The structure of the periplasmic thiol–disulfide oxidoreductase SoxS from *Paracoccus pantotrophus* indicates a triple Trx/Grx/DsbC functionality in chemotrophic sulfur oxidation

Yvonne Carius, Dagmar Rother, Cornelius G. Friedrich and Axel J. Scheidig

The structure of the periplasmic thiol–disulfide oxidoreductase SoxS from *Paracoccus pantotrophus* indicates a triple Trx/Grx/DsbC functionality in chemotrophic sulfur oxidation

**Yvonne Carius,[a] Dagmar
Rother,[b] Cornelius G. Friedrich[b]*
and Axel J. Scheidig[a]***

[a]Abteilung für Strukturbiologie, Zoologisches
Institut, Christian-Albrechts-Universität zu Kiel,
Am Botanischen Garten 1-9, 24118 Kiel,
Germany, and [b]Lehrstuhl für Technische
Mikrobiologie, Fachbereich Bio- und
Chemieingenieurwesen, Universität Dortmund,
Emil-Figge-Strasse 66, 44221 Dortmund,
Germany

Correspondence e-mail:
cornelius.friedrich@udo.edu,
axel.scheidig@strubio.uni-kiel.de

The periplasmic thiol–disulfide oxidoreductase SoxS is beneficial for the sulfur-oxidizing (Sox) phenotype of the facultative chemotrophic bacterium *Paracoccus pantotrophus* and is not part of the Sox enzyme system. SoxS combines features of thioredoxins, glutaredoxins and the thiol–disulfide oxidoreductases of the Dsb family in structure, target specificity and reaction. The structure of SoxS was solved in oxidized and reduced forms at 2.1 and 1.9 Å resolution, respectively. SoxS revealed high structural homology to typical cytoplasmic bacterial thioredoxins. In contrast, SoxS contained the active-site motif Pro-Gly-Cys-Leu-Tyr-Cys that is not present in other thioredoxins. Interestingly, the sequence of this motif is closely related to the Pro-Gly-Cys-Pro-Tyr-Cys sequence of some glutaredoxins and to the Pro-Xaa-Cys-Xaa-Tyr-Cys sequences of some members of the DsbC and DsbG subfamilies of thiol–disulfide oxidoreductases. Furthermore, the proposed substrate of SoxS, the interprotein disulfide of SoxY, Cys110^Y–Cys110^Y, is structurally similar to oxidized glutathione. However, SoxS is proposed to specifically reduce the interprotein disulfide between two SoxY subunits, releasing a heterodimeric SoxYZ as an active part of the sulfur-oxidation cycle.

1. Introduction

The facultative aerobic alphaproteobacterium *Paracoccus pantotrophus* grows chemoautotrophically using thiosulfate as an energy source and heterotrophically using a large variety of carbon sources (Rainey *et al.*, 1999). The sulfur-oxidizing (Sox) enzyme system of *P. pantotrophus* comprises the periplasmic proteins SoxYZ, SoxB, SoxCD and SoxXA. *In vitro*, these proteins together catalyze hydrogen sulfide-dependent, sulfur-dependent, thiosulfate-dependent and sulfite-dependent cytochrome *c* reduction (Rother *et al.*, 2001). The heterodimeric SoxYZ complex is the central protein of this system. The inorganic sulfur substrate is covalently linked to the thiol of the Cys110 residue of the SoxY subunit (referred to as Cys110^Y) in a two-electron-yielding step by the haem enzyme SoxXA, forming a persulfide thiosulfonate. The sulfone is subsequently hydrolyzed off by the sulfate thiohydrolase SoxB. The resulting sulfane sulfur of the SoxY-Cys110-persulfide is then oxidized to SoxY-Cys110-thiosulfonate in a six-electron-yielding step by the heterotetrameric sulfane dehydrogenase Sox(CD)$_2$. The new sulfone is finally released as sulfate by SoxB, which in turn regenerates SoxYZ for a new reaction cycle (Friedrich *et al.*, 2001).

SoxYZ has three partner proteins to which it delivers the active-site Cys110^Y. The active site of SoxYZ is freely accessible and is positioned on the peptide swinging arm of the carboxy-terminus of SoxY (Quentmeier & Friedrich, 2001; Sauve *et al.*, 2007). Thus, SoxYZ is susceptible to nonspecific chemical modifications of the thiol of Cys110^Y such as the formation of interprotein disulfides or mixed disulfides (Quentmeier & Friedrich, 2001). In fact, covalent modifications of Cys110^Y inactivate SoxYZ *in vivo* and require a reaction for reactivation (Quentmeier *et al.*, 2007).

Thioredoxin-like proteins are able to reduce protein disulfides. The *sox* gene cluster contains two genes, *soxS* and *soxW*, which encode periplasmic thioredoxin-like proteins and yield mature SoxS (99 amino acids, 11 077 Da) and SoxW (166 amino acids, 18 719 Da), respectively. Both proteins are kept in the reduced state by the cysteine–disulfide transporter SoxV, a paralogue of the apocytochrome *c*-reducing CcdA superfamily. SoxV is not involved in cytochrome *c* maturation, but is specific for chemotrophic growth with thiosulfate (Bardischewsky & Friedrich, 2001; Bardischewsky *et al.*, 2006). SoxS is beneficial for growth with thiosulfate, while SoxW is not (Orawski *et al.*, 2007). SoxS is distinct from known thioredoxins of the genus *Paracoccus*, but is highly homologous to proteins that are exclusively present in sulfur-oxidizing chemotrophic and phototrophic alphaproteobacteria (Orawski *et al.*, 2007).

Thioredoxins are small redox-active proteins with a Cys-Xaa-Xaa-Cys active-site sequence motif. The members of the thioredoxin superfamily are distinguished by their redox partners and functions in the dithiol/disulfide structure of proteins. They have been classified into five subfamilies, as reviewed by Pan & Bardwell (2006) and Kadokura *et al.* (2003). In bacteria they are involved in important cellular functions, one of which is protein repair (reviewed in Zeller & Klug, 2006). Cytoplasmic thioredoxins that perform reductive steps are re-reduced by either NADPH or glutathione. Periplasmic thioredoxins form either protein disulfides, *e.g.* DsbA and DsbB (Bader *et al.*, 1999), act as protein disulfide isomerases, *e.g.* DsbC and DsbG (reviewed by Kadokura *et al.*, 2003; Bardischewsky *et al.*, 2006), or reduce protein disulfides, as in cytochrome *c* maturation. The redox partners of periplasmic thioredoxins are membrane-integrated thiol–disulfide transporters (Kranz *et al.*, 1998; Erlendsson *et al.*, 2003). Thioredoxin-like proteins are usually nonspecific with respect to their target proteins. To date, only the thioredoxin-like protein ResA has been described as being specific with respect to its target cytochrome c_{550} of *Bacillus subtilis* and specific with respect to the reduction of the cysteine disulfide of the C*XX*CH motif (Colbert *et al.*, 2006).

Cys13 of SoxS from *P. pantotrophus* specifically binds to Cys110^Y, as shown by site-directed mutagenesis of the ^{13}Cys-Leu-Tyr-^{16}Cys motif of SoxS (Rother *et al.*, 2008). The thiol of Cys110^Y is subject to chemical modifications by interprotein or mixed disulfide bonds (Quentmeier & Friedrich, 2001; Quentmeier *et al.*, 2008). SoxS is proposed to specifically reduce protein disulfide and mixed disulfide bonds of Cys110^Y which inactivate SoxY. Here, we report the structural features of

SoxS in the oxidized and reduced states. This is the first crystal structure of a periplasmic thiol–disulfide oxidoreductase with mixed thioredoxin/glutaredoxin/DsbC-type structural characteristics that suggest different types of reactions.

2. Experimental procedures

2.1. Expression and purification of SoxS protein constructs

For the production of recombinant His$_6$-SoxS, the corresponding DNA encoding the mature periplasmic SoxS protein was amplified using the forward primer 5′-AAA**AGGATCC**-GATGACGATGACAAAGCAGAGCTGCGGCTCTTGAT-3′ and the reverse primer 5′-TTTT**AAGCTT**TCATTGCCCC-GGCTCGGCCTGA-3′ (bold sequences indicate *Bam*HI and *Hin*dIII restriction sites, respectively). The resulting fragment was cloned into the expression vector pQE-30 (Qiagen) in order to generate a fusion protein containing an N-terminal His$_6$ affinity-tag sequence followed by an enterokinase cleavage site (see Fig. 2). The plasmid was transformed into *Escherichia coli* strain M15[pREP4] (Qiagen). A 6 l culture was cultivated aerobically at 310 K in Lysogenic Broth medium supplemented with the antibiotics ampicillin (100 µg ml^{-1}) and kanamycin (30 µg ml^{-1}) to an OD$_{600}$ of 0.7. His$_6$-SoxS (referred to hereafter as SoxS) was expressed after addition of isopropyl β-D-1-thiogalactopyranoside (IPTG; final concentration 0.4 m*M*) for 4 h at 303 K. The cells were harvested by centrifugation (10 min, 6000*g*, 277 K), washed in lysis buffer [50 m*M* phosphate buffer pH 8.0, 300 m*M* NaCl, 10 m*M* imidazole, 5 m*M* 2-mercaptoethanol, 0.1 m*M* phenylmethylsulfonyl fluoride (PMSF)], frozen in liquid nitrogen and stored at 193 K.

