Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Refinement of *Haemophilus influenzae* diaminopimelic acid epimerase (DapF) at 1.75 Å resolution suggests a mechanism for stereocontrol during catalysis

Adrian John Lloyd,[a] Trevor Huyton,[b,c] Johan Turkenburg[b] and David Ian Roper[a,b]*

[a]Department of Biological Sciences, University of Warwick, Coventry CV4 7AL, England, [b]Department of Chemistry, University of York, Heslington, York YO10 5DD, England, and [c]Structural Biology, Walter and Eliza Hall Research Institute, Melbourne, Victoria 3050, Australia

Correspondence e-mail: droper@bio.warwick.ac.uk

Received 5 June 2003
Accepted 4 December 2003

PDB Reference: diaminopimelate epimerase, 1gqz, r1gqzsf.

Diaminopimelate (DAP) epimerase (DapF) is central to the biosynthesis of both lysine and cell-wall peptidoglycan in many bacteria species. The peptidoglycan layer provides great potential for the development of novel antimicrobials as it is a uniquely prokaryotic feature. Crystals of recombinant *Haemophilus influenzae* DapF that diffract to beyond 2 Å resolution have been obtained which facilitated the solution of the structure by molecular replacement at a resolution approximately 1 Å higher than that previously determined. An analysis of the structure (i) in comparison to other PLP-independent racemaces and (ii) in relation to the catalytic mechanism and stereospecificity of DapF is presented.

1. Introduction

Bacterial peptidoglycan is a macromolecule which provides the cell wall with strength and rigidity, which are vital for cell viability and osmotic integrity. Peptidoglycan is made up of a polysaccharide matrix to which are appended short pentapeptide chains, which are themselves cross-linked. Disruption of the formation of this structure is the basis of the action of important antibiotics such the penicillins and glycopeptides, including vancomycin, the use of which is now becoming limited owing to the evolution of resistance mechanisms employed by various clinically important bacteria (Bugg, 1999). The search for novel antibiotics and novel antibiotic targets is therefore of crucial importance and has renewed interest in antimicrobial drug development. In particular, the peptidoglycan biosynthetic pathway provides a potential route to novel antibiotic compounds, since many of the compounds and enzymes that act on them do not have eukaryotic homologues in either plants or animals.

In Gram-positive bacteria, the third residue in the peptidoglycan pentapeptide is usually L-lysine, whereas in Gram-negative species this residue is normally D,L-diaminopimelate (D,L-DAP; Bugg, 1999). D,L-DAP is an intermediate on the biosynthetic route to L-lysine in bacteria and higher plants but not in mammals (Bugg, 1999; Cox, 1996). As a result, the lysine biosynthetic pathway enzymes themselves are also potential novel antibiotic targets. D,L-DAP is formed from L,L-DAP by the action of diaminopimelate epimerase (DapF) without the aid of cofactors or metal ions in an analogous manner to the pyridoxal phosphate-independent proline, glutamate and aspartate racemaces (Antia *et al.*, 1957; Wiseman & Nichols, 1984; Johnston & Dixen, 1969; Nakajima *et al.*, 1986). DapF differs from these other racemases, however, because it has to discriminate between two stereocentres such that it interconverts the L,L- and D,L-diastereoisomers of DAP by epimerization but cannot further racemize the D,L-DAP diastereoisomer to the D,D-DAP diastereoisomer (Antia *et al.*, 1957). The structure of *Haemophilus influenzae* DapF was published recently by Cirilli *et al.* (1998), concurrent with our efforts to determine the structure of this enzyme. In addition, the structures of two other PLP-independent enzymes, glutamate racemace from *Aquifex pyrophilus* (Hwang *et al.*, 1999) and aspartate racemase from *Pyrococcus horikoshii* (Lui *et al.*, 2002), have also been determined, allowing limited structural comparisons to be made. We further use our data to explain the structural basis of the mechanism and stereochemical control of the reaction catalysed by DapF.

2. Materials and methods

2.1. Cloning and protein production

The gene encoding DapF from *H. influenzae* strain KW20 was amplified by PCR from chromosomal DNA and cloned into the T7 expression vector pET28b (Novagen), facilitating purification of the protein by metal-chelation affinity chromatography.

