Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary analysis of native and N-terminal truncated isoforms of toluene-4-monooxygenase catalytic effector protein

Allen M. Orville,[a]* Joey M. Studts,[b] George T. Lountos,[a] Kevin H. Mitchell[b] and Brian G. Fox[b]

[a]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA 30332-0400, USA, and [b]Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin, Madison, WI 53706-1544, USA

Correspondence e-mail: allen.orville@chemistry.gatech.edu

Received 27 September 2002
Accepted 3 January 2003

Single crystals have been obtained of the toluene 4-monooxygenase catalytic effector protein, the SeMet-enriched protein and a truncated isoform missing ten amino acids from the N-terminus. Complete X-ray diffraction data sets have been collected and analyzed to 2.0, 3.0 and 1.96 Å resolution for the native, SeMet and truncated isoform crystals, respectively. The native and SeMet proteins crystallized in space group $P6_122$ (unit-cell parameters $a = b = 86.41 \pm 0.15$, $c = 143.90 \pm 0.27$ Å), whereas the truncated isoform crystallized in space group $P2_13$ ($a = b = c = 86.70 \pm 0.47$ Å). Matthews coefficient calculations suggest either two or three molecules per asymmetric unit in the $P6_122$ space group and two molecules per asymmetric unit in the $P2_13$ space group. Experimental phases from MAD analysis of the SeMet isoform and molecular replacement of the truncated isoform confirm the presence of two molecules per asymmetric unit in each case. These crystallographic results are the first available for the evolutionarily related but functionally diversified catalytic effector proteins from the multi-component diiron monooxygenase family.

1. Introduction

Toluene 4-monooxygenase (T4MO) from *Pseudomonas mendocina* KR1 is a multi-component enzyme complex (Whited & Gibson, 1991; Pikus *et al.*, 1996) that exhibits remarkably high regiospecificity for the NADH- and O_2-dependent hydroxylation of toluene to form *para*-cresol (Fig. 1). The T4MO enzyme complex consists of an NADH oxidoreductase (T4moF, 33 kDa), a Rieske-type ferredoxin (T4moC, 12.5 kDa), a catalytic effector protein (T4moD, 11.6 kDa) and a diiron-containing hydroxylase (T4moH, 212 kDa) with a $(\alpha\beta\gamma)_2$ quaternary structure. T4MO is a member of an evolutionarily related family of oxygenases that includes four subgroups distinguished by their specificity for different natural substrates (Hemmi *et al.*, 2001; Merkx *et al.*, 2001). Methane monooxygenase is the best characterized member of this family and crystal structures are available for hydroxylase components from *Methylococcus capsulatus* Bath and *Methylosinus trichosporium* OB3b (Rosenzweig *et al.*, 1993; Elango *et al.*, 1997). Previous single-turnover and peroxide-shunt results have shown that the hydroxylase diiron center is the unique site of O_2 reactivity (Andersson *et al.*, 1991). However, protein–protein interactions involving the catalytic effector protein cause changes in the spectroscopic features of the hydroxylase diiron center (Fox *et al.*, 1991), changes in the lifetimes of reactive intermediates (Chang *et al.*, 2001) and changes in the product distributions observed from substrates capable of yielding more than one product (Froland *et al.*, 1992; Mitchell *et al.*,

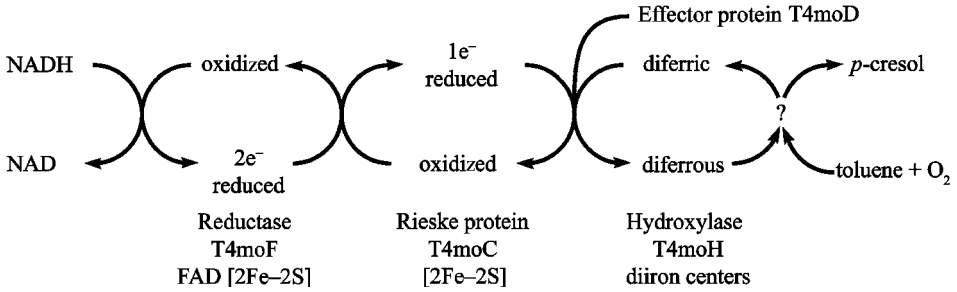

Figure 1
Toluene 4-monooxygenase enzyme complex. The catalytic effector protein (T4moD, 11.6 kDa) is the subject of this work. The natural enzyme complex produces >95% yield of *p*-cresol from NADH, O_2 and toluene (Mitchell *et al.*, 2002).