For the purification of SoxS, cells were thawed, resuspended in lysis buffer and disrupted at 110 MPa using a cell homogenizer (Avestin Inc.). After centrifugation (120 min, 75 000*g*, 277 K), the supernatant was applied onto an Ni–NTA–Superflow matrix (Qiagen) and the column was washed to baseline with 50 m*M* phosphate buffer pH 8.0 containing 300 m*M* NaCl, 5 m*M* 2-mercaptoethanol, 0.1 m*M* PMSF and 20 m*M* imidazole. SoxS was eluted in the same buffer with a linear gradient from 20 to 250 m*M* imidazole. Fractions containing SoxS (identified by SDS–PAGE with Coomassie staining) were pooled, dialyzed against 50 m*M* Tris–HCl pH 8.0 with 5 m*M* 2-mercaptoethanol and applied onto a Superdex 200 gel-filtration column (16/600, GE Healthcare) equilibrated with 50 m*M* Tris–HCl pH 7.5 and 5 m*M* 2-mercaptoethanol (flow rate 0.8 ml min^{-1}). Fractions containing pure SoxS were pooled and concentrated to 20–25 mg ml^{-1}. The yield was 16 mg per litre of culture medium. The His$_6$ tag was not removed prior to crystallization.

Selenomethionine-substituted SoxS (referred to hereafter as SoxS-SeMet) was produced in *E. coli* M15[pREP4] using a modified two-step protocol as described by Guerrero *et al.* (2001). The cells were cultivated in LB medium supplemented with ampicillin (100 µg ml^{-1}) and kanamycin (30 µg ml^{-1}) for 16 h, harvested and washed with sterile deionized water. The cells were then inoculated into 2 l SelenoMet medium

(Molecular Dimensions) supplemented with the same antibiotics and L-selenomethionine (40 μg ml^{-1}) to yield an initial OD_{600} of 1.0. After incubation for 30 min at 303 K, SoxS-SeMet was overexpressed by the addition of IPTG (0.4 mM) for 5 h at 303 K. SoxS-SeMet was purified essentially as described for the native construct. The yield was around 18 mg per litre of culture medium. The homogeneity of both proteins was further verified by MALDI–TOF.

2.2. Crystallization

Crystals of recombinant SoxS and its selenomethionine derivative were grown at 291 K using the vapour-diffusion method in hanging drops. Crystallization conditions were screened using Crystal Screen and Crystal Screen 2 (Hampton Research). 1 μl protein solution at a concentration of 20 mg ml^{-1} (in 50 mM Tris pH 7.5 with 5 mM 2-mercaptoethanol) was mixed with 1 μl reservoir solution and equilibrated against 1 ml reservoir solution. The final reservoir conditions were sodium citrate buffer pH 4.8–5.4, 0.4 M sodium dihydrogen phosphate and 0.4 M potassium dihydrogen phosphate. Spindle-shaped crystals belonging to space group $P3_121$ with unit-cell parameters $a = b = 80.9$, $c = 33.2$ Å grew within 3 d (Fig. 1). Based on analyses using *SFCHECK* from the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994), these crystals were found to display merohedral twinning. Therefore, an additive screen was performed based on the best crystallization condition identified. Crystals of the selenomethionine derivative grown with 3% D-(+)-glucose as an additive revealed significantly reduced twinning factors. For data collection under cryogenic conditions, single crystals were soaked in cryoprotectant solution [reservoir solution supplemented with 25%(v/v) glycerol] for 2 min prior to flash-cooling in liquid nitrogen.

2.3. Data collection, processing and structure refinement

The initial X-ray diffraction data sets for crystals of SoxS and its selenomethionine-substituted form were collected on the home source (Xcalibur PX Ultra System, Oxford Diffraction) at 100 K using Cu $K\alpha$ radiation ($\lambda = 1.541$ Å). These data sets were processed with *XDS/XSCALE* (Kabsch, 1988, 1993), *CrysAlis* (Oxford Diffraction) or *MOSFLM* (Leslie, 1992) and checked with *SFCHECK* from the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994). Assuming the presence of one molecule of SoxS in the asymmetric unit, the Matthews parameter (V_M; Matthews, 1968) was calculated to be 2.4 Å^3 Da^{-1}, with a solvent content of 49.0%. The first diffraction tests revealed a high degree of merohedral twinning (around 30–40% with twinning law $-h, -k, l$). To overcome potential problems during phasing and structure refinement arising from twinning, modified crystallization conditions using additives were analyzed that reduced the twinning ratio. For the collection of complete data sets, crystals with a low twinning ratio of around 4% were chosen. A selenium fluorescence scan and a four-wavelength data set were collected from a single SoxS-SeMet crystal on the HASYLAB beamline X12 at DESY, Hamburg (the indi-

vidual data sets were collected in the sequence peak, inflection point, high-energy remote, low-energy remote). Diffraction images were processed with *DENZO/SCALEPACK* (Otwinowski & Minor, 1997). The structure of selenomethionine-substituted SoxS was solved during data collection using the 3W-MAD protocol of *Auto-Rickshaw* (Panjikar *et al.*, 2005) with the first three data sets (peak, inflection point and high-energy remote). The input diffraction data were prepared and converted for use in *Auto-Rickshaw* using programs from the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994). The anomalous F_A values were calculated using the program *SHELXC* (Sheldrick *et al.*, 2001; Schneider & Sheldrick, 2002; Sheldrick, 2008). Based on initial analysis of the data, the maximum resolution for substructure determination and initial phase calculation was set to 2.4 Å. All five Se atoms requested were found using the program *SHELXD* (Schneider & Sheldrick, 2002). The correct hand for the substructure was determined using the programs *ABS* (Hao, 2004) and *SHELXE* (Sheldrick, 2002). The occupancies of all substructure atoms were refined and initial phases were calculated using the program *MLPHARE* (Collaborative Computational Project, Number 4, 1994). The initial phases were improved using density modification and phase extension to 2.4 Å resolution using the program *DM* (Cowtan & Main, 1998). 92 of 116 amino-acid residues of the model were built using the program *ARP/wARP* (Perrakis *et al.*, 1999; Morris *et al.*, 2003). The final structure of the selenomethionine derivative of SoxS was refined against the low-energy remote data set. This structure was used to solve the structure of the native protein by molecular replacement using the program *MOLREP* (Vagin & Teplyakov, 1997). Both structures were inspected and manually adjusted with *Coot* (Emsley & Cowtan, 2004) and refined with *REFMAC5* (Murshudov *et al.*, 1997). The final statistics for both structures are given in Table 1. The coordinates and structure-factor amplitudes have been deposited in the PDB under accession codes 3d4t (oxidized form) and 3dml (reduced form). Struc-
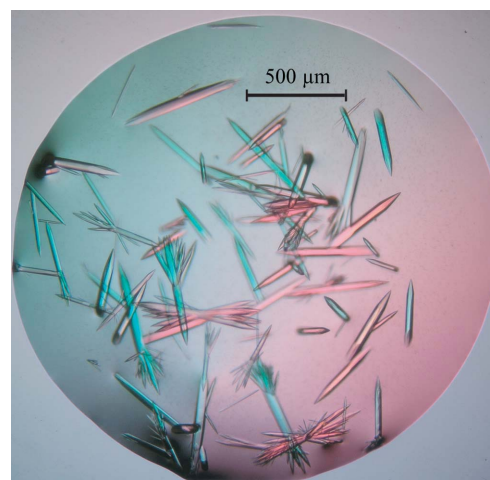
Figure 1
Crystals of SoxS grown in citrate buffer, sodium phosphate and potassium phosphate. They are spindle-shaped and twirled at the ends, which might be an indication of the twinning of the crystal lattice.

turally related proteins were identified in the PDB (Berman *et al.*, 2000) using the *DALI* server (Holm & Sander, 1996). Sequences of related thioredoxins were aligned with *ClustalW* (Thompson *et al.*, 1994).