Cultures of *Escherichia coli* BL21(DE3) containing the DapF expression construct were grown in LB media at 310 K induced with 0.5 m*M* IPTG and were harvested after 3 h. Highly purified DapF was eluted from a Ni^{2+}-chelating Sepharose column (Amersham-Pharmacia Biotech) in 50 m*M* HEPES, 100 m*M* NaCl and 500 m*M* imidizole pH 8.0 and was immediately buffer-exchanged into 10 m*M* HEPES pH 8.0 and 10 m*M* dithiothreitol. SDS–PAGE showed the protein to be free of any major or minor contaminants and it

required no further purification prior to crystallization trials.

2.2. Crystallization and data collection

Crystals of DapF were grown by vapour diffusion using the hanging-drop method. Crystals were obtained in 1 *M* sodium acetate, 0.1 *M* imidazole pH 6.5 at a protein concentration of ~10 mg ml^{-1} in 25 m*M* HEPES pH 8.0, 5 m*M* DTT at 290 K. Crystals of DapF belonging to space group $C222_1$, with unit-cell parameters a = 98.64, b = 113.87, c = 64.48 Å appeared after a few hours and were suitable for X-ray data collection within 24–48 h. Single crystals of DapF were transferred briefly to a cryoprotectant solution consisting of 1 *M* sodium acetate, 0.1 *M* imidazole, 5 m*M* DTT and 35%(*v*/*v*) PEG 300 and were then frozen at 120 K for data collection. Data were collected on beamline 9.6 at the SRS Daresbury to a resolution of 1.75 Å using a MAR345 image-plate device. The data were processed, scaled and reduced using *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

2.3. Structure refinement and solution

During our attempts to solve the structure using MIR and MAD techniques, the X-ray structure of the enzyme at 2.7 Å resolution was published by Cirilli *et al.* (1998). The coordinates for the structure became available in the PDB (code 1bwz). These data were used as a starting model for refinement with *REFMAC* (Murshudov *et al.*, 1997, 1999) using the higher resolution data collected in this study. All crystallographic calculations were performed using the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994).

In the initial model, water molecules were omitted and 5% of the data in the resolution range 15–1.75 Å were omitted from the refinement and used to calculate the $R_{\rm free}$ value (Brünger, 1992). The *R* factor for the initial model was 28.3%, with an $R_{\rm free}$ of 31.4%. After rigid-body refinement, *R* and $R_{\rm free}$ were reduced to 23.5 and 26.1%, respectively. Electron-density maps were calculated and the structure was rebuilt using the *XBUILD* module (Oldfield, 1994) within the program *QUANTA* (Accelrys, San Diego, CA, USA). The first solvent shell was modelled, followed by further refinement interspersed with remodelling. Careful monitoring of the cross-validation reflections suggested that inclusion of scattering from riding H atoms and restrained anisotropic modelling of the thermal displacement parameters was appropriate. This was justified by reductions in $R_{\rm cryst}$ and $R_{\rm free}$ of 2.3/1.4% and 0.9/1.0%, respectively. The final model consists of 2117 protein atoms and 353 water molecules, with a final $R_{\rm cryst}$ of 16.5% and an $R_{\rm free}$ of 19.6% (see Table 1). The structure was validated using the program *PROCHECK* (Laskowski *et al.*, 1993). The final model conforms to expected values for the geometry of protein structures at this resolution. 91% of the residues are in the most favoured region of the Ramachandran plot, with no residues in disallowed regions.

Table 1
Refinement and structure quality statistics for *H. influenziae* DapF at 1.75 Å resolution (PDB code 1gqz).

Values in parentheses refer to the outer resolution shell.

Data quality	
Resolution of data (Å)	56.8–1.75
$R_{\rm merge}$	0.039 (0.212)
No. independent reflections	37359
Completeness (%)	99.18 (99.9)
Mean $I/\sigma(I)$	29.98 (7.41)
Crystal parameters	
Space group	$C222_1$
Unit-cell parameters (Å)	a = 99.64, b = 113.87, c = 64.48
Matthews coefficient $V_{\rm M}$ (Å^3 Da^{-1})	2.82
Solvent content (%)	56
Refinement	
No. protein atoms	2117
No. solvent waters	353
Reflections used in refinement	37280
$R_{\rm free}$	0.196
$R_{\rm cryst}$	0.165
R.m.s. deviation 1–2 bonds (Å)	0.020
R.m.s. deviation 1–3 angles (°)	1.9
R.m.s. deviation chiral volumes (Å^3)	0.107
Overall average B factor	31.652
Average main-chain B (Å^2)	27.26
Average side-chain B (Å^2)	31.47
Average solvent B (Å^2)	45.81