2002). These results underscore the fundamental importance of effector protein–hydroxylase interactions in diiron-enzyme catalysis.

Presently, NMR structures of effector proteins are available for four of the 24 members of the family: T4MO (T4moD; Hemmi *et al.*, 2001), phenol hydroxylase (DmpM; Qian *et al.*, 1997) and methane monooxygenase (MmoB) from *Methylosinus trichosporium* OB3b (Chang *et al.*, 1999) and *Methylococcus capsulatus* Bath (Walters *et al.*, 1999). These proteins represent three of the four functional classes identified for this enzyme family (Hemmi *et al.*, 2001). As a group, these proteins exhibit similar secondary-structure topology but significant variability in tertiary structure that may be related to functional divergence and/or an incomplete set of NMR distance restraints. In this report, we describe the crystallization and preliminary analysis of crystals of native T4moD, a selenomethionine-enriched form (SeMet-T4moD) and an isoform lacking ten residues from the N-terminus (ΔN10-T4moD). The data are of sufficient quality to permit crystal structural determinations for both T4moD and ΔN10-T4moD, which are in progress.

2. Material and methods

2.1. Protein expression and purification

T4moD was expressed in *Escherichia coli* BL21 (DE3) and purified as previously reported (Studts & Fox, 1999). SeMet-T4moD was expressed in the Met auxotroph *E. coli* B834 (DE3) in a minimal medium augmented with selenomethionine. ESI–MS analysis indicated that the N-terminal selenomethionine was removed during expression and that ∼95% incorporation was obtained at the two remaining methionine residues. The ΔN10-T4moD isoform was constructed by PCR using pJDP01 as the template (Xia *et al.*, 1999). The following oligonucleotides primers were used: DtruncF (5′-gctttaCATATGaataacgttggaccgattatccg-3′) and NtermR (5′-caaggggttatgctagttattgctcagcggt-3′). Capitals indicate the *Nde*I site used for cloning, which also placed the start codon two residues before β-strand 1, the first identifiable secondary-structure element in the NMR structure (Hemmi *et al.*, 2001).

2.2. Crystallization and X-ray data collection

The initial crystallization studies were performed using screening kits from Hampton Research and were optimized by screening additives and cryoconditions. All crystals were grown using the hanging-drop vapor-diffusion technique with standard 24-well Linbro plates from ∼2 µl of protein (typically in 25 m*M* MOPS pH 6.75, 7% glycerol, 0.15 *M* NaCl) mixed with 2 µl reservoir solution. T4moD crystallized at 298 K with 975 m*M* sodium/potassium phosphate pH 4.7, 400 m*M* NaCl, 50 m*M* succinate pH 5.5 at a protein concentration of approximately 40 mg ml^{-1}. The approximately 0.2 × 0.2 × 0.2 mm crystals grew reproducibly within approximately one week under aerobic conditions. Crystals of SeMet-T4moD (approximately 0.15 × 0.15 × 0.15 mm) were grown in an anaerobic chamber (Coy Labs, maintained with a 95% N_2 and 5% H_2 atmosphere) under nearly identical crystallization conditions supplemented with 1 m*M* dithiothreitol. Single crystals of T4moD and SeMet-T4moD were transferred to the mother liquor augmented with 30% glycerol as the cryoprotectant immediately prior to flash-freezing in a 100 K cold stream. For ΔN10-T4moD, the reservoir solution contained 2.0 *M* ammonium sulfate and 5%(*v*/*v*) 2-propanol and the drop (10 mg ml^{-1} protein) was augmented with 1 µl of 7.5%(*v*/*v*) 1,2,3-heptanetriol. Crystals appeared after 3 d of equilibration at 277 K and reached maximum dimensions of 0.2 × 0.2 × 0.2 mm within one week (Fig. 2). The ΔN10-T4moD crystals were transferred to Paratone-N (Hampton Research), the excess mother liquor was pulled away and the crystals were flash-frozen by rapid submersion in liquid N_2.

