be a relatively frequent situation in protein crystals twinned by merohedry or pseudomerohedry.

1. Introduction

Crystal twinning is believed to be a common growth anomaly, yet is infrequently reported for macromolecules owing to the inherent difficulties associated with solving and refining such a structure (reviewed in Yeates, 1997). Twinning occurs when a single macroscopic crystal is composed of two or more microscopic domains whose underlying lattices differ in orientation (Buerger, 1960; Koch, 1992). In merohedral twinning, the lattices of the separate crystal domains superimpose in all three dimensions, in which case a twin operator may be derived that relates the differently oriented microscopic domains. The superposition of these lattices in the crystal (real space) leads to superposition of the measured reflections (reciprocal space). Therefore, any measured reflection $I_{obs}(\mathbf{h}_1)$ has a twin-related reflection $I_{obs}(\mathbf{h}_2)$ and each is a weighted sum of distinct reflections $I(\mathbf{h}_1)$ and $I(\mathbf{h}_2)$ from the separate crystal domains,

$$I_{obs}(\mathbf{h}_1) = (1 - \alpha)I(\mathbf{h}_1) + \alpha I(\mathbf{h}_2) \qquad (1)$$

$$I_{obs}(\mathbf{h}_2) = \alpha I(\mathbf{h}_1) + (1 - \alpha)I(\mathbf{h}_2), \qquad (2)$$

where α represents the fractional volume of the crystal's minor twin domain ($0 < \alpha \leq 0.5$). This linear system of equations can be solved easily for $I(\mathbf{h}_1)$ and $I(\mathbf{h}_2)$ – a process referred to as detwinning – if α is known accurately. This procedure for detwinning fails in the limit of perfect twinning where $\alpha = 0.5$, in which case the determinant $(1 - \alpha)(1 - \alpha) - \alpha^2 = 0$. From these equations, it is clear that as α approaches 0.5, the twin-related reflections $I_{obs}(\mathbf{h}_1)$ and $I_{obs}(\mathbf{h}_2)$ begin to resemble one another. Therefore, the R_{merge} values from scaling typically give the false impression of a higher apparent symmetry, which often contributes to ambiguity in structure determination and refinement.

True merohedral twinning only occurs in crystal systems where the unit cell is capable of supporting a higher symmetry than the true symmetry of the crystal. Here, the twin operator is part of the crystal system but not of the Laue group and, therefore, merohedral twinning only occurs in crystal systems with more than one Laue group (3, 32, 4, 6 and 23 point groups; Yeates, 1997). However, pseudomerohedral twinning is possible in lower point groups in cases of fortuitous unit-cell geometry. For example, when $\beta = 90°$ in a monoclinic space group, pseudomerohedral twinning can emulate an orthorhombic space group, and when $a = b$ in an orthorhombic space group, pseudomerohedral twinning can mimic a tetragonal space group.

