Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of *Escherichia coli* YfdW, a type III CoA transferase

Arhonda Gogos,[a] Jason Gorman[b] and Lawrence Shapiro[a,b,c]*

[a]Department of Biochemistry and Molecular Biophysics, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032, USA, [b]Department of Ophthalmology, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032, USA, and [c]Naomi Berrie Diabetes Center, Columbia University College of Physicians and Surgeons, 630 West 168th Street, New York, NY 10032, USA

Correspondence e-mail: shapiro@convex.hhmi.columbia.edu

Received 21 August 2003
Accepted 2 January 2004

PDB References: YfdW at 2.35 Å resolution, 1pqy, r1pqysf; YfdW at 1.6 Å resolution, 1q7e, r1q7esf; YfdW–CoA complex, 1q6y, r1q6ysf.

Crystal structures are reported for free and coenzyme A (CoA) bound forms of the YfdW protein from *Escherichia coli*, a representative type III CoA transferase. The structures reveal a two-domain protomer with interdomain connections forming a ring-like structure with a large central hole. Two protomers associate to form a highly intertwined dimer in which the hole of each ring is filled by the partner molecule. Each protomer binds a single CoA molecule and these CoA-binding sites are distant from one another in the dimer.

1. Introduction

CoA transferases, which are found in all free-living organisms, are enzymes that catalyze the transfer of coenzyme A from a donor to an acceptor. Most of these enzymes belong to two well characterized families, the type I and type II CoA transferases (Heider, 2001). Recent work has revealed the existence of a distinct third family of type III enzymes, which have unique amino-acid sequences unrelated to the type I and type II enzymes (Elssner *et al.*, 2001; Heider, 2001). Sequence analysis shows that type III CoA transferases are found in bacteria, archaea and eukarya.

The catalytic activities of some of the known type III CoA transferases have been characterized. They transfer CoA to various small organic molecules from a variety of CoA donors with high specificity. These include a formyl-CoA:oxalate CoA transferase from *Oxalobacter formigenes* (Baetz & Allison, 1990; Sidhu *et al.*, 1997), a succinyl-CoA:(*R*)-benzylsuccinate CoA transferase from *Thaura aromatica* (Leutwein & Heider, 1999, 2001), an (*E*)-cinnamoyl-CoA:(*R*)-phenylacetate CoA transferase from *Clostridium sporogenes* (Dickert *et al.*, 2000) and a butyrobetainyl-CoA:(*R*)-carnitine CoA transferase (CaiB) from *Escherichia coli* and *Proteus* sp. (Eichler *et al.*, 1994; Elssner *et al.*, 2001).

The enzyme studied here, *E. coli* YfdW, has an amino-acid sequence representative of type III CoA transferases and is a member of the this family (PF02515). Neither its biological function nor its catalytic properties have yet been characterized. We determined high-resolution crystal structures of YfdW both in the free form by SAD analysis of the selenomethionyl protein and in complex with CoA by molecular replacement. The enzyme adopts a complex two-domain α/β fold arranged in a ring-like structure. Two monomers interlock to form a highly intertwined dimer, each protomer having a separate CoA-binding site distant from the other. The amino-acid sequence of YfdW is 57% identical to formyl-CoA transferase (FRC) of *O. formigenes*, for which a similar structural study was recently reported (Ricagno *et al.*, 2003). Furthermore, while this manuscript was under review, another

Table 1
Statistics from the crystallographic analysis.

$R_{sym} = I - \sum |I - \langle I \rangle| / \sum I$, where I is the observed intensity and $\langle I \rangle$ is the average intensity. $R_{cryst} = 100 \times \sum \big||F_{obs}| - |F_{calc}|\big| / \sum |F_{obs}|$, where F_{obs} are the observed structure factors and F_{calc} are the calculated structure factors. The crystallographic R factor, R_{cryst}, is based on 95% of the data used in refinement and the free R factor, R_{free}, is based on 5% of the data withheld for the cross-validation test. R.m.s.d., root-mean-square deviation. Over 90% of the main-chain dihedrals fall within the 'most favoured regions' of the Ramachandran plot (Laskowski *et al.*, 1993). Values in parentheses are for the highest resolution shell.