3. Results

3.1. Overall structure and comparison with members of the thioredoxin protein family

SoxS was produced heterologously in *E. coli* without its N-terminal signal sequence but with an N-terminal His$_6$ tag

followed by an enterokinase cleavage site. The affinity tag was not removed for crystallization. The protein was expressed without and with substitution of methionine by seleno-methionine. The correct molecular masses of the recombinant proteins were verified by MALDI mass spectrometry. Based on SDS–PAGE analysis and MALDI mass spectrometry, no proteolytic cleavage was observed in dissolved protein crystals. The polypeptide chain is well defined in the electron-density maps for both structures except for the N-terminal His$_6$ tag, the enterokinase site and the last three C-terminal amino-acid residues. The alignment of the SoxS protein

Table 1
Diffraction data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

	SoxS-SeMet (reduced form)				SoxS (oxidized form)
	Peak	Inflection point	High-energy remote	Low-energy remote	Native
Data collection and processing					
X-ray source	X12/EMBL				Xcalibur PX Ultra
Detector	MAR Mosaic CCD detector (225 mm)				Onyx CCD (165 mm)
Wavelength (Å)	0.97833	0.97838	0.95369	1.0332	1.54
Temperature (K)	100	100	100	100	100
Crystal-to-detector distance (mm)	250	250	250	180	80
Exposure time (s)	30	30	30	50	120
Oscillation range (°)	1	1	1	0.5	1
Total oscillation range (°)	180	180	180	126.5	180
Space group	$P3_121$	$P3_121$	$P3_121$	$P3_121$	$P3_121$
Unit-cell parameters					
$\quad a = b$ (Å)	80.9	80.9	80.9	80.9	81.0
$\quad c$ (Å)	33.2	33.2	33.2	33.2	33.1
$\quad \alpha = \beta$ (°)	90	90	90	90	90
$\quad \gamma$ (°)	120	120	120	120	120
Resolution limit (Å)	2.4 (2.46–2.4)	2.40 (2.46–2.4)	2.4 (2.46–2.4)	1.9 (1.94–1.9)	2.05 (2.16–2.05)
Completeness (%)	99.9 (100)	99.9 (100)	99.9 (100)	99.7 (98.3)	99.5 (100.0)
No. observations (overall/unique)	51623/5050	53837/5050	54214/5050	64193/10007	268214/7981
Average redundancy	10.2 (9.4)	10.6 (9.8)	10.7 (9.9)	6.4 (5.2)	33.6 (19.7)
$\langle I/\sigma(I)\rangle$	19.4 (8.9)	19.4 (8.5)	19.1 (8.3)	30.9 (6.4)	36.0 (6.2)
$R_{\text{r.i.m.}}$† (%)	10.7 (28.9)	10.8 (30.8)	10.9 (33.0)	9.0 (49.6)	2.2 (11.5)
$R_{\text{p.i.m.}}$‡ (%)	3.3 (9.4)	3.3 (9.7)	3.3 (10.4)	3.4 (21.5)	12.6 (51.8)
B factor from Wilson plot (Å^2)	30.9	31.9	31.1	18.4	17.3
FOM (*DM*)	0.665 (0.871)	0.665 (0.871)	0.665 (0.871)		
Refinement statistics					
Resolution limit (Å)				1.9 (1.95–1.9)	2.05 (2.1–2.05)
No. of unique reflections				9530 (684)	7621 (547)
Completeness of data (%)				99.7 (97.3)	100 (100)
R_{cryst}§ (%)				15.4 (19.9)	17.4 (22.7)
R_{free}¶ (%)				20.1 (23.4)	21.6 (33.4)
No. of non-H atoms					
$\quad$Overall				961	904
$\quad$Protein				799	789
$\quad$Solvent				162	115
Ramachandran plot††, residues in (%)					
$\quad$Favoured regions				96.0	94.7
$\quad$Allowed regions				4.0	5.3
$\quad$Generously allowed regions				0	0
Coordinate error‡‡				0.124	0.155
R.m.s. deviations from ideal values					
$\quad$Bond lengths (Å)				0.012	0.013
$\quad$Bond angles (°)				1.512	1.608
Mean B factor (Å^2)					
$\quad$Backbone				16.7	14.0
$\quad$Side chain				19.2	15.3
$\quad$Solvent				43.6	29.6

† $R_{\text{r.i.m.}} = 100 \times \sum_{hkl}[N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl} \sum_i I_i(hkl)$. ‡ $R_{\text{p.i.m.}} = 100 \times \sum_{hkl}[1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith individual measurement of the reflection with Miller indices hkl and $\langle I_i(hkl)\rangle$ is the mean intensity of all measurements of $I(hkl)$, calculated for $I \geq 3\sigma(I)$; N is the redundancy or multiplicity of the observed reflection (Weiss, 2001; Diederichs & Karplus, 1997). § $R_{\text{cryst}} = 100 \times \sum ||F_{\text{obs}}| - |F_{\text{calc}}||/\sum |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively. ¶ R_{free} is equivalent to R_{cryst} but calculated with reflections (5%) omitted from the refinement process (Brünger, 1992; Tickle *et al.*, 2000). †† Calculated using the program *PROCHECK* (Laskowski *et al.*, 1993). ‡‡ Calculated based on a Luzzati plot using the program *SFCHECK* (Vaguine *et al.*, 1999).

sequence with all known members of the thioredoxin family (329 sequences) and the glutaredoxin family (173 sequences) as listed in the UniProtKB/Swiss-Prot database (April 2008; http://www.expasy.ch/) did not identify any close relation to other members of these families. However, the closest relatives outside the group of SoxS proteins from other organisms were all members of the DsbC and DsbG subfamilies which had some relation to the glutaredoxins regarding the active site (Fig. 2). SoxS is composed of four β-strands surrounded by two α-helices (Fig. 3). A search for structurally related proteins in the PDB (Berman et al., 2000) performed with the DALI server (Holm & Sander, 1996) identified E. coli thio-
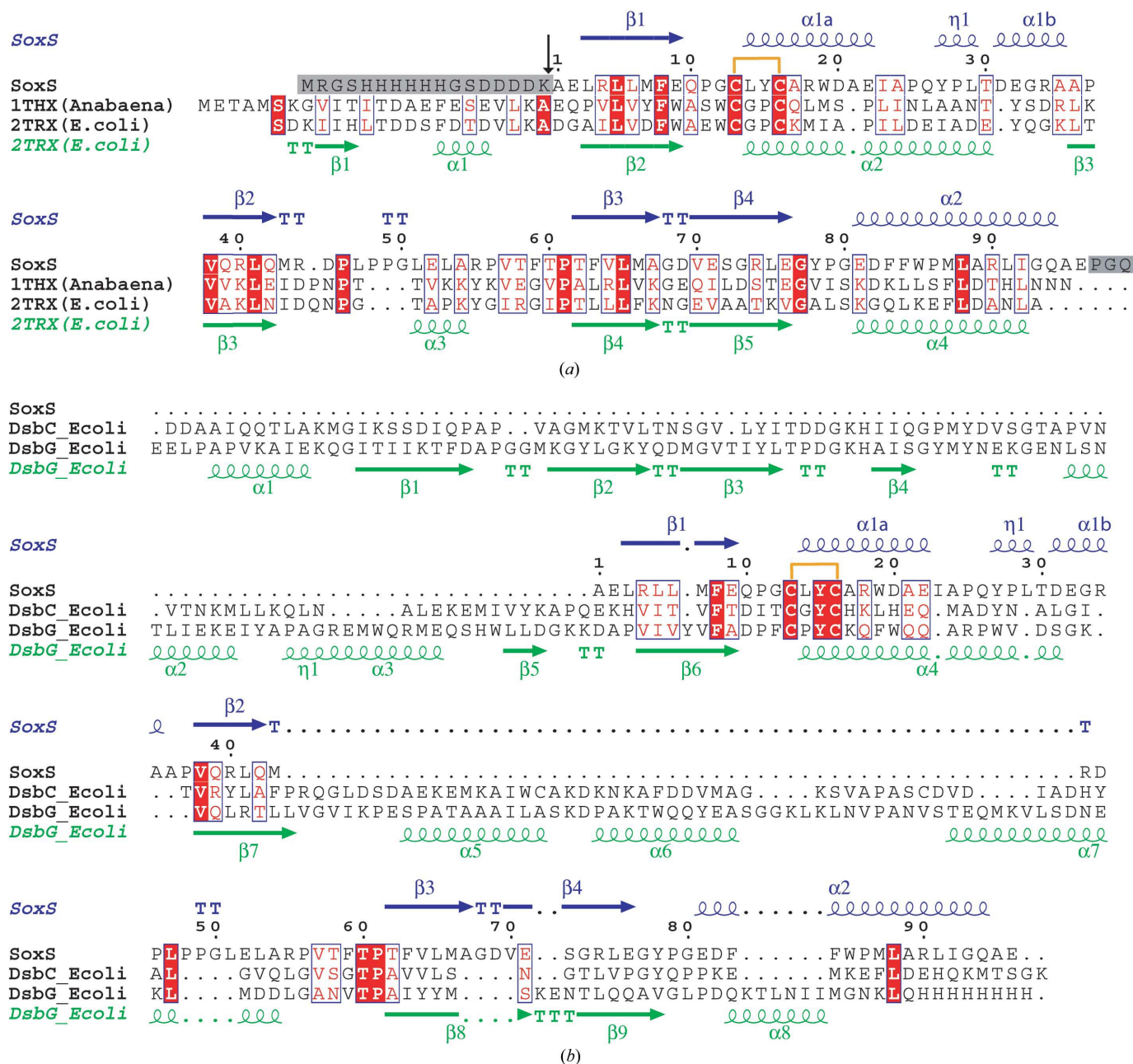
Figure 2
(a) Sequence alignment between the proteins SoxS (P. pantotrophus), thioredoxin 2 (Anabaena sp.; PDB code 1thx; Saarinen et al., 1995) and thioredoxin 1 (E. coli; PDB code 2trx; Katti et al., 1990). (b) Sequence alignment between the proteins SoxS (P. pantotrophus), DsbC (E. coli; PDB code 1tjd; Banaszak et al., 2004) and DsbG (E. coli; PDB code 1v57; Heras et al., 2004). The sequence alignment as initially performed with the program ClustalW (Thompson et al., 1994) had to be manually adjusted in order to fit secondary-structure elements based on analysis with the program DALILITE (Holm & Sander, 1997). The final figure was prepared with the program ESPript (Gouet et al., 1999) using the similarity score 0.7 0.5 B. The secondary-structure elements were determined with the program DSSP (Kabsch & Sander, 1983). The residues highlighted in grey at the N-terminus in the SoxS sequence in (a) indicate the His$_6$ affinity tag and the enterokinase recognition sequence within the expression construct, which were not visible in the electron density. The arrow ($\downarrow$) indicates the enterokinase cleavage site, which is identical to the cleavage site for the periplasmic leader peptide. The residues highlighted in grey in (a) at the C-terminus are also not visible in the electron density. The orange brackets highlight the two cysteinyl residues of the active site.