Because merohedral or pseudomerohedral twinning affects intensity statistics – the percentage of weak reflections is reduced – twinning can often be detected by comparing the observed intensity distribution with that expected from Wilson statistics (Chandra *et al.*, 1999; Stanley, 1972; Wilson, 1949). Rather than comparing complete distributions, it is often more convenient to calculate a single parameter to describe the distribution. For example, for non-centrosymmetric space groups $\langle |E^2 - 1| \rangle = 0.736$, where E is the normalized structure factor (Herbst-Irmer & Sheldrick, 1998; Sheldrick, 1990–1995). An observed value that is significantly lower may indicate twinnin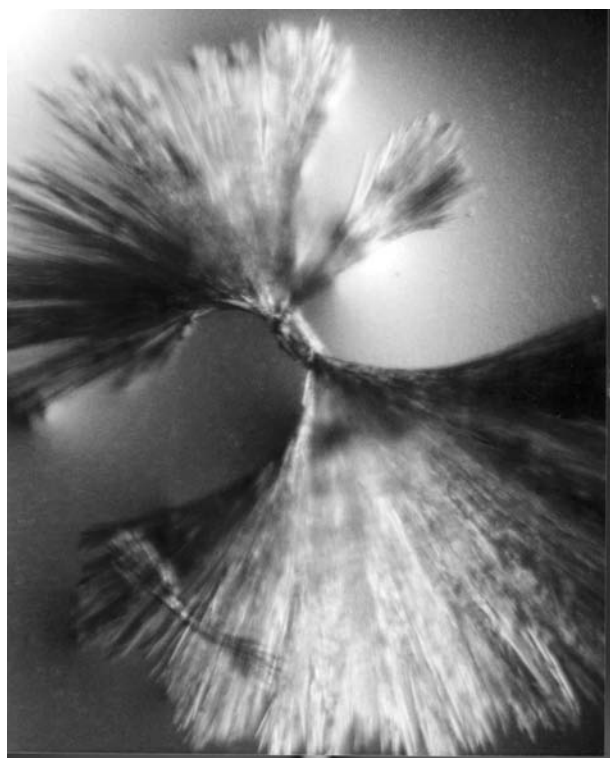g, while a value that is significantly higher may indicate high anisotropy. A particularly simple test is to calculate $\langle I^2(\mathbf{h}) \rangle / \langle I(\mathbf{h}) \rangle^2$ in separate shells or after normalizing in shells (Yeates, 1997). A value of 2.0 is expected for ordinary data (acentric reflections), while fully or 'perfectly' twinned data ($\alpha = 0.5$) give a value of 1.5, with partial twinning giving intermediate values (Stanley, 1972; Yeates, 1997). These simple calculations were critical in detecting and overcoming the unusual type of merohedral twinning encountered in the present work.

2. Materials and methods

2.1. Protein preparation

The murine catalytic antibody 15A10 (Yang *et al.*, 1996) was obtained from cultured hybridoma cells and is composed of a λ light chain and an IgG$_1$ heavy chain. The IgG was digested to Fab and Fc fragments in 1%(w/w) papain and 10 mM cysteine for 4 h at 310 K and stopped with 20 mM iodoacetamide. The Fab was purified by protein A (Repligen) and protein G (APBiotech) affinity chromatography to remove the unwanted Fc fragments and any remaining traces of papain, respectively. The purified Fab was dialyzed into 0.2 M sodium acetate buffer pH 5.5 and concentrated to 20 mg ml^{-1} for crystallization.

2.2. Crystallization

Crystals of Fab 15A10 were grown from 12% PEG 4K, 12–14% 2-propanol, 100 mM MES pH 5.4 at 296 K in sitting drops equilibrated by vapor diffusion. The crystals appear overnight and form a cluster of needles that appear to stem from a single point of nucleation (Fig. 1). One such cluster was shattered into multiple fragments in an attempt to isolate single crystals and flash-cooled in liquid nitrogen using the mother liquor substituted with 20%(v/v) glycerol as a cryoprotectant.

2.3. Data collection and processing

One of the isolated crystal fragments (see §2.2) gave a relatively clean diffraction pattern that was readily indexed by both *DENZO* (Otwinowski & Minor, 1997) and *MOSFLM* (Arndt & Wonacott, 1977; Figs. 2a and 2b). Intense reflections ($I\sigma_I^{-1} > 20$) at low resolution were used to determine the dimensions of the integration box during indexing and data processing. During integration with *DENZO*, the agreement between predicted and actual reflections was monitored by visual inspection and was consistent with correctly refined unit-cell parameters, detector and crystal parameters and crystal-to-detector distance, and had acceptable positional errors (χ^2 values < 1 throughout). However, refinement of the crystal mosaicity was unstable during integration. On closer inspection, select zones of the diffraction pattern appear to contain split reflections (Fig. 2c) and hence *DENZO* attempted to refine the mosaicity as 2.4° or greater. Visual inspection of the predicted reflections and actual reflections indicated that the inflated mosaicity correctly predicted the split reflections, but incorrectly predicted split reflections in

Figure 1
Fab 15A10 crystals. The structure was determined from a similar cluster of crystals that was shattered to isolate single crystals. Seeding and systematic variation of buffer, additives and temperature were attempted to obtain a more favorable morphology, but these efforts were unsuccessful. The long axis of the crystal corresponds to the short axis (*a*) of the unit cell.