(*a*) Diffraction data statistics.

Crystal	1 (SeMet)	2 (SeMet)	3 (SeMet)	4 (Native + ligand)
Beamline	BNL X9A	APS 31ID	APS 31ID	APS 31ID
Wavelength (Å)	0.9793	0.9793	0.9793	0.9793
Resolution range (Å)	27.88–2.35 (2.43–2.35)	20.0–2.1 (2.17–2.10)	20.0–1.6 (1.66–1.60)	20.0–2.0 (2.07–2.00)
Measured reflections	133632	186497	180448	96770
Unique reflections	18133	25227	56025	27082
Completeness (%)	98.3 (97.8)	98.4 (97.7)	97.5 (95.8)	93.6 (72.3)
R_{sym}	0.093 (0.18)	0.166 (0.284)	0.098 (0.295)	0.066 (0.265)
$\langle I \rangle / \langle \sigma(I) \rangle$	24.3 (14.9)	15.6 (9.4)	11.8 (4)	10.5 (5)

(*b*) Refinement statistics.

Crystal	1 (SeMet)	3 (SeMet)	4 (Native + ligand)
Resolution range (Å)	27.99–2.35	20.0–1.60	20.0–2.0
No. reflections (observed)	18078	53302	25765
No. reflections (R_{free})	877	2723	1317
R factor/R_{free}	0.177/0.211	0.156/0.186	0.160/0.207
R.m.s.d. bond lengths (Å)	0.009	0.012	0.02
R.m.s.d. bond angles (°)	1.4	1.4	2.0

X-ray crystallographic study of *E. coli* YfdW in crystal forms different from ours was reported (Gruez *et al.*, 2003). This study examined YfdW in the apo form, in complex with acetyl-CoA and in complex with acetyl-CoA and oxalate.

2. Materials and methods

The *yfdW* gene was cloned into the pSGX-1 expression vector, which encodes a C-terminal six-His tag. Selenomethionine (SeMet) substituted protein was expressed in *E. coli* BL21 cells and purified to homogeneity by affinity chromatography and gel filtration. The protein was concentrated to 13.5 mg ml^{-1} in 10 m*M* HEPES pH 7.5, 150 m*M* NaCl, 10 m*M* methionine and 10% glycerol. Crystals were obtained by vapour diffusion in 1 µl hanging drops that contained 0.6 µl protein solution and 0.4 µl well solution containing 70% 2-methyl-2,4-pentanediol (MPD), 0.1 *M* bis-Tris pH 6.5 at 277 K. The crystals belong to space group *C*2, with unit-cell parameters a = 92.4, b = 69.6, c = 73.2 Å, β = 108.6° and one molecule per asymmetric unit.

Crystals of the liganded form were obtained after incubation of the native protein with 3 m*M* CoA and 3 m*M* oxalate, using a well solution consisting of 65% MPD and 0.1 *M* bis-Tris pH 6.5 at 277 K. They are essentially isomorphous to the unliganded form, belonging to space group *C*2, with unit-cell parameters a = 92.1, b = 66.8, c = 73.1 Å, β = 108.7° and one molecule per asymmetric unit.

The crystals were flash-frozen at 100 K without addition of cryoprotectant. Data were collected at beamlines X9A of the National Synchrotron Light Source (NSLS) and 31ID of the Advanced Photon Source. The data were processed and merged using the *HKL* program suite (Otwinowski & Minor, 1997). Se-atom positions were located using *SOLVE* (Terwilliger & Berendzen, 1999), refined with *SHARP* (de La Fortelle & Bricogne, 1997) and the first model was built with the program *RESOLVE* (Terwilliger, 2000). Initial refinement was performed with *CNS* (Brünger *et al.*, 1998) and subsequently with *REFMAC*5.0 (Murshudov *et al.*, 1997) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Figures were produced with the program *SETOR* (Evans, 1993).