All X-ray diffraction data were collected from crystals held at approximately 100 K. The T4moD data set was collected with 0.7433 Å X-rays at the Advanced Photon Source (APS), Argonne National Laboratory. The four-wavelength MAD data set from the SeMet-T4moD crystal was collected at beamline 1-5 at the Stanford Synchrotron Radiation Source. An ADSC Quantum 4 detector was used to collect two 16° wedges of unique data and their inverse-beam wedges at the wavelengths indicated in Table 1. Each image was exposed for 45 s with a 1° oscillation. Diffraction data from ΔN10-T4moD crystals were collected during the commissioning of beamline 22-ID by the South East Regional Collaborative Access Team (SER-CAT) at APS using a MAR CCD 165 detector. Each image was collected with a 1 s exposure and a 1° oscillation range. The T4moD data set was processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). The SeMet-T4moD and ΔN10-T4moD data sets were processed with *MOSFLM* (Leslie, 1991; Powell, 1999) and *SCALA* from the *CCP*4 suite of programs (Dodson *et al.*, 1997; Collaborative Computational Project, Number 4, 1994).

The four-wavelength MAD data set of SeMet-T4moD was further scaled and analyzed with *FHSCALE* and *SCALEIT* from the *CCP*4 suite of programs (Colla-

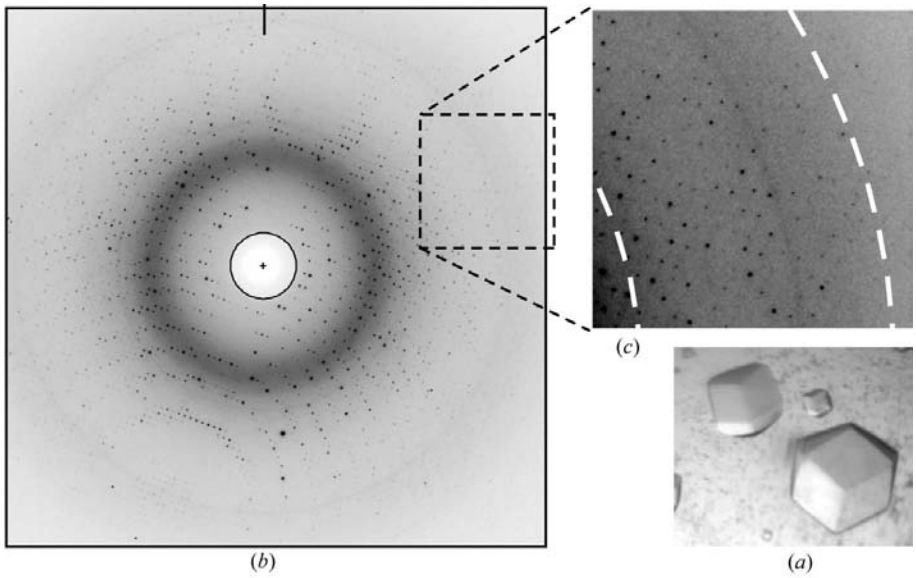

Figure 2
(*a*) Crystals of ΔN10-T4moD measuring ∼0.3 mm across the diagonal. (*b*) The X-ray diffraction pattern obtained with 1 s exposure and 1° oscillation along the vertical axis (vertical line). The data were collected at SER-CAT beamline 22-ID with a MAR CCD 165 area detector and a 103 mm crystal-to-detector distance. (*c*) An expanded and contrast-adjusted view of the diffraction pattern between 2.0 and 1.5 Å resolution (inner and outer dashed arcs), respectively.

Table 1
X-ray data-collection statistics for T4moD.

Unit-cell parameters for $P6_122$ crystals are $a = b = 86.41 \pm 0.15$, $c = 143.9 \pm 0.27$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$ and for $P2_13$ crystals are $a = b = c = 86.7 \pm 0.47$, $\alpha = \beta = \gamma = 90^\circ$. Values in parentheses are for the highest resolution shell.