zones that were ostensibly normal (Fig. 2b). Hence, the mosaicity was fixed at 1.4°, which appeared qualitatively to achieve the most reasonable compromise for the predictions in the split and normal zones. Using this lower mosaicity, the number of pixels separating a pair of split reflections dictated what fraction of the pair was fully encompassed by the integration box. Difficulty in refining the mosaicity during processing was also encountered with post-refinement during scaling (Otwinowski & Minor, 1997). Hence, the mosaicity was fixed at 1.4° throughout data processing and scaling, which facilitated post-refinement of the crystal/detector orientations and the unit-cell parameters (Tables 1 and 2).

2.4. Structure determination and refinement

In principle, structure determination by molecular replacement should not be hampered by crystal twinning (Redinbo
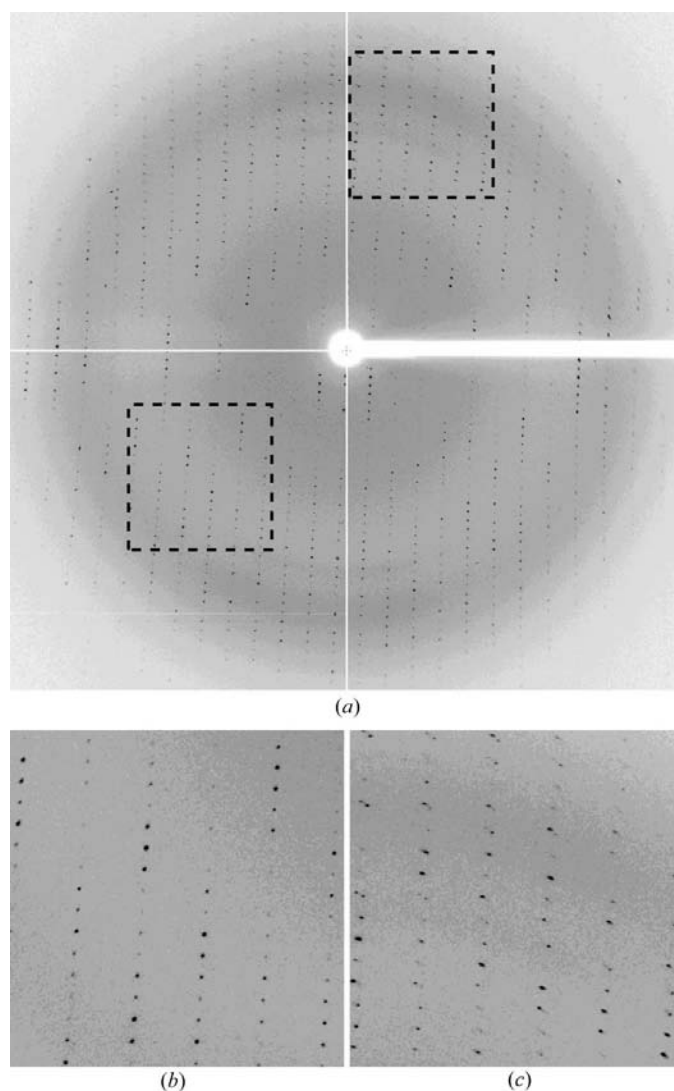

Figure 2
Diffraction pattern from 15A10. (a) The overall pattern shows clearly defined zones. Dashed boxes denote the regions expanded in (b) and (c). (b) Close inspection shows reflections that are ostensibly normal. (c) Select reflections in this zone are apparently split.

& Yeates, 1993) and, therefore, the twinned data were used to solve the structure. Initially, *MERLOT* (Fitzgerald, 1988) was used to rapidly screen ∼100 Fabs for potential solutions to the rotation function. Two of the top ten solutions contained a λ light chain and an IgG_1 heavy chain [PDB codes 1mfa (Zdanov et al., 1994) and 1ine (Love et al., 1993)], in agreement with the 15A10 isotype. The structure was determined using the variable domain of 1mfa and the constant domain of 1ine as search models in *EPMR* (Kissinger et al., 1999), which randomly places the molecules in the asymmetric unit and employs an evolutionary search algorithm to optimize the most promising solutions.