Crystals 1 and 2 in Table 1 were used in phasing. Although data set 2 had the better resolution, initial refinement was performed against data set 1 because of its better data statistics. This structure was deposited in the Protein Data Bank (PDB) with accession code 1pqy. We subsequently collected data from crystal 3 (Table 1), which was used for the

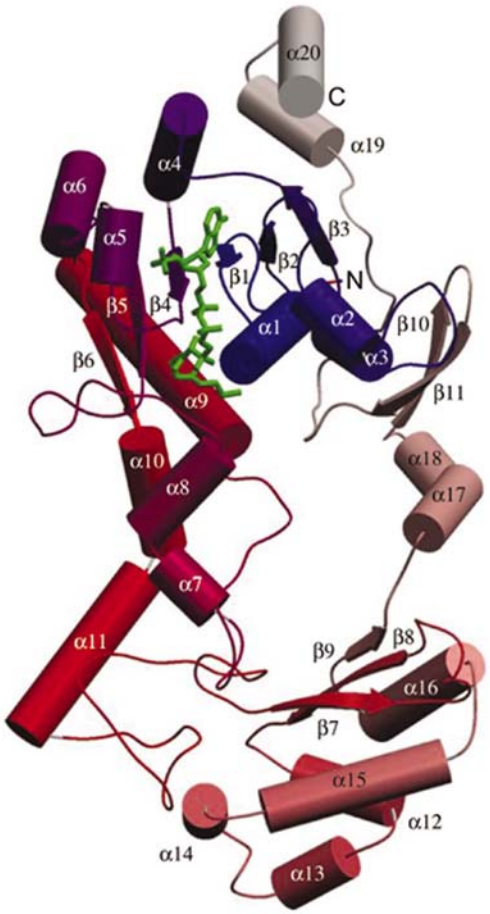

Figure 1
Ribbon diagram of *E. coli* YfdW protein protomer, color-coded blue to white from the NH_2- to the COOH-terminus. Helices are shown as cylinders and the NH_2- and COOH-termini are indicated. CoA is shown in green in an all-atom representation.

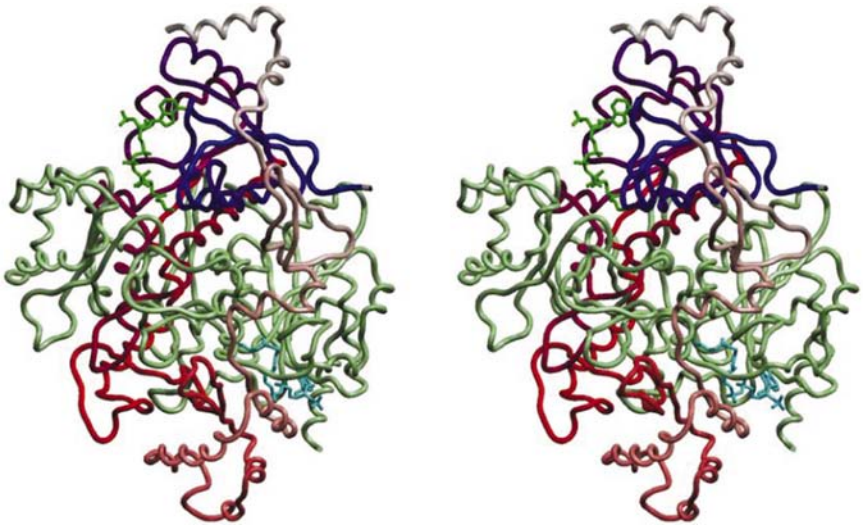

Figure 2
C^{α} worm diagram of YfdW dimer. One protomer is color-coded from the NH_2- to the COOH-terminus and the other is shown in green. The two CoA molecules bound to the protomers are shown in green and blue.