	SeMet-T4moD†							
	Inflection (f')	Peak (f'')	Remote 1	Remote 2	T4moD‡	ΔN10-T4moD§	ΔN10-T4moD¶	ΔN10-T4moD¶
Wavelength (Å)	0.979880	0.979571	0.925256	1.068830	0.7433	1.5418	1.000	1.000
Resolution range (Å)	33–3.0 (3.08–3.0)	33–3.0 (3.08–3.0)	33–3.0 (3.08–3.0)	37–3.0 (3.08–3.0)	37–2.05 (2.11–2.05)	23–2.2 (2.32–2.20)	16–1.96 (2.07–1.96)	17–2.00 (2.11–2.00)
Space group	$P6_122$	$P6_122$	$P6_122$	$P6_122$	$P6_122$	$P2_13$	$P2_13$	$P2_13$
Total reflections	44128	44088	44703	39510	34207	235752	339875	333434
Unique reflections	6781	6791	6800	6293	19009	11161	15889	15236
Multiplicity	6.5 (5.9)	6.5 (6.0)	6.6 (6.8)	6.3 (5.3)	1.8 (1.2)	21.1 (21.1)	21.4 (21.0)	21.9 (21.4)
Completeness (%)	98.9 (98.9)	98.9 (91.3)	99.2 (99.2)	99.0 (91.0)	96.5 (78.2)	99.9 (99.9)	99.8 (99.8)	99.8 (99.8)
Anomalous completeness (%)	98.1	98.0	98.6	88.3	—	—	—	—
R_{sym}†† (%)	7.1 (27.5)	7.1 (29)	7.2 (32.7)	7.5 (33)	6.6 (29.4)	10.9 (30.0)	6.4 (31.7)	5.8 (34.3)
$I/\sigma(I)$‡‡	4.8 (2.7)	4.8 (2.6)	4.7 (2.3)	5.0 (2.5)	6 (2.5)	6.2 (2.4)	9.6 (2.3)	10.1 (2.2)

† Data collected at beamline 1-5 of SSRL with an ADSC Quantum 4 detector. ‡ Data collected at APS. § Data collected with rotating Cu anode and an MSC R-AXIS IV++ detector. ¶ Data collected at beamline 22-ID at APS with a MAR CCD 165 detector. †† $R_{sym}(I)$ gives the average agreement between the independently measured intensities such as $\sum_h \sum_i |I_i - I| / \sum_h \sum_i I$, where I is the mean intensity of the i observations of reflection h. ‡‡ $I/\sigma(I)$ is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

borative Computational Project, Number 4, 1994). The Se atoms were located with *SOLVE* (Terwilliger & Berendzen, 1999) using data from all four wavelengths and X-ray cross-sectional estimates of anomalous scattering factors at each wavelength as determined with *CROSSEC* (Collaborative Computational Project, Number 4, 1994). The Se sites were refined with *SHARP* (La Fortelle & Bricogne, 1997) to produce the experimental phases currently in use for model building and refinement. The resulting phasing statistics from *SHARP* are presented in Table 2.

Table 2
Phasing statistics for SeMet-T4moD.

Compiled from *SHARP* (La Fortelle & Bricogne, 1997).

Data set	Inflection (f')	Peak (f'')	Remote 1	Remote 2
Wavelength (Å)	0.979880	0.979571	0.925256	1.068830
Resolution range (Å)	33–3.0	33–3.0	33–3.0	37–3.0
R_{cullis} centric (iso)†	0.35	0.34	—	0.78
R_{cullis} acentric (iso/ano)	0.39/0.93	0.39/0.78	—/0.83	0.74/0.97
R_{kraut} centric (iso/ano)‡	0.06/0.12	0.07/0.13	0.08/0.13	0.09/0.14
R_{kraut} acentric (iso/ano)	0.02/0.04	0.02/0.04	0.03/0.04	0.03/0.04
Phasing power, centric (iso)§	4.11	4.00	0	0.12
Phasing power, acentric (iso/ano)	4.43/1.04	4.11/1.86	0/1.65	0.12/0.86
Figure of merit¶	0.521 (1593 centric reflections); 0.519 (5181 acentric reflections) 0.896 (after solvent flattening, 6815 reflections)			

