Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase from *Pseudomonas putida*

B. J. Bell,[a] L. Watanabe,[b] J. L. Rios-Steiner,[a]† A. Tulinsky,[a] L. Lebioda[c] and R. K. Arni[b]*

[a]Department of Chemistry, Michigan State University, East Lansing, MI 48824, USA, [b]Department of Physics, IBILCE/UNESP, R. Cristóvão Colombo 2265, São José do Rio Preto-SP, CEP 15054-000, Brazil, and [c]Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC 29208, USA

† Present address: Department of Chemistry, UPR, Mayaguez, PR 00681-9019, Puerto Rico.

Correspondence e-mail: arni@df.ibilce.unesp.br

Received 10 March 2003
Accepted 12 June 2003

PDB Reference: KDPG aldolase, 1mxs, r1mxssf.

2-Keto-3-deoxy-6-phosphogluconate (KDPG) aldolase from *Pseudomonas putida* is a key enzyme in the Entner–Doudoroff pathway which catalyses the cleavage of KDPG *via* a class I Schiff-base mechanism. The crystal structure of this enzyme has been refined to a crystallographic residual $R = 17.1\%$ ($R_{\rm free} = 21.4\%$). The N-terminal helix caps one side of the torus of the $(\beta\alpha)_8$-barrel and the active site is located on the opposite, carboxylic side of the barrel. The Schiff-base-forming Lys145 is coordinated by a sulfate (or phosphate) ion and two solvent water molecules. The interactions that stabilize the trimer are predominantly hydrophobic, with the exception of the cyclically permuted bonds formed between Glu132 OE1 of one molecule and Thr129 OG1 of a symmetry-equivalent molecule. Except for the N-terminal helix, the structure of KDPG aldolase from *P. putida* closely resembles the structure of the homologous enzyme from *Escherichia coli*.

1. Introduction

2-Keto-3-deoxy-6-phosphogluconate aldolases (KDPG aldolases; EC 4.1.2.14) play key roles in the Entner–Doudoroff pathway (Entner & Doudoroff, 1952), a variation of the Embden–Meyerhoff pathway, which was first discovered in *Pseudomonas saccharophil* (Entner & Doudoroff, 1952; MacGee & Doudoroff, 1954). The Entner–Doudoroff pathway encountered in Gram-negative bacteria probably predates the Embden–Meyerhoff–Parnas pathway (Romano & Conway, 1996). In the Entner–Doudoroff pathway, glucose-6-phosphate is converted by a dehydrase to 6-phosphogluconate, which loses water to form KDPG. Subsequently, KDPG aldolase cleaves KDPG, forming pyruvate and glyceraldehyde-3-phosphate; the latter is converted to pyruvate in the normal way. In the reverse reaction, KDPG aldolase catalyses the addition of pyruvate to D-glyceraldehyde phosphate in a highly diastereoselective manner. Unlike other pyruvate aldolases examined to date, the KDPG aldolases accept short-chain non-carbohydrate electrophilic aldehydes as substrates, providing a general methodology for the construction of the 4-hydroxy-2-ketobutyrate skeleton.

KDPG aldolase belongs to the class I aldolases, since it catalyses the cleavage of KDPG *via* a Shiff-base-assisted mechanism. Thus, catalysis proceeds in a similar fashion to that in fructose bisphosphate (FBP) aldolase. A Shiff-base ketimine (azomethine) intermediate is formed by the condensation of the carbonyl group of a sugar and the active-site lysine. KDPG aldolase is relatively non-specific in forming Schiff bases with carbonyl compounds, except for dihydroxyacetone, hydroxypyruvate and 2-ketogluconate.