The structure was initially refined against intensities (Chandra et al., 1999) in *CNS* (Brünger et al., 1998) with the twinned data. The initial maps had a reasonable figure of merit (FOM = 0.51), but unacceptable R values after the first round of rebuilding: $R_{cryst} = 33.5\%$ and $R_{free} = 41.7\%$. Inclusion of the twinning operator $(h, -k, -l)$, an estimated twinning fraction and NCS restraints in *SHELXL* (Sheldrick & Schneider, 1997) drastically reduced the R values. Prior to solving the structure, the twin fraction was estimated from a statistical comparison of twin-related reflections (Yeates, 1997) to be 0.37. This value was optimized during atomic refinement. The final refined structure has twinned $R_{cryst} = 18\%$ and $R_{free} = 27\%$ and a final refined twinning fraction $\alpha = 0.43$. These twinned R values probably underestimate the true crystallographic R values to some degree (Redinbo & Yeates, 1993; Rees, 1982), but this issue has not yet been explored thoroughly in the literature.
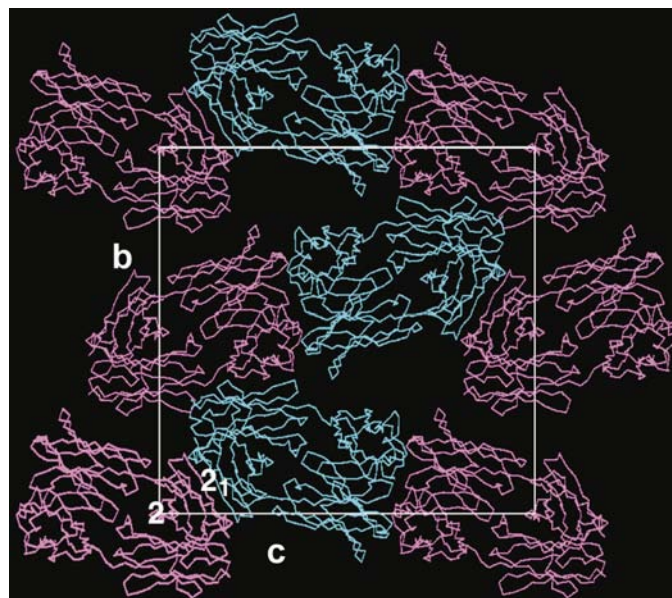
Figure 3
Packing in the unit cell. The crystal packing is viewed down the *a* axis or the long axis of the crystal. The asymmetric unit for the $P2_1$ space group is $(0, 0, \frac{1}{2})$ and the 2_1 crystallographic symmetry operator lies along the *b* axis. The two molecules (green and magenta) in the asymmetric unit are indicated with respect to the unit cell and are related to one another by a NCS 2_1 screw axis. In this view, the NCS symmetry operator and the twin operator are normal to the plane of the paper and are denoted in the figure by 2_1 and 2, respectively.

Table 1
Options for 15A10 lattice.

Twinned 15A10 data were integrated in space group $P1$ and merged in triclinic, monoclinic and orthorhombic point groups.

Lattice	Unit-cell parameters (Å, °)	R_{merge}† (%)
Triclinic	$a = 37.5$, $b = 108.6$, $c = 111.4$, $\alpha = 90.2$, $\beta = 89.9$, $\gamma = 90.2$	6.7
Monoclinic	$a = 37.5$, $b = 108.4$, $c = 111.3$, $\beta = 90$	8.3
Monoclinic	$a = 108.4$, $b = 37.5$, $c = 111.3$, $\beta = 90$	13.6
Monoclinic	$a = 37.5$, $b = 111.3$, $c = 108.4$, $\beta = 90$	14.1
Orthorhombic	$a = 37.5$, $b = 108.3$, $c = 111.3$	14.5

† $R_{merge} = [\sum_h \sum_i |I_i(h) - \langle I(h)\rangle|/\sum_h \sum_i I_i(h)] \times 100$, where $I(h)$ is the average intensity of i symmetry-related observations of reflections with Bragg index h.