† R_{cullis} = ⟨phase-integrated lack of closure⟩/⟨$|F_{ph} - F_p|$⟩, where F_{ph} is the structure factor obtained from the inflection, peak or remote 2 data sets and F_p is obtained from the 0.925256 Å data set. The isomorphous and anomalous differences are designated as iso and ano, respectively. ‡ $R_{kraut} = \sum \varepsilon_{iso} / \sum \Delta_{iso}$ for isomorphous differences and $\Sigma\varepsilon_{ano}/\sum\Delta_{Bijvoet}$ for anomalous differences, where ε_{iso} and ε_{ano} are the isomorphous and anomalous phase-integrated lack of closure, respectively, Δ_{iso} is the isomorphous difference and $\Delta_{Bijvoet}$ is the Bijvoet difference. § Phasing power = ⟨[$|F_{h(calc)}|$/phase-integrated lack of closure]⟩, where $F_{h(calc)}$ is

3. Results and discussion

The *tmoD* gene (NCBI accession No. M65106; Yen *et al.*, 1991) encodes a 103-residue 11 618 Da protein that contains no Cys or Trp residues. There are three Met residues (1, 74 and 103) in T4moD, but the N-terminal Met is efficiently removed during expression in *E. coli* BL21 (DE3). The calculated instability index of 49.7 suggests that T4moD may be 'unstable' (Guruprasad *et al.*, 1990). However, purified T4moD is stable and a high-resolution NMR structure has recently been completed (Hemmi *et al.*, 2001). Moreover, previous structural and catalytic studies suggest a potential functional role for the amino-terminal region of the catalytic effector protein families (Brandstetter *et al.*, 1999; Chang *et al.*, 2001). To initiate comparative functional studies of the T4MO complex, a truncated isoform of the T4moD was constructed, purified and crystallized.

The diffraction data from native T4moD and SeMet-T4moD crystals are consistent with space groups $P6_122$ or $P6_522$ ($a = b = 86.41 \pm 0.15$, $c = 143.9 \pm 0.27$ Å). Matthews coefficient and solvent-content analysis (Matthews, 1968) suggest either two ($V_M = 3.5$ Å^3 Da^{-1}, 64.8% solvent) or three ($V_M = 2.3$ Å^3 Da^{-1}, 47.2% solvent) molecules per asymmetric unit. Although *SOLVE* did not produce a reasonable result from analysis of the $P6_522$ space-group data, three Se sites were identified in the asymmetric unit for the $P6_122$ space group. The overall estimated figure of merit reported from *SOLVE* was 0.34, with a Z score of 15.22 and peak intensities at the Se sites of 39.5, 39.1 and 26.8σ. The presence of three well defined Se atoms is consistent with V_M calculations assuming a trimer in the asymmetric unit given that, for example, Met74 is well ordered and that Met103 is disordered in the crystal lattice. However, an alternative postulate of a dimer in the asymmetric unit is tenable by assuming that the two monomers are structurally inequivalent and that one of the C-terminal SeMet residues is disordered. The Se sites were refined with *SHARP* (Table 2) and the resulting experimental phases and electron-density maps clearly show that the asymmetric unit contains two T4moD molecules. The overall folds of the monomers are similar to the NMR structure (Hemmi *et al.*, 2001). However, there are at least two regions where the structure of each monomer appears to differ significantly from the other and from the NMR structure, which includes the amino-terminal regions.

Crystals of the ΔN10-T4moD isoform are colorless non-birefringent cubes (Fig. 2) and the diffraction data is consistent with the cubic space group $P2_13$ (unit-cell parameters $a = b = c = 86.75$ Å). Despite the different space group, the unit-cell edges of the cubic form are nearly identical to the two short cell edges of the hexagonal crystal form. A typical data set from ΔN10-T4moD consists of 339 875 observations of 15 889 unique reflections in the resolution range 15.83–1.96 Å. The overall completeness of 99.8% (99.8% in the 2.07–1.96 Å highest resolution shell) and overall R_{sym} of 6.4% (31.7% in the highest resolution shell) are both excellent.

The data are strong as indicated by the overall $I/\sigma(I)$ of 9.6. Assuming two ΔN10-T4moD monomers in the asymmetric unit, the calculated V_M of the crystals is 2.64 Å^3 Da^{-1}, corresponding to a solvent content of 53%. Molecular replacement is currently under way using the partially refined structure of the appropriately truncated native T4moD model and confirms the Matthews coefficient estimates for this isoform.