3. Results and discussion

The higher resolution and quality of the data collected in this study compared with that of the previously available model prompted us to continue with refinement of this structure following the publication of the Cirilli structure. Our data are 99.3% complete, compared with the published structure (Cirilli *et al.*, 1998) which had 90% complete data at nearly 1 Å lower resolution. The quality of the electron-density maps was very high; an example of which is given in Fig. 1, which details electron density across the disulfide-bonded active-site residues. During our refinement and model building, it became apparent that residue 111 in our structure is an aspartate in agreement with the *H. influenzae* KW20 genome locus entry HI0750 for DapF and not methionine as assigned in the originally deposited 2.7 Å structure (PDB code 1bwz). Sequencing of our expression construct for DapF confirmed the presence of an aspartate at position 111 as per the sequence database entry and we can only conclude therefore that the 1bwz structure has been incorrectly assigned at that residue.

Figure 1
View of the observed electron density for the disulfide bond formed between Cys73 and Cys217 within the DapF active site. The map shown is a $2F_{\rm obs} - F_{\rm calc}$ map and is contoured at 2σ, with residues in loops in the immediate vicinity of the active site marked.

The structures of two other PLP-independent racemases, *A. pyrophilus* glutamate racemace (MurI) and *P. horikoshii* aspartate racemase, were published shortly after the Cirilli structure of *H. influenzae* DapF. All three enzymes have a two-domain α/β structure and have an apparently similar reaction chemistry in which two conserved cysteines act as a base and an acid to respectively remove the C^{α} proton from an amino acid and replace it, with a resulting L→D inversion of stereochemistry at this position (Hwang *et al.*, 1999; Lui *et al.*, 2002).

Comparison of our DapF structure with sequence alignments of DapF from 57 prokaryotes using *ClustalW* (Thompson *et al.*, 1994) demonstrated that the vast majority of the residues that are conserved line an active-site cleft formed from both domains. Appended to either side of the active-site cleft are the two conserved cysteines (73 and 217 in *H. influenzae*) which constitute the active site and catalyse L→D inversion of L,L-DAP to D,L-DAP. Comparative analysis of the alignments of 55 MurI and eight aspartate racemase sequences with the corresponding *A. pyrophilus* MurI and *P. horikoshii* aspartate racemase structures (PDB codes 1b74 and 1jfl, respectively) also demonstrate a similar arrangement of conserved active-site cysteines associated with a clustering of conserved residues within an active-site cleft. Despite these shared characteristics, a high level of sequence and structural

divergence between the three groups of enzymes has resulted during the course of evolution.

Although DTT was added to our DapF preparations immediately after purification, it is evident that a disulfide bond is formed between the active-site cysteines 73 and 217 in our structure as found in the Cirilli *et al.* (1998) model. This strongly suggests that a disulfide bond is rapidly formed during purification prior to crystallization. In contrast to Cirilli *et al.* (1998), the inclusion of DTT in our crystallization experiments increased the quality, size and diffraction limit of the crystals produced. The close proximity of the sulfydryl groups of Cys73 and Cys217 to each other that leads to their ready oxidation to cystine is consistent with their proposed role in the catalytic cycle of DapF, where they abstract protons from the L- and D-stereocentre C^{α} atoms of L,L- and D,L-diaminopimelic acid, respectively (Koo *et al.*, 2000).