KDPG and other aldolases have recently received much attention owing to their potential application as catalysts for stereochemically specific carbon–carbon bond formation (Shelton *et al.*, 1996). Aldolases demonstrate high specificity for the donor substrate (*i.e.* the nucleophilic enolate) but are relatively tolerant with respect to the acceptor (*i.e.* electrophilic) group. Thus, aldolases are of importance for the organic synthesis of carbohydrates (Drauz & Waldmann, 1995). These molecules have gained further attention since they serve as intermediates in the synthesis of complex bioactive molecules such as glycosyltransferase and glycosidase inhibitors. FBP aldolase has been used for the construction of cyclic imine sugars which serve as inhibitors of glycosidases and as building blocks for the synthesis of glycosyltransferase inhibitors (Machajewski & Wong, 2000). FBP aldolase has additionally found application in the synthesis of bicyclic sugars and disaccharide mimics (Zannetti *et al.*, 1999). The acetaldehyde-dependent 2-deoxyribose-5-phosphate aldolase (DERA) catalyses the condensation between two aldehydes and has been used for the synthesis of epothilones (Machajewski & Wong, 2000), which belong to a new class of anticancer agents. The glycine-dependent D- and L-threonine aldolases produce β-hydroxy-α-amino acids that are important components for the synthesis and production of numerous natural products (Machajewski & Wong, 2000). Transaldolases and transketo-aldolases have been used for catalyzing C—C bond formation. The activities of enzymes such as fucosidase and fucosyl-transferase are implicated in inflammation in cancer and in other diseases.

The crystal structure of KDPG aldolase from *Escherichia coli* has been determined at 1.95 Å (Allard *et al.*, 2001). Recently, directed evolution has been employed to engineer a new active site in the KDPG aldolase from *E. coli* in which the location of the Shiff-base-forming lysine has been moved to a neighboring β-strand (Wymer *et al.*, 2001). The modified enzyme was demonstrated to be catalytically more efficient with altered specificity and was able to catalyze the synthesis of both D- and L-sugars from non-phosphorylated aldehydes and pyruvate (Fong *et al.*, 2000).

The KDPG aldolases from *P. putida* and *E. coli* share approximately 49% sequence identity. Here, we report the crystal structure of the KDPG aldolase from *P. putida* refined to a crystallographic residual of R = 17.1% ($R_{\rm free}$ = 21.4%).

2. Experimental

2.1. Crystallization

Briefly, crystals of KDPG aldolase from *P. putida* were obtained by microdialysis of protein solution at 15 mg ml^{-1} against sequentially higher concentrations of ammonium sulfate to a final concentration of 1.5 *M* ammonium sulfate, 0.1 *M* KH_2PO_4 pH 3.5 in plexiglass buttons (Vandlen *et al.*, 1973). The crystals were subsequently stored in the aforementioned solution for over 14 years.

2.1.1. Data collection. Attempts to stabilize these crystals in cryoprotectants were unsuccessful. Over 30 crystals were tested, but they were either very highly mosaic or decayed rapidly when exposed to X-rays. Finally, a single-crystal (minimum dimension of 0.7 mm) was mounted in a capillary and we were able to collect diffraction data at room temperature (296 K). Monochromatic Cu $K\alpha$ radiation was produced by a Rigaku RU-200 rotating-anode generator operating at 50 kV and 120 mA equipped with Yale/MSC mirrors and a helium beam path. Diffraction intensities were measured utilizing a Rigaku R-AXIS IIC imaging-plate detector and the data were reduced to a maximum resolution of 2.2 Å using the program *PROCESS* (Higashi, 1990). The completeness of the high-resolution data is low owing to reflection overlap and crystal decay. A plot of completeness as a function of $\sin\theta/\lambda$ is shown in Fig. 1. The data-processing statistics are presented in Table 1. The crystals belong to the cubic system, space group $P2_13$, with a unit-cell constant of 103.2 Å.

Table 1
Diffraction data-collection and refinement statistics for KDPG aldolase from *P. putida*.

Space group	$P2_13$
Unit-cell parameter (Å)	103.2
$V_{\rm M}$ (Å^3 Da^{-1})	3.25
Solvent content (%)	61.8
$R_{\rm sym}$† (%)	4.6
Completeness (%)	52.3
No. of observations	66178
No. of unique reflections	9901
Refinement resolution (Å)	∞–2.2
R factor‡ (%)	17.1
$R_{\rm free}$§ (%)	21.4
No. of solvent molecules	56
Average B values¶ (Å^2)	33.3
R.m.s.d. from ideal values††	
Bond length (Å)	0.006
Bond angle (°)	1.2

† $R_{\rm sym} = 100 \times \sum |I(h) - \langle I(h)\rangle| / \sum I(h)$, where $I(h)$ is the observed intensity and $\langle I(h)\rangle$ is the mean intensity of reflection h over all measurements of $I(h)$. ‡ $R_{\rm factor} = 100 \times \sum |F_o - F_c| / \sum (F_o)$, the sums being taken over all reflections with $F/\sigma(F) > 0$ cutoff. § $R_{\rm free}$ = R factor for 10% of the data which were not included during crystallographic refinement. ¶ B values are average B values for all non-H atoms. †† R.m.s.d., root-mean-square deviation.