This research was supported by the Howard Hughes Medical Institutes (fellowship support to AMO from PI Dr B. W. Matthews, University of Oregon), funds from the Georgia Tech Research Corporation and the Georgia Institute of Technology Office of the Vice Provost for Research to AMO by National Science Foundation MCB-9733734 to BGF. Portions of this research were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL), operated by Stanford University on behalf of the US Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research and by the National Institutes of Health, National Center for Research Resources, Biomedical Technology Program and the National Institute of General Medical Sciences. Portions of the work were also carried out at beamline 22-ID in the SER-CAT facilities at the APS. The use of APS is supported by the US Department of Energy, Basic Energy Sciences, Office of Science under Contract No. W-31-109-Eng-38.

References

Andersson, K. K., Froland, W. A., Lee, S.-K. & Lipscomb, J. D. (1991). *New J. Chem.* **15**, 411–415.

Brandstetter, H., Whittington, D. A., Lippard, S. J. & Frederick, C. A. (1999). *Chem. Biol.* **6**, 441–449.

Chang, S.-L., Wallar, B. J., Lipscomb, J. D. & Mayo, K. H. (1999). *Biochemistry*, **38**, 5799–5812.

Chang, S.-L., Wallar, B. J., Lipscomb, J. D. & Mayo, K. H. (2001). *Biochemistry*, **40**, 9539–9551.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dodson, E. J., Winn, M. & Ralph, A. (1997). *Methods Enzymol.* **277**, 620–633.

Elango, N., Radhakrishnan, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscomb, J. D. & Ohlendorf, D. H. (1997). *Protein Sci.* **6**, 556–568.

Fox, B. G., Liu, Y., Dege, J. E. & Lipscomb, J. D. (1991). *J. Biol. Chem.* **266**, 540–550.

Froland, W. A., Andersson, K. K., Lee, S. K., Liu, Y. & Lipscomb, J. D. (1992). *J. Biol. Chem.* **267**, 17588–17597.

Guruprasad, K., Reddy, B. V. B. & Pandit, M. W. (1990). *Protein Eng.* **4**, 155–161.

Hemmi, H., Studts, J. M., Chae, Y. K., Song, J., Markley, J. L. & Fox, B. G. (2001). *Biochemistry*, **40**, 3512–3524.

La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.

Leslie, A. G. W. (1991). *Crystallographic Computing 5*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 51–61. Oxford University Press.

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

Merkx, M., Kopp, D. A., Sazinsky, M. H., Blazyk, J. L., Müller, J. & Lippard, S. J. (2001). *Angew. Chem. Int. Ed.* **40**, 2782–2807.

Mitchell, K. H., Studts, J. M. & Fox, B. G. (2002). *Biochemistry*, **41**, 3176–3188.

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

Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Münck, E., Steffan, R. J., McClay, K. & Fox, B. G. (1996). *Biochemistry*, **35**, 9106–9119.

Powell, H. R. (1999). *Acta Cryst.* D**55**, 1690–1695.

Qian, H., Edlund, U., Powlowski, J., Shingler, V. & Sethson, I. (1997). *Biochemistry*, **36**, 495–504.

Rosenzweig, A. C., Frederick, C. A., Lippard, S. J. & Nordlund, P. (1993). *Nature (London)*, **366**, 537–543.

Studts, J. M. & Fox, B. G. (1999). *Protein Expr. Purif.* **16**, 109–119.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Walters, K., Gassner, G. T., Lippard, S. J. & Wagner, G. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 7877–7882.

Whited, G. M. & Gibson, D. T. (1991). *J. Bacteriol.* **173**, 3010–3016.

Xia, B., Pikus, J. D., Xia, W. D., McClay, K., Steffan, R. J., Chae, Y. K., Westler, W. M., Markley, J. L. & Fox, B. G. (1999). *Biochemistry*, **38**, 727–739.

Yen, K. M., Karl, M. R., Blatt, L. M., Simon, M. J., Winter, R. B., Fausset, P. R., Lu, H. S., Harcourt, A. A. & Chen, K. K. (1991). *J. Bacteriol.* **173**, 5315–5327.