redoxin 1 (Katti *et al.*, 1990; r.m.s.d. of 1.9 Å for 84 of 108 C^α atoms, 14% sequence identity and a *Z* score of 10.2) and *Anabaena* thioredoxin 2 (Saarinen *et al.*, 1995; r.m.s.d. of 2.2 Å for 85 of 108 C^α atoms, 19% sequence identity and *Z* score of 9.2) as the closest structural homologues of SoxS. Thus, despite the low sequence identity, SoxS from *P. pantotrophus* super-imposes very well with these classical thioredoxins. The thioredoxin-like fold is characterized by a $\beta\alpha\beta$ motif linked by an α-helix to a $\beta\beta\alpha$ motif (Pan & Bardwell, 2006). Major deviations between SoxS and the structurally related thioredoxins were observed in the N-terminal region and the linker region between the two core motifs. SoxS does not have the first β-strand $\beta1$ and α-helix $\alpha1$ of these two classical thioredoxins. The native structure of SoxS as released in the peri-plasm after removal of the peri-plasmic leader sequence starts with a β-strand corresponding to $\beta2$ of Trx-1 (see Figs. 2 and 3). In addition, SoxS does not have an α-helix linking the two core motifs; instead it displays an extended loop between β-strands $\beta2$ and $\beta3$. This loop reaches into space partially occupied by helix $\alpha1$ of Trx-1 (see Fig. 3) and is very well structured owing to having a Pro-Leu-Pro-Pro motif of increased rigidity and making packing interactions with the N-terminal part of SoxS. Further-more, the long helix $\alpha2$ of Trx-1 is interrupted in SoxS by a π-helical turn (see Fig. 3), forming an extended surface patch along this helix.

3.2. Active site

Based on the presence of only one redox-active C*XX*C motif, SoxS belongs to the thioredoxin 1 family founded by *E. coli* thioredoxin 1. The active site comprises the residues Gly-[13]Cys-Leu-Tyr-[16]Cys-Ala (the numbering is based on the peri-plasmic form of the SoxS protein without the periplasmic leader peptide). This C*XX*C motif is positioned between the end of β-strand $\beta1$ and the beginning of α-helix $\alpha1$ (Fig. 2). Whereas a common theme in most thior-edoxins and glutaredoxins is the incorporation of a proline residue at the first position of the C*XX*C
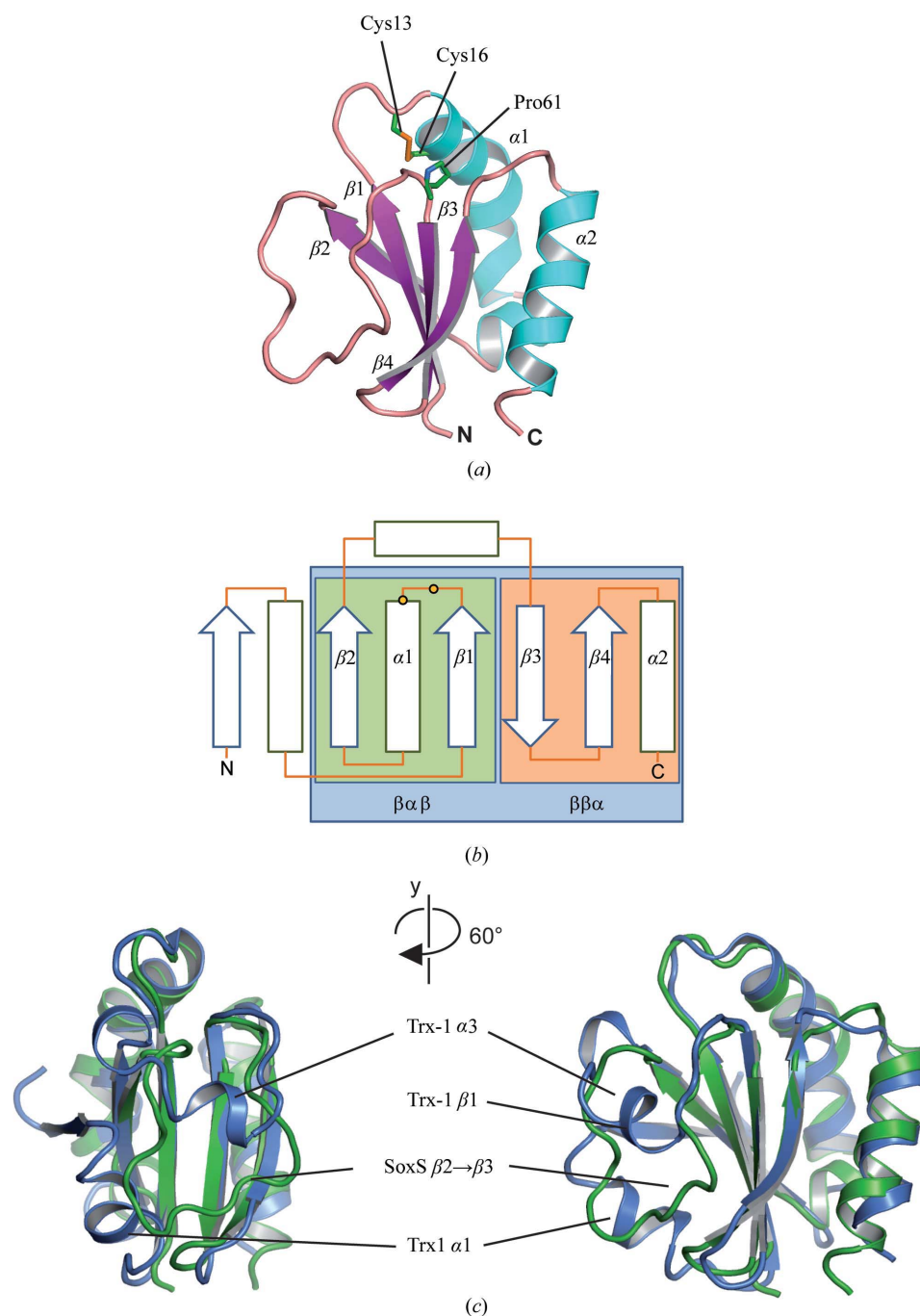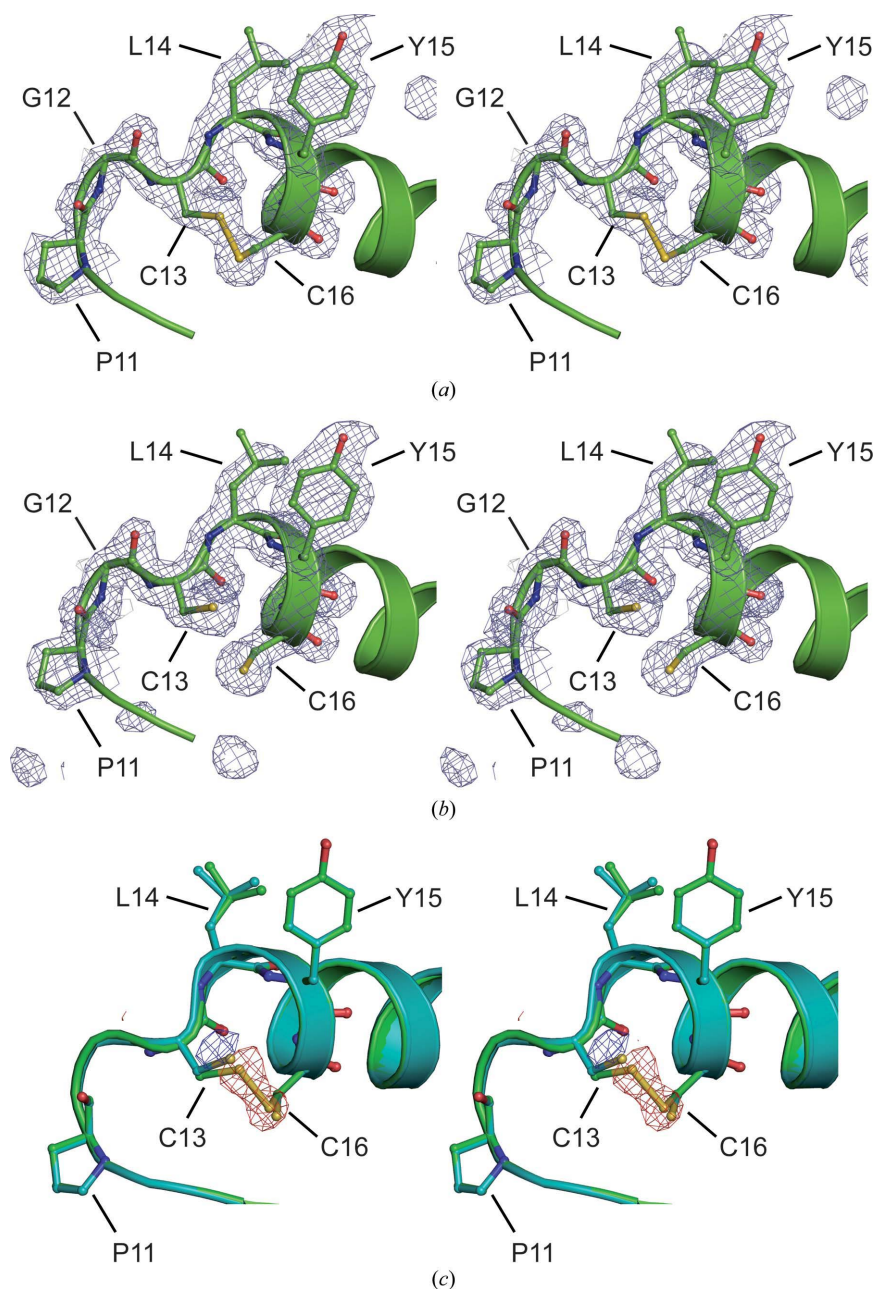