3. Results and discussion

The split-reflection pathology (Fig. 2) and unrefined crystal mosaicity provided the first indication of possible pseudo-merohedral twinning in the 15A10 crystals. The data were initially processed in the $P1$ space group ($R_{merge} = 6.7\%$) and imported into *XPREP* (Sheldrick, 1990–1995) to search for higher symmetry (Table 1). Assignment of the space group from inspection of the R_{merge} values alone (Table 1) was inconclusive. The data scaled well in a monoclinic space group, but the moderate R_{merge} of 14.5% in the orthorhombic space group suggested the possible presence of additional symmetry not accounted for by the monoclinic space group. Sometimes, twinning can be diagnosed by analysis of the crystal packing density, where the asymmetric unit of the higher symmetry cell is not large enough to accommodate the protein (Gomis-Rüth *et al.*, 1995; Redinbo & Yeates, 1993). Here, however, the Matthews coefficient would be a reasonable 2.21 $\text{Å}^3\,\text{Da}^{-1}$ (solvent content = 44%) for an orthorhombic cell, assuming one molecule per asymmetric unit. The question about symmetry and the correct space group was resolved by calculation of the intensity distribution, $\langle I^2(\mathbf{h})\rangle/\langle I(\mathbf{h})\rangle^2 = 1.7$ and $\langle |E^2 - 1|\rangle = 0.604$, implying that the true space group is monoclinic and that the data are pseudomerohedrally twinned (Yeates, 1997). Assignment of the monoclinic space group is further supported by the split reflections, which would imply that the β angle does not exactly equal 90°, but may in fact deviate by ±0.1° as implied by the refined triclinic cell (Table 1), since no restrictions or assumptions are imposed on the angles for this space group.

Two Fab molecules were anticipated in the monoclinic asymmetric unit. Strong solutions to the rotation function were obtained from twinned data in space group $P2_1$ with the programs *MERLOT* (Fitzgerald, 1988) and *AMoRe* (Navaza, 1994). For twinned data, solutions to the rotation function are typically found for each of the twin domains and placed accordingly with shrewd application of the translation function (Redinbo & Yeates, 1993). In the case of 15A10, however, *AMoRe* apparently found rotation solutions for only one of the twin domains, which may be explained by the coincidence of the NCS and twin operator (Fig. 3). Curiously, *AMoRe* was unable to find solutions for the translation function. The correlation coefficients after translation in *AMoRe* were significantly worse than those obtained from the rotation

Table 2
Final data-collection statistics.

Parentheses denote outer shell statistics.

Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 37.5$, $b = 108.4$, $c = 111.3$, $\beta = 90$
Resolution range	30–2.5 (2.45–2.35)
Observations	91574
Unique reflections	32009
Completeness (%)	85.4 (76.6)
Average $I/\sigma(I)$	14.4 (2.8)
R_{merge} (%)	8.3 (39.9)

function alone; moreover, inspection of the packing revealed overlapping interactions between crystallographic and non-crystallographic symmetry mates. This failure to pack is likely to have occurred because *AMoRe* was attempting to pack both twin domains simultaneously. Ultimately, the structure was solved by searching for the two anticipated constant (C) domains in *EPMR* (Kissinger *et al.*, 1999), fixing those solutions as a partial structure and searching again for the remaining two variable (V) domains. Inspection of the packing shows the V and C associate to form an Fab as expected, with two molecules in the asymmetric unit and no overlapping or clashing domains (Fig. 3).