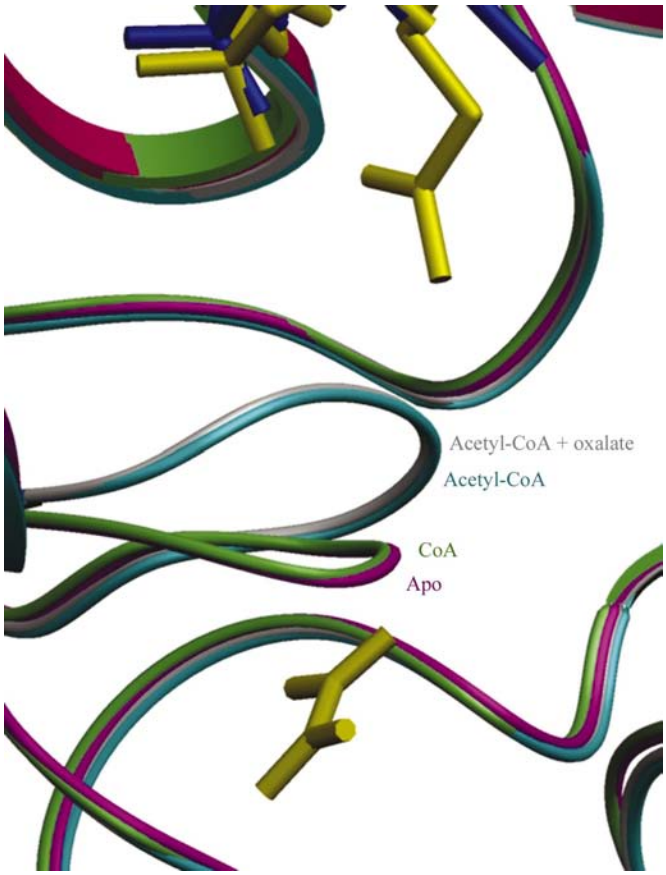

Figure 3
Superposition of four YfdW structures: (i) apo form, shown in magenta, (ii) bound to CoA, shown in green, (iii) bound to acetyl-CoA (Gruez *et al.*, 2003), shown in cyan, and (iv) bound to acetyl-CoA and oxalate (Gruez *et al.*, 2003), shown in gray. CoA and acetyl-CoA are partly shown at the top of the picture, in blue and yellow, respectively. Oxalate is depicted in yellow at the bottom of the picture in an all-atom representation. Note the shift in conformation of the glycine-rich loop apparently induced by acetyl-CoA.

high-resolution refinement of the free enzyme (PDB code 1q7e). Crystal 4 was used in refinement of the CoA–enzyme complex (PDB code 1q6y).

3. Results and discussion

We determined the structure of the free form of YfdW from *E. coli* using phases derived from SAD experiments performed on two different crystals of SeMet-substituted protein. The final model was refined against a third data set to 1.6 Å to an *R* factor of 15.6%. This structure was used as a search model for determination of the CoA-bound structure, which was refined to 2.0 Å resolution, to an *R* factor of 16.0%. Both models comprise residues 2–416. The region 100–106 appears to be disordered in the structure of the free enzyme and is only observed in the liganded form.

3.1. Structure description

The YfdW protein is a highly intertwined dimer in the crystal. Each protomer consists of two large domains joined together by two linker regions at distal ends of the domains, thus forming a ring-like structure with a large hole in the center (Fig. 1). Two of these ring-like protomers interlock, each filling the hole of the other (Fig. 2).

The C-terminal domain of one protomer, consisting of a helix, a two-stranded β-sheet and two α-helices, folds around the other protomer to form a clamp that is stabilized by many interactions with both the partner molecule and the large N-terminal domain of its own protomer. This highly intertwined arrangement suggests that substantial conformational changes are necessary in the assembly of the dimer. Consistent with this observation, biochemical evidence suggests that type III CoA transferases function as dimers (Heider, 2001).

In the dimer, the large N-terminal domain, which includes a Rossmann fold, interacts with the small C-terminal domain of the other protomer. A loop rich in glycine residues (amino acids 246–249) belongs to this small domain. In the FRC structure (Ricagno *et al.*, 2003), this loop is seen in an 'open' conformation in the free state and a 'closed' conformation in the liganded state. In the closed conformation of FRC, this loop caps the putative active site and it has been proposed that this capping may be important to substrate binding (Ricagno *et al.*, 2003). In the YfdW structures reported here, the glycine-rich loop appears to be in the open conformation irrespective of ligand-binding state. Comparison with the YfdW structures determined by Gruez *et al.* (2003) reveal nearly identical structures, which differ significantly only by their adoption of the 'open' or 'closed' conformation for the glycine-rich loop of the active site (Fig. 3). Although the apo and CoA-bound enzymes adopt the 'open' conformation, acetyl-CoA and acetyl-CoA/oxalate complexes are found in the 'closed'