Figure 3
(*a*) Ribbon representation of the main-chain trace of oxidized SoxS. The disulfide bridge formed by the two cysteinyl residues of the active-site C*XX*C motif together with the proline residue Pro61 are displayed in stick representation. The secondary-structure elements are labelled. (*b*) Schematic representation of the general thioredoxin fold. The two motifs of the core fold are indicated with green and red backgrounds. The additional secondary-structural elements as observed in Trx-1 from *E. coli* are drawn outside the blue box of the core thioredoxin fold. The two active-site cysteinyl residues are indicated by yellow circles. (*c*) Superposition of SoxS (green) and Trx-1 (*E. coli*; blue). The two views are related by a 60° rotation as indicated. Trx-1 has two extra α-helices (Trx-1 $\alpha1$ and $\alpha3$) and one additional β-sheet (Trx-1 $\beta1$). The extended loop region in SoxS linking β-strands $\beta2$ and $\beta3$ substitutes for Trx-1 helix $\alpha3$ and partly takes up the space occupied by Trx-1 helix $\alpha1$.

motif and an aromatic residue in the second position, in SoxS the first position of the *CXXC* motif is replaced by a leucine residue. Furthermore, the frequently observed tryptophan residue preceding the first active-site cysteine is replaced by a glycine residue. We have determined the structure of SoxS in two different states: the oxidized state in which the two active-site cysteines form a disulfide bond and a partially reduced state that indicates free cysteinyl residues. 2-Mercaptoethanol was used as a reductant throughout purification and crystallization. However, in order to allow spontaneous oxidation to occur in the crystals of native SoxS, we did not replenish the reductant after setting up the crystals and we harvested the crystals after a couple of weeks. In contrast, the crystals of SeMet-modified SoxS were harvested within a couple of days after setting up with fresh 2-mercaptoethanol. We did not observe significant changes between these crystals in diffraction quality (mosaic spread) or unit-cell parameters. Structural comparison of refined native and SeMet-substituted SoxS revealed only a few differences. The most prominent was observed between Cys13 and Cys16 (Fig. 4). Based on refinement with OMIT maps and $F_{obs} - F_{obs}$ maps, we concluded that Cys13 and Cys16 are linked by a disulfide bond in the native SoxS crystals, whereas significant evidence of reduction was observed in the case of SeMet-substituted SoxS. Since it has been well documented that synchrotron radiation at high doses can reduce disulfide groups (Weik *et al.*, 2000, 2002), we analyzed and refined the individual data sets within the four-wavelength synchrotron data set. The distance between the cysteine S atoms of Cys13 and Cys16 was measured and compared with that observed in a SoxS-SeMet crystal structure for which the data set was collected in-house using a sealed-tube X-ray generator. For the synchrotron data set a stepwise increase in the distance could be observed: 2.6 Å (peak), 3.0 Å (inflection point), 3.1 Å (high-energy remote) and 3.3 Å (low-energy remote). The distance between these two S atoms in the oxidized form (collected in-house) is 2.1 Å, which is close to the theoretical value of 2.04 Å for a disulfide bridge. The distance between these two S atoms in a crystal derived from the SoxS-SeMet form kept under reducing conditions during purification, crystallization and crystal mounting was refined from an in-house data set to 3.0 Å, which is close to the theoretical value (3.1 Å) for two neighbouring thiol groups. Based on this analysis, we conclude that (i) the crystals of native SoxS are predominantly in the oxidized form, (ii) the crystals of selenomethionine-substituted SoxS are in a mixed state but with an excess of the reduced form and (iii) further reduction of the disulfide bridge took place during data collection at the synchrotron. For the rest of
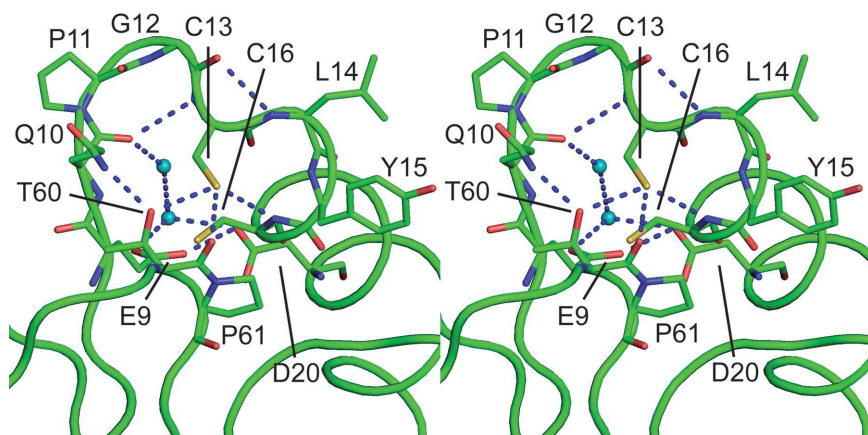

Figure 4
Stereo representation of the active site. (*a*) Difference σ_A-weighted $2mF_{obs} - DF_{calc}$ electron-density map for the oxidized form of SoxS (contoured at 1.5σ above the mean value of the map) and (*b*) for the reduced form of SoxS (contoured at 1.5σ above the mean value of the map). (*c*) Active sites of reduced SoxS (cyan) and oxidized SoxS (green), respectively, superimposed with a difference electron-density map calculated between the data sets from the reduced and oxidized form ($F_{obs, reduced} - F_{obs, oxidized}$, $\alpha_{calc, oxidized}$), contoured at 5.0σ (blue) and -5.0σ (red), respectively. The weighted R factor between these two data sets is 11.2% as calculated by the program *SCALEIT* from the *CCP*4 program package (Collaborative Computational Project, Number 4, 1994). This figure was prepared using *PyMOL* (DeLano, 2004).

the protein the two structures superimpose very well with each other (the r.m.s. distance in C^α positions is only 0.2 Å). No conformational changes in the main-chain trace or in side-chain rotamers could be observed, indicating a very rigid structure, especially within the active site.

Upon reduction of the disulfide bond, the S atoms of Cys13 and Cys16 move away from each other by a side-chain rotation along χ_1 of 16° and 10°, respectively. Only small adjustments in the main-chain trace of Pro11 and Gly12 are coupled to this disulfide-bond cleavage. This also reflects the tight and rigid conformation of the loop ^{10}Gln-Pro-Gly-Cys-Leu. In addition to the proline-specific restriction in its main-chain conformation, this loop is stabilized by hydrogen bonds that are formed between Gln10 O and Cys13 N (d = 3.23 Å), Gly12 O and Leu14 N (d = 3.14 Å) and Cys13 O and Cys16 N (d = 3.18 Å) (Fig. 5). An additional contribution to the rigidity of the ^{13}Cys-Leu-Tyr-Cys-containing loop comes from the side-chain groups of Leu14 and Tyr15, which are tightly packed against each other.