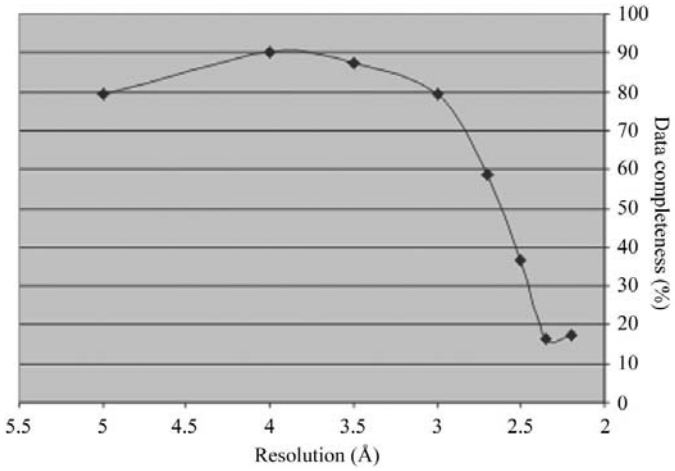

Figure 1
Plot of data completeness (%) *versus* resolution (Å). The first point corresponds to the range ∞–5 Å, the second 5–4 Å *etc*.

2.1.2. Structure determination and refinement. The structure was solved by molecular replacement with the program *AMoRe* (Navaza, 1994) as implemented in the *CCP*4 package, using the atomic coordinates of the *E. coli* enzyme (Allard *et al.*, 2001). To produce the search model, non-conserved amino acids were mutated to glycine residues and solvent molecules were deleted. Although there was no outstanding peak in the results of the rotation search, the 16th peak of the rotation search provided a clear solution in the translation search, with a correlation coefficient of 24.5 and an *R* factor of 49.5% for data between 20.0 and 4.0 Å. Rigid-body refinement of this solution resulted in a correlation coefficient of 31.6 and an *R* factor of 46.8% for data in the resolution range 7.0–3.5 Å.

The electron-density maps were examined and the amino-acid sequence of KDPG aldolase from *P. putida* was introduced. The atomic model was periodically rebuilt using the *TURBO-FRODO* software and the structure was refined using simulated-annealing and *B*-factor refinement as implemented in the program *CNS* (Brünger *et al.*, 1998). The electron-density maps provided good orientation of the sulfate ions, especially considering the low completeness at high resolution of the data (Fig. 4). The solvent structure was obtained using the *CNS* software and from visual inspections of electron-density maps. Root-mean-square differences from ideal geometry for the atomic bond lengths and bond angles are presented in Table 1. The overall stereochemical quality of the final model was assessed by the program *PROCHECK* (Laskowski *et al.*, 1993). The only Ramchandran outlier is Arg28, which is located at the end of the N-terminal helix, is exposed to the solvent and is characterized by poor electron density. To enable structural comparisons, the atomic models of the enzyme from *P. putida* and *E. coli* were superposed using the program *WHATIF* (Vriend, 1990).

Figure 2
View along the axis of the β-barrel of KDPG aldolase from *P. putida*. The N-terminal helix is colored differently. Lys145 and the sulfate ions S1 and S2 are included to indicate the position of the active site. Figs. 2, 4 and 5 were generated with *PYMOL* (DeLano, 2002).