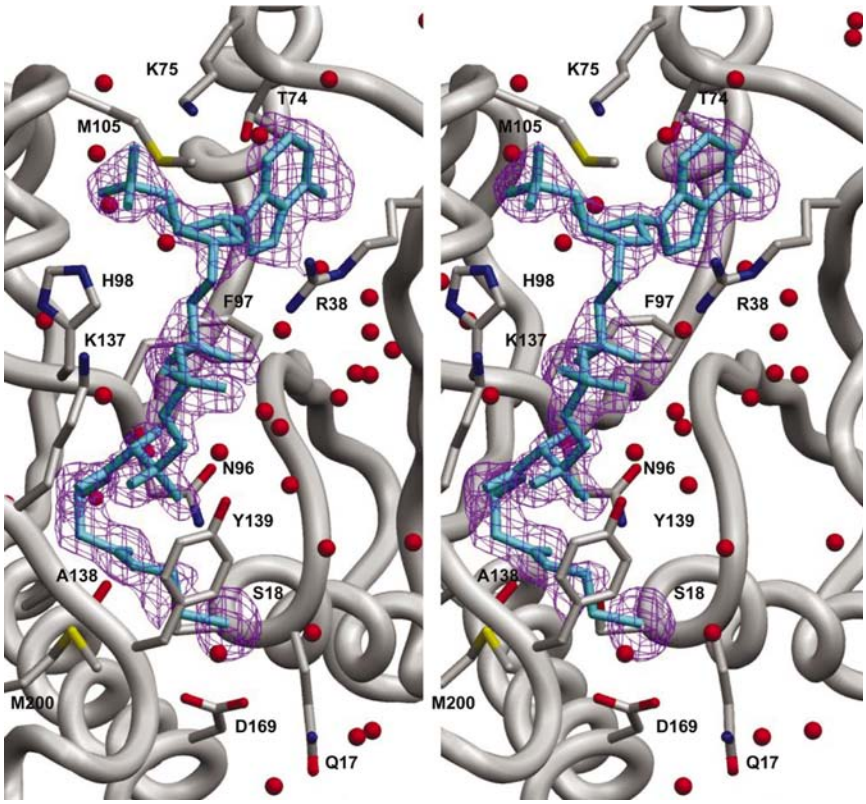

Figure 4
Binding of CoA to the putative active site of YfdW. A CoA omit map contoured at 2σ is shown in magenta. CoA is depicted in blue in an all-atom representation and the water molecules are shown as red spheres. All CoA interactions are within the large domain. The adenine moiety is positioned between the side chains of Arg38, Thr74, Phe97 and Met105. There are hydrogen-bonding interactions between adenine N6 and the carbonyl O atom of Leu72 and a water-mediated hydrogen bond between adenine N7 and the carbonyl O atom of Ile36. For the ribose moiety, O4′ forms a hydrogen bond to Arg38 NH1 and O2′ forms a water-mediated hydrogen bond to the carbonyl O atom of Ala101. The ribose phosphate moiety forms hydrogen bonds to His98 $N^{\varepsilon 2}$ through O9 and O7 and to Lys75 N^{ζ} through O8. The binding of the CoA pantetheine chain is stabilized by van der Waals interactions with Met200 and several hydrogen-bonding interactions. AO1 of the pyrophosphate moiety hydrogen bonds to Arg38 $N^{\eta 2}$, AO4 to the hydroxyl of Tyr139 through two water molecules and AO5 to Lys137 N^{ζ}. The pantetheine hydroxyl group hydrogen bonds to the amide of His98 and the pantetheine amino PN4 interacts with the carbonyl O atom of Ala138 and PN8 with the carbonyl O atom of Asn96. The pantetheine carbonyl O atom PO5 interacts with Asn96 ND2, which is stabilized by a hydrogen bond to Ser18 OG. The pantetheine SH group PS1 hydrogen bonds to the amide of Asn17 and is also in close contact with Asp169 OD2. The SH group also appears to be stabilized by a helix dipole interaction with the N-terminus of α-helix 1.

conformation. This shift appears to originate from interaction of the acetyl group with the glycine-rich loop.