Inspection of our structure revealed that one domain of the enzyme can be overlaid on the other by rotation around the active-site cleft. Inspection of the amino-acid sequence of the two halves of the primary sequence reveals low sequence similarity (35%) with the exception of the active-site region, where a greater similarity exists. As found by Cirilli *et al.* (1998), the secondary structure of the enzyme is dominated by a β-sheet equally partitioned between the two domains of DapF. However, close inspection of the arrangement of the major α-helical elements of our structure reveals a hitherto unreported structural feature likely to be essential to the catalytic function of DapF. Both domains possess an internal α-helix cradled by a β-sheet. These helicies are orientated end-on towards each other, comprising residues Cys217–Gln231 and Gly76–Phe87. The first of these helicies directly positions Cys217 adjacent to Cys73, whilst the latter helix orientates a tripeptide terminated by Cys73 such that this cysteine is now adjacent to Cys217 (Fig. 2*a*). The orientation and length of these two helices is therefore crucial to the relative proximity of Cys217 and Cys73 and essential to maintenance of the optimal orientation of these residues for their interaction with the C^{α} proton of DAP. Similarly, the active-site cysteines Cys82 and Cys194 of the *P. horikoshii* aspartate racemase structure are mounted on the ends of two α-helicies.

The dipole moments formed by helices Cys217–Met230 and Gly76–Phe83 will generate positive charges at their N-termini (Creighton, 1984). The negative influence that this has on the value of the pK_a of the Cys217 and Cys73 thiol groups will differ because the former residue is directly linked to the N-terminus of helix Cys217–Met230, whilst Cys73 is linked to the N-terminus of the Gly76–Phe83 helix *via* a dipeptide. This could well account for the differential reduction in pK_a of these two groups from ~9 to either 7.0 or 6.1 observed by Koo & Blanchard (1999) and ensure that protonation of one thiol relative to the other is maintained, as required by the catalytic mechanism of DapF.

DapF differs from the other PLP-independent racemases thus far characterized in that its substrate (L,L or D,L-DAP) contains two chemically identical stereochemical centres instead of one (D or L-glutamate and D- or L-aspartate for glutamate and aspartate racemases, respectively). As DapF does not generate D,D-DAP from L,L-DAP (Antia *et al.*, 1957), this enzyme exerts a degree of stereochemical control over the reactions it catalyses that prevents the conversion of D,L-DAP to D,D-DAP that is clearly not required by glutamate or aspartate racemases. The structural basis of this control has not been previously addressed with respect to DapF. We therefore examined our DapF structure and the available glutamate racemase and aspartate racemase structures and the sequence alignments of these enzymes in order to gain further insight into how DapF regulates the stereochemistry of the reactions it catalyses.

Sequence alignments of DapF in comparison with the glutamate and aspartate racemases[1] revealed that 16 residues are conserved exclusively in all the DapF sequences we analysed. Reference to our DapF crystal structure revealed that three of these residues (Arg78, Glu208 and Arg209)

[1] The sequence alignments have been deposited in the IUCr electronic archive (Reference: hv5003). Details for accessing this material are described at the back of the journal.