Upon reduction and associated site-chain rotation, the thiol group of Cys13 becomes more solvent-accessible (oxidized form, 18.7 Å^2; reduced form, 24.3 Å^2). A similar accessibility (23.1 Å^2) can be calculated for the thiol group in human glutaredoxin (PDB code 2fls; C. Johansson, C. Smee, K. L. Kavanagh, J. Debreczeni, U. Oppermann & M. Sundstrom, unpublished work) if the bound glutathione is removed from the model. In the reduced form, Cys13 is within hydrogen-bonding distance of Thr60 O (d = 3.39 Å), Thr60 OG1 (d = 3.26 Å), Cys16 N (d = 3.17 Å) and Cys16 SG (d = 3.36 Å). The interaction with Thr60 is conformationally stabilized by a hydrogen bond between Thr60 OG1 and Gln10 NE2 (d = 2.89 Å) and by the adjacent residue Pro61, which also shields Cys16 from the solvent (Fig. 5). In the oxidized state of SoxS, the hydrogen bonds to Thr60 are elongated by around 0.2 Å. The C-terminal cysteine Cys16 is almost completely buried and is shielded by residue Pro61 from the protein surface.

There are no polar atoms in proximity to its S atom apart from the S atom of the N-terminal cysteine Cys13. Therefore, based on similar observations in other members of the thioredoxin family, it can be assumed that deprotonation of the sulfhydryl group of Cys16 is rather disfavoured compared with that of the sulfhydryl group of Cys13 and only occurs upon oxidation and disulfide-bridge formation.

A negative charge potential is created below Cys16 by Glu9, which forms two hydrogen bonds, one via Glu9 OE2 to a water molecule and another via Glu9 OE1 to Asp20 OD2. The negative charge of these two carboxylic side chains is only partially reduced via hydrogen-bond interactions of Asp20 OD1 with Arg40 [d(Arg40 NE) = 3.28 Å and d(Arg40 NH2) = 2.78 Å]. In addition to this hydrogen-bond network, the conformation of the active-site loop is further stabilized by hydrogen bonds and tight packing between residues from the preceding β-strand $\beta1$ and the following α-helix $\alpha1$ as well as β-strand $\beta3$. These observations, the activation of the sulfhydryl group of Cys13 and nearly no conformational change between the reduced and oxidized form of SoxS, might indicate a favourable conformation for the oxidized state of SoxS. Thus, a high reduction potential towards its substrate can be rationalized as observed in vitro by the reduction of oxidized Cys110^Y in SoxYZ. However, there are currently no quantitative data available for the pK_a values of the active-site cysteines and for the redox potential of the protein.

3.3. Substrate-binding site

SoxS is exclusively specific for the substrate SoxY of the SoxYZ complex (Rother et al., 2008). In general, thioredoxins are nonspecific with respect to the target proteins, but exhibit a specific thiol–disulfide oxidoreductase redox reaction. Members of the Dsb subgroup are involved in cytochrome c maturation. The Dsb subgroup displays a Trx fold with insertions or extensions which can provide additional specificity determinants for the recognition of target proteins [as observed for Bradyrhizobium japonicum DsbE (Edeling et al., 2002), B. japonicum TlpA (Capitani et al., 2001) and Bacillus subtilis ResA (Colbert et al., 2006)]. SoxS displays the minimal Trx fold without fold insertions or extensions and its substrate specificity towards SoxYZ is unprecedented among thioredoxin-like proteins. Based on the structural features known from other thioredoxin-like proteins, a potential substrate site can be envisaged [e.g. glutaredoxin covalently bound to glutathione (Bushweller et al., 1994; PDB code 2fls) and human thioredoxin bound to peptides derived from NF-κB target sites (Qin et al., 1995, 1996). In Fig. 6 the CPK surface representation and molecular-surface representation of SoxS are displayed, indicating the potential binding site for the target protein. Starting from the

Figure 5
Stereo representation of the hydrogen-bonding network around the thiolate of cysteine Cys13 in the reduced form of SoxS. The main-chain trace of the protein is displayed as a loop representation. The side chains of relevant residues are depicted in ball-and-stick representation and coloured green, with associated N, O and S atoms in blue, red and yellow, respectively. Direct and water-mediated hydrogen bonds are represented by cyan dashed lines. This figure was prepared using PyMOL (DeLano, 2004).

reactive cysteinyl residue Cys13, a shallow rim can roughly be highlighted formed by residues Gln10, Pro11, Gly12, Cys13, Tyr15, Phe59, Thr60, Pro61, Gly77, Tyr78, Pro79, Phe83 and Met87. These residues form a shallow groove on both sides of the active site. The highly conserved *cis*-proline (Pro61) is frequently discussed as one determinant for target binding as the *cis*-conformation exposes the main-chain O atom of the preceding residue for interaction with potential peptide or protein substrates (Qin *et al.*, 1995). Therefore, it is likely that the adjacent residues form part of the interaction surface in SoxS. In close proximity to this Pro61 and the active-site Cys13, a remarkable feature is observed in which the two aromatic residues Tyr15 and Phe59 form a close clamp. It can be envisaged that this clamp serves as a selectivity filter for glycine residues as for example observed in the C-terminal Gly-Gly-[110]Cys-Gly-Gly-COOH motif of SoxY. A similar feature was observed for human glutaredoxin 2 covalently bound to glutathione with a threonine residue in place of Phe59 (PDB code 2fls). Based on the surface topology, no obvious contribution to substrate binding can be assumed for Leu14 other than supporting the orientation of Tyr15.

4. Discussion

Analysis of the structure of SoxS, a protein that is beneficial for chemotrophic sulfur oxidation *in vivo* and that is also specific for its only substrate SoxY *in vitro*, revealed the following general and distinct structural features. (i) The overall structure of SoxS consists of four β-strands surrounded by two α-helices and shows all of the features of the typical thioredoxin fold that is common to all thioredoxins in bacteria and eukaryotes. (ii) This fold contains an N-terminal $\beta\alpha\beta$ motif and a C-terminal $\beta\beta\alpha$ motif connected by a loop in which a third helix is incorporated. In SoxS, this third helix is replaced by a large meandering loop. (iii) Structure-based alignment (Holm & Sander, 1997) identified the cytoplasmic thioredoxin 1 from *E. coli* and the cytoplasmic thioredoxin 2 from *Anabaena* as the proteins with the highest structural homology to the periplasmic SoxS, irrespective of their very low sequence identity.

The structure of SoxS from *P. pantotrophus* was solved both in an oxidized and in a mostly reduced state at 2.1 and 1.9 Å resolution, respectively. The structural alignment of these two SoxS structures showed no significant overall differences in the three-dimensional fold (0.2 Å r.m.s.d.) except for the side-chain rotamer of the two

cysteine residues in the active site and the main-chain conformation of the preceding loop of α-helix $\alpha1$.

SoxS exhibits a kink within α-helix $\alpha1$, which contains the active site at its N-terminal end, caused by a conserved proline residue (Pro25) as is common for other thioredoxins. As a result of this kink, the active site becomes more accessible to the environment. Pro25 is positioned more towards the C-terminal end of this helix, in contrast to the cytoplasmic thioredoxins from *E. coli* and *Anabaena*. In SoxS this specific feature causes a pronounced bulge in this α-helix (Fig. 2). Furthermore, this helix can be divided into two parts, $\alpha1a$ and $\alpha1b$, caused by a short break at residues 23–30 (including a π-helix turn) at the C-terminal end of this helix. A similar feature has been observed in DsbDγ from *E. coli* (Kim *et al.*, 2003).

(a)