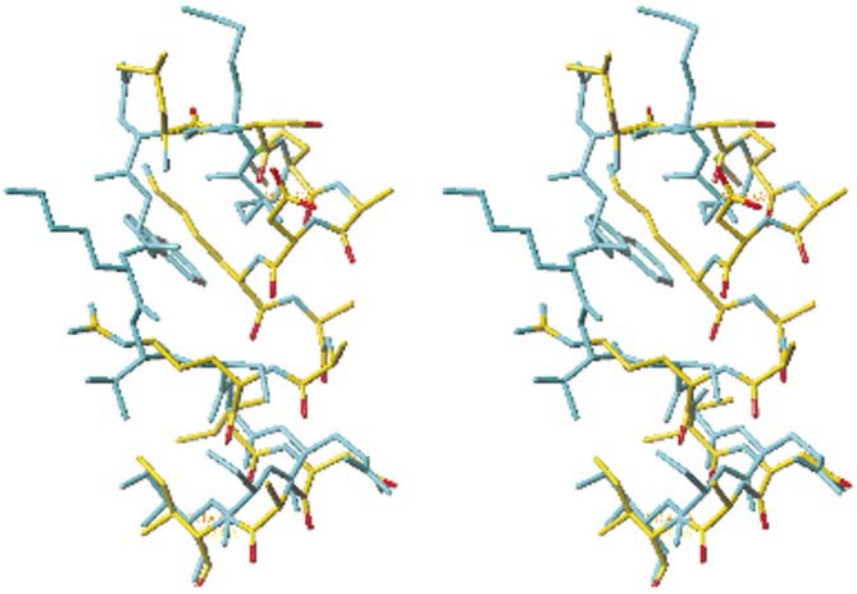

Figure 3
Superposition of the N-terminal segments of KDPG aldolase from *P. putida* (residues 11–23, in yellow) and *E. coli* (residues 1–11, in blue) shows unwinding of part of the N-terminal helix in the latter. Figs. 3 and 6 were generated with *TURBO-FRODO*.

2.2. Results and discussion

The crystal structure of KDPG aldolase from *P. putida* was one of the early enzyme structures to be determined, originally at 3.5 Å resolution (Mavridis & Tulinsky, 1976) and subsequently at 2.8 Å resolution (Mavridis *et al.*, 1982). The PDB currently only holds the positions for the Cα atoms obtained by interpretation of electron-density maps utilizing a Richards Box (PDB code 1kga). We present here a more complete structure refined at a higher resolution. The shape of the monomer of KDPG aldolase can be described as an oblate ellipsoid (Fig. 2) and the enzyme belongs to the β-barrel $(\beta\alpha)_8$ class of enzymes. The central cavity of the barrel has a diameter of approximately 6 Å, while the diameter and height of the β-barrel are approximately 40 and 25 Å, respectively. The first nine N-terminal amino acids could not be located in the electron-density maps. Amino acids 10–26 form a four-turn α-helix (helix I) which is almost perpendicular to the axis of the torus of the β-barrel and caps its side, thus preventing access to the N-terminal part of the barrel cavity.

An alignment of the amino-acid sequences of the aldolases from *P. putida* and *E. coli* yields 49% identity with no gaps in the structurally conserved region, Arg20–Ala225 in the enzyme from *P. putida* and Ser7–Leu213 in the *E. coli* enzyme. Superpositioning the two structures results in an r.m.s. distance of 0.96 Å between the

equivalent Cα positions, with a maximum deviation of 2.32 Å. The largest structural difference is at the N-terminus (Fig. 3), where residues Leu11–Arg19 in the *P. putida* enzyme form two additional turns of helix I, whereas residues Met1–Ser7 in *E. coli* enzyme form a random coil located in the same region of the molecule. Post-translationally processed *P. putida* KDPG aldolase has 11 residues more than the *E. coli* enzyme at the N-terminus; however, most of these amino acids could not be traced in the electron-density maps, probably owing to disorder. The additional residue at the C-terminus of the *P. putida* enzyme, Asn226, also could not be located in the maps.

The active site is located on the opposite side of the barrel in a shallow depression; it is freely accessible to solvent and is formed by the highly conserved residues Glu57, Arg61, Thr85, Val104, Pro106 and Lys145. As expected, similarity between the *E. coli* and *P. putida* structures is closest in the active-site region. Two sulfate ions were present in the *E. coli* structure with an S–S distance of ∼8 Å (Allard *et al.*, 2001). In the *P. putida* structure there also are two tetrahedral ions which form analogous interactions. They correspond to either sulfate or phosphate ions or a mixture of the two, as both species were present in the crystallization solution at concentrations of 1.5 and 0.1 *M*, respectively. A somewhat longer distance between the central atoms, ∼10 Å, suggests that at least one of the sites may be predominantly occupied by phosphate ions, which at pH 3.5 have a different charge and hydrogen-bonding properties to the sulfate ion. The Schiff-base-forming Lys145 is coordinated by one of the sulfate/phosphate ions (S1) and two solvent water molecules (Fig. 4). The second sulfate/phosphate ion (S2) is linked to the latter and to Lys145 N$^{\zeta}$ *via* a solvent bridge formed by three water molecules; it is likely that this binding site is weak and an artifact of crystallization.