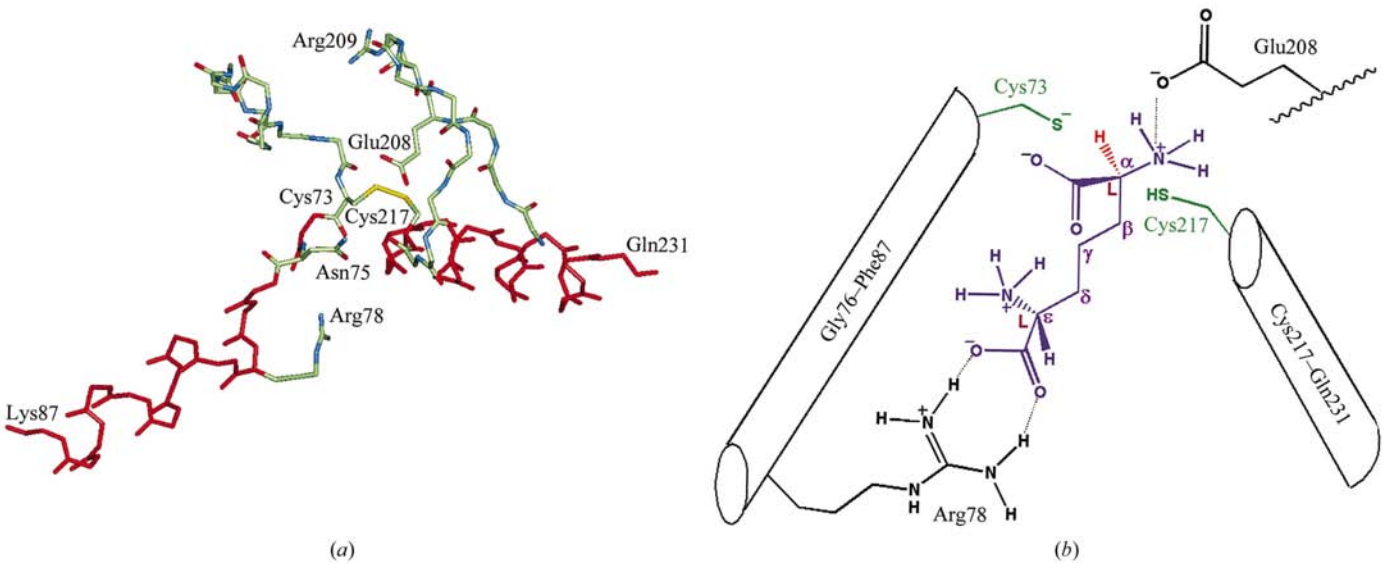

Figure 2
Orientation of *H. influenzae* DapF active-site residues and substrate by α-helices Gly76–Phe87 and Cys217–Gln231. (*a*) α-Helices are coloured red; the universally conserved residues Arg78 and Glu208, implicated in binding the ε-carboxyl and α-amino groups of the DAP substrate and Asn75 are labelled. The two catalytic cysteines Cys73 and Cys217 are labelled in yellow. (*b*) Schematic representation of the active site, highlighting the proposed relationship between the chiral centres of DAP and the active-site residues identified in the text. Active-site cysteines are identified in yellow and the helices from which they emanate are represented as a cylinder.

capable of strong electrostatic interactions with the amino and carboxyl groups of DAP are in the vicinity of the active site. Inspection of the structure by ourselves and Cirilli *et al.* (1998) revealed that the Glu208 δ-carboxyl was 3.4 Å from the Cys73 S atom. This distance is sufficient to allow a charge interaction with the amino group attached to the C^{α} atom of DAP at which the epimerization is taking place.

Further inspection of our structure revealed that the distances between the Cys73–Cys217 disulfide and the guanidinium groups of Arg78 (7.3 Å) or Arg209 (9.7 Å) were sufficient to accommodate the 1-amino pentanoic acid side chain (6.9 Å in length) that extends from the C^{α} atom of the stereocentre of the DAP undergoing epimerization to the ε-carboxyl O atoms of the distal non-reacting L-stereocentre. However, interaction of the DAP ε-carboxyl with Arg209 could probably be discounted because its guanidinium group was incorrectly orientated and binding of DAP between the Cys73–Cys217 disulfide would be sterically hindered by Gln72. In contrast, the guanidinium group of Arg78 was orientated towards the Cys73–Cys217 disulfide and there were no residues that would sterically prevent the binding of the substrate between these residues.

In the light of these findings, we wish to propose that the guanidinium side chain of Arg78 binds the carboxyl of the distal C^{ε} L-stereocentre of DAP which does not undergo epimerization (Fig. 2*b*). The relative proximity of Glu208 to the Cys73–Cys217 disulfide suggests that it interacts with the DAP amino group attached to the C^{α} undergoing D–L epimerization. It is also relevant to note here that the orientation of the helices that control the position of Cys73 and Cys217 relative to each other (see above) crucially maintains the correct position of these residues with respect to Glu208 in order to allow simultaneous juxtaposition of the C^{α} and α-amino group of DAP with the DapF Cys73/217 sulfydryls and the Glu208 carboxyl, respectively.

We therefore postulate that the role of Arg78 is to electrostatically anchor the DAP ε-carboxyl at the invariant L-stereocentre of DAP and that this interaction is specific for the L-stereocentre at this position within the diamino acid, because juxtaposition of the C^{ε} D-stereocentre in place of the C^{ε} L-stereocentre of DAP would bring an amino group into close enough proximity to Arg78 to be strongly disfavoured. Furthermore, it is likely that the interaction of Arg78 with DAP would make an essential contribution to overcoming the activation energy of the epimerization reaction. These factors would ensure that one stereocentre within DAP is maintained in the L-conformation and thus conversion of L,L-DAP to D,D-DAP cannot occur. Consistent with the involvement of an arginine in the stereochemical control of the DapF reaction is the observation that the ionization of the groups responsible for such control occurred at pH values greater than 9 (Koo & Blanchard, 1999).