(b)
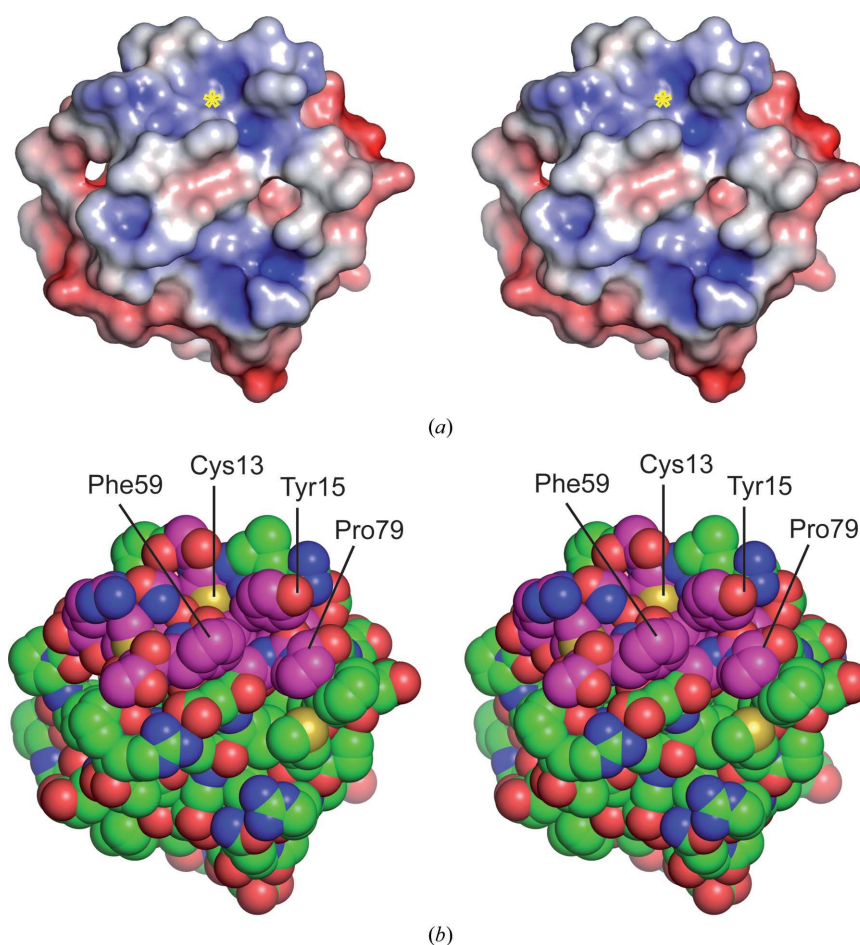
Figure 6
Molecular-surface representation of SoxS. (a) Representation of the molecular surface coloured according to the electrostatic potential. The yellow asterisk indicates the position of the redox-active Cys13. The molecular surface is coloured according to the electrostatic potential as calculated with the program *APBS* (Baker *et al.*, 2001). The molecular surface is colour-ramped according to the electrostatic potential, with red indicating negative potential and blue indicating positive potential; fully saturated colours indicate a potential of $\geq\pm4kT$/e (assuming an ionic strength of 150 m*M*, a protein interior dielectric of 2 and a solvent dielectric of 78.5). The rendered surface representation was prepared with *PyMOL* (DeLano, 2004). (b) The putative binding cleft on the surface of SoxS. The spheres are coloured according to the type of the underlying atom (carbon, green; nitrogen, blue; oxygen, red; sulfur, yellow). For the putative substrate-binding epitope the C atoms are coloured magenta. The S atom of the redox-active cysteinyl residue Cys13 is labelled as well as the aromatic amino-acid residues located at the surface near the active site.

Owing to small variations in their general fold and in the amino-acid sequence within the active site, the members of the thioredoxin family differ significantly in their reactivity and perform a number of different biochemical tasks. The thiols of the Cys-Xaa-Xaa-Cys motif can reversibly form a disulfide bond, enabling them to perform thiol–disulfide exchange reactions. The local concentration of thiol- and disulfide-containing targets together with the redox potential of the thioredoxin characterize the catalyzed reaction as more reductive or more oxidative (Zapun *et al.*, 1993; Aslund *et al.*, 1997; Huber-Wunderlich & Glockshuber, 1998; reviewed by Kadokura *et al.*, 2003). Considerable effort has been made to resolve the molecular determinants of the redox potential of the proteins in the thioredoxin family and its relation to the reduced pK_a of the first active-site cysteine and how the active-site environment modulates this pK_a. Structural, mutational and modelling studies revealed that the pK_a value of the N-terminal cysteine thiol can be modulated by the type of amino-acid residue in the -*X*-*X*- positions of the C*XX*C motif, which affect its nucleophilicity and the overall redox potential (Dyson *et al.*, 1997; Mossner *et al.*, 1998). In addition, different structural features have been discussed for the stabilization of the most stable oxidation state of the protein. Throughout the family, a conserved contribution to the stabilization of the reduced state with a thiolate anion of the first cysteine in the active site has been attributed to the charge of the following α-helix dipole and through hydrogen bonding to the sulfhydryl of the partner cysteinyl residue (Kallis & Holmgren, 1980; Gleason *et al.*, 1990; Foloppe & Nilsson, 2004). This stabilization is further assisted by the type of amino-acid residue that flanks the active-site cysteines, enabling a characteristic hydrogen-bond network (Foloppe & Nilsson, 2004). The deprotonated thiolate state of SoxS Cys13 is favoured by specific hydrogen bonds and charges. Another contribution comes from conformational hydrogen-bond stabilization of the active-site loop. Furthermore, the conformation of a disulfide bridge between the two active-site cysteines can be favoured to a greater or lesser extent by the packing between the β-strand and the α-helix containing the active site within their connecting loop (Goulding *et al.*, 2004; Guddat *et al.*, 1998).

Based on multiple sequence alignment with all known members of the thioredoxin family, SoxS is grouped with the members of the DsbC subfamily. These proteins are localized in the periplasmic space, as is SoxS. The disulfide oxidoreductases (Dsbs) of the extracellular compartment are proteins that catalyze disulfide-bond formation and/or isomerization of incorrect disulfides by reducing disulfide bonds, making them accessible for reoxidation by the action of DsbA. New evidence indicates a similar function for SoxS. However, whereas DsbC is predominantly kept in its reduced form owing to its dimerization, which prevents its oxidation by DsbB (Bader *et al.*, 2001), there is no experimental evidence for the dimerization of SoxS. The mechanism by which SoxS is kept in or converted into its reduced state is unknown. Potential candidates are SoxW and SoxV.

The active-site sequence Pro-Gly-[13]Cys-Leu-Tyr-[16]Cys of SoxS could not be identified in any other known member of the thioredoxin or glutaredoxin superfamilies. Thus, we cannot directly compare SoxS with any other characterized members of these thiol/disulfide oxidoreductase families, but have to draw conclusions based on similarities in structure and function. The active-site sequence of SoxS is most closely related to the Pro-Gly-Cys-Pro-Tyr-Cys sequence of some glutaredoxin proteins and to the Pro-Xaa-Cys-Xaa-Tyr-Cys sequences of some members of the DsbC and DsbG subfamilies. Interestingly, the physiological function of these proteins is closer to the reduction of disulfide bridges, such as deglutathionylation in the case of glutaredoxins or the breakage of misformed disulfide bridges for disulfide-bridge isomerization in the case of DsbC and DsbG. These similarities in the active-site sequence are partially reflected by some potential similarities based on the physiological role of SoxS. SoxS is most likely to have high substrate specificity, which is in contrast to most thioredoxins and similar to the Dsb proteins of the cytochrome *c* maturation pathway. The substrate of SoxS (the C-terminus of SoxY) has structural features that are comparable with those of oxidized glutathione. By reductive activation of SoxY, the physiological role of SoxS is closer to the reduction of cysteinyl groups than to the oxidations performed by DsbC and DsbG.

Interestingly, most thioredoxin-like disulfide-bond oxidoreductases predominantly act on proteins with intramolecularly formed disulfide bonds. The active site of SoxYZ is Cys110[Y], which is exposed at the carboxy-terminal end of SoxY and located in the Gly-Gly-[110]Cys-Gly-Gly-COOH moiety (Quentmeier & Friedrich, 2001; Sauve *et al.*, 2007). The thiol of Cys110[Y] undergoes nonspecific oxidation, *e.g.* the formation of a Cys110[Y]–Cys110[Y]-linked catalytically inactive Sox(YZ)$_2$ heterotetramer (Quentmeier *et al.*, 2007, 2008). As SoxS is target-specific for SoxY, SoxS is proposed to specifically reduce the interprotein disulfide bridge between the two SoxY proteins within the tetramer (Rother *et al.*, 2008). Recently, the structures of SoxY alone (Stout *et al.*, 2007) and of SoxY in complex with SoxZ (Sauvé *et al.*, 2007) have been determined by X-ray crystallography. In all crystal structures, the C-terminal part of SoxY with the proposed active cysteine residue (Cys110) was not well resolved, indicating rather flexible properties of the Gly-Gly-[110]Cys-Gly-Gly-COOH moiety. The disulfide cross-linked C-terminal peptides share structural similarities with oxidized glutathione (GSSG). Therefore, the similarities of SoxS to the glutaredoxin (Grx) protein family based on its active-site sequence can be rationalized. The glutaredoxins are distantly related to thioredoxin and operate *via* a similar mechanism. Glutaredoxins are kept in their reduced state by reaction with reduced glutathione in the cytoplasm. In contrast to the active site of thioredoxins, glutaredoxins have a high-affinity site that only interacts with the GSH moiety and not with the protein (Foloppe & Nilsson, 2004). Some members of the dithiol Grxs have been shown to reduce intramolecular protein disulfides using a mechanism similar to that described for Trx (Berndt *et al.*, 2008). However, the majority of Grxs act on reducing oxidized cysteinyl

groups *via* cysteinyl-glutathione intermediates. In this deglutathionylation reaction of Grxs only the N-terminal cysteine residue within the active site is used (Berndt *et al.*, 2007), forming a glutathionylated Grx. Transferring the Grx mechanism to the SoxS–SoxYZ system would result in the activation of one SoxYZ heterodimer, leaving the second SoxYZ heterodimer covalently attached to SoxS. To further release the second reduced and active SoxYZ protein, the cleavage of the covalently bound SoxYZ by nucleophilic attack of the C-terminal active-site cysteine is needed as observed in the classical Trx mechanism. Current research is focused on the reduction reactions of variously modified $Cys110^Y$ of SoxYZ. Further analysis of substrate-bound structures of SoxS and/or mutagenesis studies are needed to localize the regions of the partner proteins that are specific for the protein–protein interaction.