3.2. CoA binding at putative active sites

The two CoA-binding sites, which mark the putative active sites, are far from each other in the dimer. No large conformational changes are observed between the free and CoA-bound structures, except that helix α5 (amino acids 100–106), which is disordered in the free structure, becomes visible in the liganded form. This differs from the FRC structure in which helix α5 can be seen in both liganded and free states (Ricagno *et al.*, 2003) and the structure of YfdW in a different crystal lattice reported by Gruez *et al.* (2003). In our YfdW structure, helix α5 appears to be stabilized upon ligand binding through a hydrophobic interaction between the adenine ring of CoA and the side chain of Met107. In the absence of CoA, a free methionine molecule, a component of the protein solution used as a reductant, is bound at each CoA-interaction site.

FRC, the nearest biochemically characterized homolog of YfdW, is known to function in the transfer of CoA to an oxalate acceptor (Baetz & Allison, 1990). For this reason, we included oxalate in the crystallization solutions for the liganded form. However, no electron density corresponding to oxalate could be located. This recapitulates the observation for FRC, where oxalate also could not be found in electron-density maps (Ricagno *et al.*, 2003). However, in the YfdW structure reported by Gruez *et al.* (2003) the oxalate ion can be placed in density, but it is far away from the carbonyl groups of acetyl CoA. Thus, these authors have described the oxalate as adopting a 'resting position' that is not reflective of its position in catalysis. It should be noted, however, that neither the CoA donor nor the acceptor has been identified for YfdW experimentally. CoA-binding interactions are described in detail in Fig. 4.

The catalytic mechanism of characterized CoA transferases involves the formation of a thioester intermediate between the SH group of CoA and an acidic side chain of the enzyme (Dickert *et al.*, 2000; Leutwein & Heider, 2001). In YfdW, the SH group of CoA is positioned near the side-chain carboxylate of Asp169 and we suggest that this residue is likely to be the target of thioester formation. Details of the catalytic mechanism, however, must await further experimental studies.

We gratefully acknowledge Thirumuruhan Radhakannan at beamline X9A of the National Synchrotron Light Source (NSLS) and Kevin D'Amico at beamline 31ID of the Advanced Photon Source for assistance with data collection. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. W-31-109-Eng-38. This is a contribution of the New York Structural Genomics Research Consortium (NIH S/G IP50 GM62529).

References

Baetz, A. L. & Allison, M. J. (1990). *J. Bacteriol.* **172**, 3537–3540.
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.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Dickert, S., Pierik, A. J., Linder, D. & Buckel, W. (2000). *Eur. J. Biochem.* **267**, 3874–3884.
Eichler, K., Schunck, W. H., Kleber, H. P. & Mandrand-Berthelot, M. A. (1994). *J. Bacteriol.* **176**, 2970–2975.
Elssner, T., Engemann, C., Baumgart, K. & Kleber, H. P. (2001). *Biochemistry*, **40**, 11140–11148.
Evans, S. V. (1993). *J. Mol. Graph.* **11**, 134–138.
Gruez, A., Roig-Zamboni, V., Valencia, C., Campanacci, V. & Cambillau, C. (2003). *J. Biol. Chem.* **278**, 34582–34586.
Heider, J. (2001). *FEBS Lett.* **509**, 345–349.
La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Leutwein, C. & Heider, J. (1999). *Microbiology*, **145**, 3265–3271.
Leutwein, C. & Heider, J. (2001). *J. Bacteriol.* **183**, 4288–4295.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Ricagno, S., Jonsson, S., Richards, N. & Lindqvist, Y. (2003). *EMBO J.* **22**, 3210–3219.
Sidhu, H., Ogden, S. D., Lung, H. Y., Luttge, B. G., Baetz, A. L. & Peck, A. B. (1997). *J. Bacteriol.* **179**, 3378–3381.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.