We also observed that another absolutely conserved residue Asn75 lies next to the path that the alkyl chain of DAP is likely to follow between Arg78 and Cys73/Cys217. In our model, its position is likely to orientate the DAP to dock between Arg78, Cys217 and Glu208, as Asn75 partially blocks simultaneous access to Cys73 and Arg78. Koo *et al.* (2000) have demonstrated that Cys217 is likely to be the residue that abstracts a proton in the D,L→L,L epimerization of DAP. We therefore suggest that our crystal structure may represent the form of DapF that binds the D,L diastereoisomer of DAP. The potential involvement of Arg78 and Asn75 have hitherto remained unreported. The roles played by these residues and Glu208 in the steric control exerted by DapF is currently being pursued in this laboratory.

An active-site inhibitor of *E. coli* DapF has been identified (Higgins *et al.*, 1989). 2-(4-Amino-4-carboxybutyl) aziridine-2-carboxylate (azi-DAP) was shown to irreversibly alkylate Cys73, suggesting a direct role for this residue in the active site. Whilst some work has been published (Wiseman & Nichols, 1984; Koo & Blanchard, 1999) on the mechanism of the enzyme, no structures exist for active-site mutants or substrate/inhibitor complexes, which would greatly assist in the elucidating the catalytic mechanism. We have recently synthesized an oxa analogue of azi-DAP (Cox *et al.*, 2002) which may provide an alternative route to such a complex. We are currently working toward determining a structure of the enzyme–inhibitor complex to shed further light on the mechanism of these enzymes. Studies may well furnish new DapF inhibitors which possess antimicrobial activity.

The authors would like to thank the staff of the Daresbury Synchrotron Source for data-collection facilities and members of the York Structural Biology Laboratory for help, advice and support. We thank Miss A. Blewett for critical reading of the manuscript. DIR was supported by a Medical Research Council Career Development Fellowship. The York Structural Biology Laboratory is supported by the Biotechnology and Biological Sciences Research Council.

References

Antia, M., Hoare, D. S. & Work, E. (1957). *Biochem. J.* **65**, 448–459.
Bugg, T. D. H. (1999). *Comprehensive Natural Products Chemistry*, Vol. 3, edited by M. Pinto, pp. 241–294. Oxford: Elsevier.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Cirilli, M., Zheng, R., Scapin, G. & Blanchard, J. S. (1998). *Biochemistry*, **37**, 16452–16458.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cox, R. J. (1996). *Nat. Prod. Rep.* **13**, 29–43.
Cox, R. J., Durston, J. & Roper, D. I. (2002). *J. Chem. Soc. Perkin Trans. I*, pp. 1029–1039.
Creighton, T. E. (1984). *Proteins: Structures and Molecular Properties*, 2nd ed., pp. 171–199. New York: W. H. Freeman & Co.
Higgins, W., Tardif, C., Richard, C., Krivanek, M. A. & Cardin, A. (1989). *Eur. J. Biochem.* **186**, 137–145.
Hwang, K. Y., Cho, C. S., Kim, S. S., Yu, Y. G. Y. & Cho, Y. (1999). *Nature Struct. Biol.* **6**, 422–426.
Johnston, M. M. & Dixen, W. F. (1969). *J. Biol. Chem.* **244**, 5414–5420.
Koo, C. W. & Blanchard, J. S. (1999). *Biochemistry*, **38**, 4416–4422.
Koo, C. W., Sutherland, A., Vederas, J. C. & Blanchard, J. S. (2000). *J. Am. Chem. Soc.* **112**, 6122–6123.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lui, L., Iwata, K., Kita, A., Kawarabayasi, Y., Yohda, M. & Miki, K. (2002). *J. Mol. Biol.* **319**, 474–489.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.
Nakajima, N., Tanizawa, K., Tanaka, H. & Soda, K. (1986). *Agric. Chem. Biol.* **50**, 2823–2830.
Oldfield, T. J. (1994). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 15–18. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Thompson, J. D., Higgins, G. D. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Wiseman, J. S. & Nichols, J. S. (1984). *J. Biol. Chem.* **259**, 8907–8194.