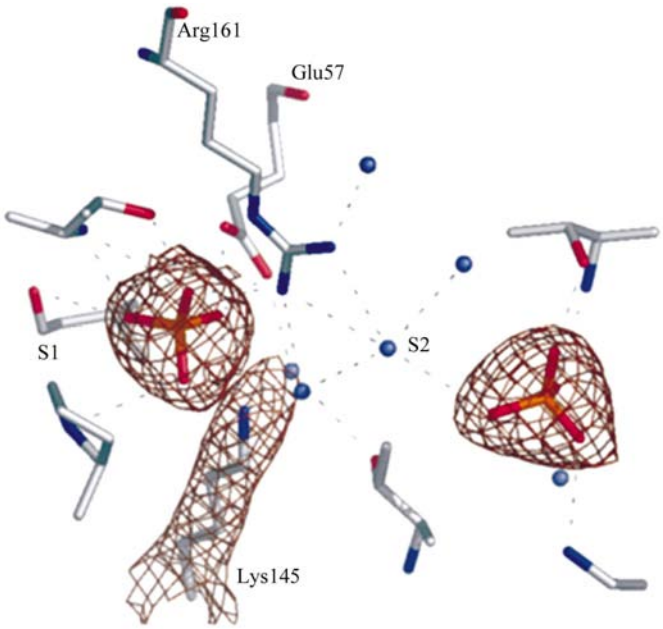

Figure 4
The interactions in the active site of KDPG aldolase from *P. putida* are illustrated. For clarity, the electron density (contoured at 1σ) is shown only for Lys145 and the sulfate ions S1 and S2. Hydrogen bonds are indicated by broken lines.

It has been proposed that only residues from one subunit participate in catalysis (Allard *et al.*, 2001), but that other subunits are involved in supporting the active-site framework. Currently, no data on the activity of monomers are available as the conditions required to disrupt the trimer also denature the protein. The asymmetric unit contains one subunit and space-group symmetry operations generate physiological trimers. The threefold-symmetry axis of the trimer (Fig. 5) is only approximately orthogonal to the axis of the β-barrel, so that the molecule resembles a ship's propeller (Mavridis *et al.*, 1982). About 1230 Å^2 per subunit or 13% of the protomer surface is buried upon trimer formation. Most of the buried

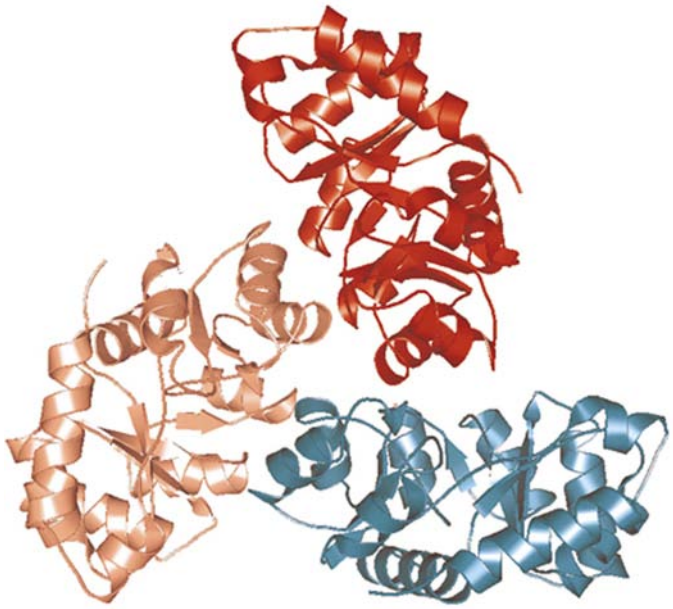

Figure 5
Trimer of KDPG aldolase showing the propeller-like arrangement of subunits.