Several examples of pseudomerohedrally twinned crystal structures have been reported in the literature (Ito *et al.*, 1995; Ban *et al.*, 1999; Frazão *et al.*, 1999; Yang *et al.*, 2000; Declercq & Evrard, 2001). Three of these examples involve a twinned $P2_1$ space group where a is nearly equal to c and there is a higher apparent symmetry corresponding to the C-centered orthorhombic space group (Ito *et al.*, 1995; Ban *et al.*, 1999; Yang *et al.*, 2000). For example, GpD, the capsid-stabilizing protein of bacteriophage λ, crystallizes in space group $P2_1$ with a nearly equal to c, which corresponds to a subgroup of space group $C222_1$ with the twofold twinning operator along the diagonal between a and c (Yang *et al.*, 2000). In contrast, rubredoxin oxygen oxidoreductase forms pseudomero-hedrally twinned crystals in orthorhombic space group $P2_12_12$ with a fortuitously near b, emulating the tetragonal space group $P4_22_12$ (Frazão *et al.*, 1999). In this case, two molecules in the asymmetric unit are related by a NCS twofold coincident with the twinning operator along the pseudo-diagonal between the a and b axes, which generates the near-fourfold NCS axis parallel to c (Frazão *et al.*, 1999). The case of 15A10 differs from these reported examples, since 15A10 involves a twinned monoclinic cell that mimics the primitive orthorhombic space group rather than the C-centered ortho-rhombic or tetragonal space groups.

There is cumulating evidence that non-crystallographic symmetry (NCS) may promote twinning, as the NCS operator is often nearly parallel to the twinning operator (Declercq & Evrard, 2001; Frazão *et al.*, 1999; Liang *et al.*, 1996; Rees & Lipscomb, 1980). In the case of 15A10, the twinning operator is a twofold along the a axis, while the NCS operator is a 2_1 screw axis nearly parallel to the a axis and shifted from where the true crystallographic axis would be if the space group were orthorhombic (Fig. 3). If the NCS operator were a true crys-

tallographic screw axis, the true crystal symmetry would be orthorhombic and the possibility of twinning would be lost – the two microscopic domains with different lattice orientation in the crystal would become indistinguishable from one another. In the present case, the protein molecules are apparently unable to pack in a way that would combine the 2_1 crystallographic axis with the NCS axis to produce orthorhombic symmetry. Instead, 15A10 crystals grow as twins with the NCS axis parallel to the twinning axis.

4. Concluding remarks

The present study demonstrates the determination of a macromolecular crystal structure in the presence of pseudomerohedral twinning. The phenomenon is made possible by fortuitous unit-cell geometry and is apparently promoted by non-crystallographic symmetry. Since it is possible to determine crystal structures by molecular replacement from twinned data, the presence of twinning may remain unnoticed until refinement. In such cases, higher than expected R factors during refinement may provide the only hint that the data are twinned. The occurrence of twinning, even in space groups that normally do not allow it, demonstrates the importance of routinely checking all newly acquired diffraction data sets for twinning.

References

Arndt, U. W. & Wonacott, A. J. (1977). Editors. *The Rotation Method in Crystallography: Data Collection from Macromolecular Crystals*. New York: North-Holland.

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B. & Steitz, T. A. (1999). *Nature (London)*, **400**, 841–847.

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. (1998). *Acta Cryst.* D**54**, 905–921.

Buerger, M. J. (1960). *Crystal Structure Analysis*. New York: John Wiley & Sons.

Chandra, N., Acharya, K. R. & Moody, P. C. E. (1999). *Acta Cryst.* D**55**, 1750–1758.

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

Fitzgerald, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.

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**, 1465–1476.

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

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

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

Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.

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

Liang, J., Ealick, S., Nielsen, C., Schreiber, S. L. & Clardy, J. (1996). *Acta Cryst.* D**52**, 207–210.

Love, R. A., Villafranca, J. E., Aust, R. M., Nakamura, K. K., Jue, R. A., Major, K. G. Jr, Radhakrishnan, R. & Butler, W. F. (1993). *Biochemistry*, **32**, 10950–10959.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

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

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

Rees, D. C. & Lipscomb, W. N. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 3288–3291.

Sheldrick, G. M. (1990–1995). *SHELXTL* 5.03. Siemens Industrial Automation Inc., Madison, WI, USA.

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

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

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

Yang, G., Clun, J., Arakawa-Uramoto, H., Gawinowicz, M. A., Zhao, K. & Landry, D. W. (1996). *J. Am. Chem. Soc.* **118**, 5881–5890.

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

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

Zdanov, A., Li, Y., Bundle, D. R., Deng, S. J., MacKenzie, C. R., Narang, S. A., Young, N. M. & Cygler, M. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 6423–6427.