5. Conclusions

Our work describes the three-dimensional structure of SoxS from *P. pantotrophus*, an unusual periplasmic thioredoxin-like protein. SoxS resembles a chimera formed of Trx-, Grx- and DsbC-type proteins. SoxS displays the basic core thioredoxin-type fold. From the Grx family it has obtained an affinity towards GSSG-like peptide features required for the recognition of the intermolecular disulfide bridge at the C-terminal end of dimeric SoxY. SoxS shares its disulfide-reduction potential, releasing two reduced cysteinyl residues from the partner protein, with the Trx and DsbC families. Based on its proposed physiological function, SoxS is best described as a thiol–disulfide oxidoreductase specific for a single substrate, SoxYZ. The unique structural features of SoxS suggest a variety of biochemical reduction reactions. This analysis is in accordance with the biochemical data, which suggest re-activation of SoxYZ rather than maturation. Together with the other structures of the *sox* gene cluster, the structures of SoxS will help to understand the mode of activation of the sulfur-oxidizing cycle in *P. pantotrophus* and similar organisms.

References

Aslund, F., Berndt, K. D. & Holmgren, A. (1997). *J. Biol. Chem.* **272**, 30780–30786.

Bader, M. W., Hiniker, A., Regeimbal, J., Goldstone, D., Haebel, P. W., Riemer, J., Metcalf, P. & Bardwell, J. C. (2001). *EMBO J.* **20**, 1555–1562.

Bader, M., Muse, W., Ballou, D. P., Gassner, C. & Bardwell, J. C. (1999). *Cell*, **98**, 217–227.

Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.

Banaszak, K., Mechin, I., Frost, G. & Rypniewski, W. (2004). *Acta Cryst.* D**60**, 1747–1752.

Bardischewsky, F., Fischer, J., Holler, B. & Friedrich, C. G. (2006). *Microbiology*, **152**, 465–472.

Bardischewsky, F. & Friedrich, C. G. (2001). *FEMS Microbiol. Lett.* **202**, 215–220.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Berndt, C., Lillig, C. H. & Holmgren, A. (2007). *Am. J. Physiol. Heart Circ. Physiol.* **292**, H1227–H1236.

Berndt, C., Lillig, C. H. & Holmgren, A. (2008). *Biochim. Biophys. Acta*, **1783**, 641–650.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Bushweller, J. H., Billeter, M., Holmgren, A. & Wüthrich, K. (1994). *J. Mol. Biol.* **235**, 1585–1597.

Capitani, G., Rossmann, R., Sargent, D. F., Grütter, M. G., Richmond, T. J. & Hennecke, H. (2001). *J. Mol. Biol.* **311**, 1037–1048.

Colbert, C. L., Wu, Q., Erbel, P. J., Gardner, K. H. & Deisenhofer, J. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4410–4415.

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

Cowtan, K. & Main, P. (1998). *Acta Cryst.* D**54**, 487–493.

DeLano, W. L. (2004). *The PyMOL Molecular Graphics System.* http://www.pymol.org.

Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.

Dyson, H. J., Jeng, M. F., Tennant, L. L., Slaby, I., Lindell, M., Cui, D. S., Kuprin, S. & Holmgren, A. (1997). *Biochemistry*, **36**, 2622–2636.

Edeling, M. A., Guddat, L. W., Fabianek, R. A., Thony-Meyer, L. & Martin, J. L. (2002). *Structure*, **10**, 973–979.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Erlendsson, L. S., Acheson, R. M., Hederstedt, L. & Le Brun, N. E. (2003). *J. Biol. Chem.* **278**, 17852–17858.

Foloppe, N. & Nilsson, L. (2004). *Structure*, **12**, 289–300.

Friedrich, C. G., Rother, D., Bardischewsky, F., Quentmeier, A. & Fischer, J. (2001). *Appl. Environ. Microbiol.* **67**, 2873–2882.

Gleason, F. K., Lim, C. J., Gerami-Nejad, M. & Fuchs, J. A. (1990). *Biochemistry*, **29**, 3701–3709.

Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.

Goulding, C. W., Apostol, M. I., Gleiter, S., Parseghian, A., Bardwell, J., Gennaro, M. & Eisenberg, D. (2004). *J. Biol. Chem.* **279**, 3516–3524.

Guddat, L. W., Bardwell, J. C. & Martin, J. L. (1998). *Structure*, **6**, 757–767.

Guerrero, S. A., Hecht, H.-J., Hofmann, B., Biebl, H. & Singh, M. (2001). *Appl. Microbiol. Biotechnol.* **56**, 718–723.

Hao, Q. (2004). *J. Appl. Cryst.* **37**, 498–499.

Heras, B., Edeling, M. A., Schirra, H. J., Raina, S. & Martin, J. L. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 8876–8881.

Holm, L. & Sander, C. (1996). *Science*, **273**, 595–603.

Holm, L. & Sander, C. (1997). *Nucleic Acids Res.* **25**, 231–234.

Huber-Wunderlich, M. & Glockshuber, R. (1998). *Fold. Des.* **3**, 161–171.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 67–72.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Kabsch, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.

Kadokura, H., Katzen, F. & Beckwith, J. (2003). *Annu. Rev. Biochem.* **72**, 111–135.

Kallis, G. B. & Holmgren, A. (1980). *J. Biol. Chem.* **255**, 10261–10265.

Katti, S. K., LeMaster, D. M. & Eklund, H. (1990). *J. Mol. Biol.* **212**, 167–184.

Kim, J. H., Kim, S. J., Jeong, D. G., Son, J. H. & Ryu, S. E. (2003). *FEBS Lett.* **543**, 164–169.

Kranz, R., Lill, R., Goldman, B., Bonnard, G. & Merchant, S. (1998). *Mol. Microbiol.* **29**, 383–396.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–495.

Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.

Mossner, E., Huber-Wunderlich, M. & Glockshuber, R. (1998). *Protein Sci.* **7**, 1233–1244.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Orawski, G., Bardischewsky, F., Quentmeier, A., Rother, D. & Friedrich, C. G. (2007). *Microbiology*, **153**, 1081–1086.

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

Pan, J. L. & Bardwell, J. C. (2006). *Protein Sci.* **15**, 2217–2227.

Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Qin, J., Clore, G. M., Kennedy, W. M., Huth, J. R. & Gronenborn, A. M. (1995). *Structure*, **3**, 289–297.

Qin, J., Clore, G. M., Kennedy, W. P., Kuszewski, J. & Gronenborn, A. M. (1996). *Structure*, **4**, 613–620.

Quentmeier, A. & Friedrich, C. G. (2001). *FEBS Lett.* **503**, 168–172.

Quentmeier, A., Janning, P., Hellwig, P. & Friedrich, C. G. (2007). *Biochemistry*, **46**, 10990–10998.

Quentmeier, A., Li, L. & Friedrich, C. G. (2008). *FEBS Lett.* **582**, 3701–3704.

Rainey, F. A., Kelly, D. P., Stackebrandt, E., Burghardt, J., Hiraishi, A., Katayama, Y. & Wood, A. P. (1999). *Int. J. Syst. Bacteriol.* **49**, 645–651.

Rother, D., Henrich, H. J., Quentmeier, A., Bardischewsky, F. & Friedrich, C. G. (2001). *J. Bacteriol.* **183**, 4499–4508.

Rother, D., Ringk, J. & Friedrich, C. G. (2008). *Microbiology*, **154**, 1980–1988.

Saarinen, M., Gleason, F. K. & Eklund, H. (1995). *Structure*, **3**, 1097–1108.

Sauvé, V., Bruno, S., Berks, B. C. & Hemmings, A. M. (2007). *J. Biol. Chem.* **282**, 23194–23204.

Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.

Sheldrick, G. M. (2002). *Z. Kristallogr.* **217**, 644–650.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

Sheldrick, G. M., Hauptman, H. A., Weeks, C. M., Miller, R. & Usón, I. (2001). In *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold. Dordrecht: Kluwer Academic Publishers.

Stout, J., Van Driessche, G., Savvides, S. N. & Van Beeumen, J. (2007). *Protein Sci.* **16**, 589–601.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Tickle, I. J., Laskowski, R. A. & Moss, D. S. (2000). *Acta Cryst.* D**56**, 442–450.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Vaguine, A. A., Richelle, J. & Wodak, S. J. (1999). *Acta Cryst.* D**55**, 191–205.

Weik, M., Bergès, J., Raves, M. L., Gros, P., McSweeney, S., Silman, I., Sussman, J. L., Houée-Levin, C. & Ravelli, R. B. G. (2002). *J. Synchrotron Rad.* **9**, 342–346.

Weik, M., Ravelli, R. B., Kryger, G., McSweeney, S., Raves, M. L., Harel, M., Gros, P., Silman, I., Kroon, J. & Sussman, J. L. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 623–628.

Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.

Zapun, A., Bardwell, J. C. & Creighton, T. E. (1993). *Biochemistry*, **32**, 5083–5092.

Zeller, T. & Klug, G. (2006). *Naturwissenschaften*, **93**, 259–266.