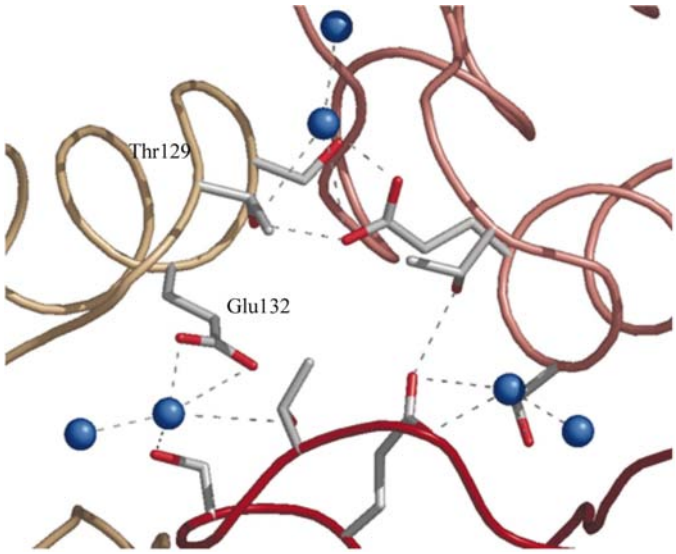

Figure 6
Interactions between symmetry-related molecules that stabilize the trimer.

surface is hydrophobic. An exception to this is a cluster of three symmetry-equivalent glutamic acid residues Glu132 that are involved in the stabilization of the trimer (Fig. 6). The buried cluster of three Glu132 residues presents an interesting aspect of the structure that unfortunately cannot be fully addressed at the pH at which the data were collected (pH 3.5). The glutamic acid is conserved in the *E. coli* enzyme and also forms a cluster of three carboxylates directed towards a solvent molecule or ion. This structure has also been determined at low pH (4.5), where the carboxylates may be protonated. On the other hand, the solvent peak may be an ammonium ion that was present at a very high concentration. It is not unlikely that at physiological pH the charges of the buried Glu132 residues are balanced by a metal ion binding at this position and stabilizing the trimer, but this hypothesis has to be verified.

KDPG aldolases functioning in the Entner–Doudoroff pathway catalyze the addition of pyruvate to D-glyceraldehyde phosphate in a highly diasteroselective manner. A comparison of the crystal structures of the enzyme from *P. putida* and *E. coli* demonstrates that they are very similar and that the enzyme from either organism can be a potentially useful industrial biocatalyst.

We wish to thank Dr J. Sygusch for providing us with the atomic coordinates of the *E. coli* enzyme prior to publication. This work was supported in part by NSF grant MCB-9873606 to LL and by grants from FAPESP (99/0961-4), FAPESP/SMOLBNet (01/07537-2) and CNPq (520081/95-1 NV). BJB is supported by a NIH GM65062 fellowship; LW is the recipient of a FAPESP fellowship.

References

Allard, J., Grochulski, P. & Sygusch, J. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 3679–3684.
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.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA.
Drauz, K. & Waldmann, H. (1995). *Enzyme Catalysis in Organic Synthesis.* Weinheim: Wiley.
Entner, N. & Doudoroff, M. J. (1952). *J. Biol. Chem.* **196**, 853–862.
Higashi, T. (1990). *J. Appl. Cryst.* **23**, 253–257.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Fong, S., Machajewski, T. D., Mak, C. C. & Wong, C. H. (2000). *Chem. Biol.* **11**, 873–883.
MacGee, J. & Doudoroff, M. J. (1954). *J. Biol. Chem.* **210**, 617–626.
Machajewski, T. D. & Wong, C. H. (2000). *Angew. Chem. Int. Ed. Engl.* **39**, 1352–1374.
Mavridis, I. M. & Tulinsky, A. (1976). *Biochemistry*, **15**, 4410–4417.
Mavridis, I. M., Hatada, M. H., Tulinsky, A. & Lebioda, L. (1982). *J. Mol. Biol.* **162**, 419–444.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Romano, A. H. & Conway, T. (1996). *Res. Microbiol.* **147**, 448–455.
Shelton, M. C., Cotterill, I. C., Novak, S. T. A., Poonawala, R. M., Sudarshan, S. & Toone, E. C. (1996). *J. Am. Chem. Soc.* **118**, 2117–2125.
Vandlen, R. L., Ersfeld, D. L., Tulinsky, A. & Wood, W. A. (1973). *J. Biol. Chem.* **248**, 2251–2253.
Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.
Wymer, N., Buchanan, L. V., Henderson, D., Metha, N., Botting, C. H., Pocivavsek, L., Fierke, C. A., Toone, E. J. & Naismith, J. H. (2001). *Structure*, **9**, 1–9.
Zannetti, M. T., Walter, C., Knorst, M. & Fessner, W. D. (1999). *Chem. Eur. J.* **5**, 1882–1890.