Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of OsmC from *Escherichia coli*: a salt-shock-induced protein

Dong Hae Shin,[a]‡ In-Geol Choi,[b]‡ Didier Busso,[b] Jaru Jancarik,[b] Hisao Yokota,[a] Rosalind Kim[a] and Sung-Hou Kim[a,b]*

[a]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA, and [b]Department of Chemistry, University of California, Berkeley, California 94720-5230, USA

‡ These authors contributed equally.

Correspondence e-mail: shkim@cchem.berkeley.edu

Received 12 December 2003
Accepted 3 March 2004

PDB Reference: OsmC, 1nye, r1nyesf.

The crystal structure of an osmotically inducible protein (OsmC) from *Escherichia coli* has been determined at 2.4 Å resolution. OsmC is a representative protein of the OsmC sequence family, which is composed of three sequence subfamilies. The structure of OsmC provides a view of a salt-shock-induced protein. Two identical monomers form a cylindrically shaped dimer in which six helices are located on the inside and two six-stranded β-sheets wrap around these helices. Structural comparison suggests that the OsmC sequence family has a peroxiredoxin function and has a unique structure compared with other peroxiredoxin families. A detailed analysis of structures and sequence comparisons in the OsmC sequence family revealed that each subfamily has unique motifs. In addition, the molecular function of the OsmC sequence family is discussed based on structural comparisons among the subfamily members.

1. Introduction

Bacterial cells often encounter unfavorable growth conditions in natural environments. *Escherichia coli*, a non-sporulating enterobacteriae, undergoes a global programmed modification of its gene-expression pattern and this results in the acquisition of resistance to chemical and physical stresses, such as heat, oxidative agents or hyperosmotic shock (Kolter *et al.*, 1993). These properties provide a better chance of survival of the cells under adverse conditions. One key regulator of this genetic programme is the product of the *rpoS* gene, an RNA polymerase sigma factor called RpoS or σ^s (Lange & Hengge-Aronis, 1991). Many genes controlled by σ^s are inducible by elevated osmolarity (Hengge-Aronis, 1996). These include genes such as *otsA* and *otsB* that encode enzymes for trehalose biosynthesis (Kaasen *et al.*, 1992) and several members of yet unknown function such as *OsmB* and *OsmE* that encode outer membrane lipoproteins (Jung *et al.*, 1990), *OsmY* that encodes a periplasmic protein (Yim & Villarejo, 1992) and *OsmC* that encodes a putative inner membrane protein (Völker *et al.*, 1998). These four genes are induced upon the entry of the cell into stationary phase in a σ^s-dependent manner and also by elevated osmolarity. Recently, genes homologous to *OsmC* have been identified in a variety of bacterial species (Fig. 1).

The OsmC sequence family, which is found only in prokaryotes, is composed of three subfamilies in the Pfam database (Bateman *et al.*, 2000): subfamily I, also called the OsmC subfamily (seven members), subfamily II, called the Ohr subfamily (organic hydrogen peroxide resistance protein, 20 members), and subfamily III of unknown function (39 members). Of these, crystal structures of members of subfamilies II and III have been reported. The crystal structure

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

Table 1
Statistics of the peak-wavelength SAD data set.

Values in parentheses refer to the highest resolution shell, which is 2.44–2.40 Å.

Wavelength (Å)	0.9793
Resolution (Å)	46.6–2.4
Redundancy	3.8 (3.4)
Unique reflections	74939 (3736)
Completeness (%)	98.2 (97.9)
$I/\sigma(I)$	25.6 (2.2)
R_{sym}† (%)	6.0 (57.7)

† $R_{sym} = \sum_{hkl}\sum_i |I_{hkl,i} - \langle I\rangle_{hkl}|/\sum|\langle I\rangle_{hkl}|$.

and biochemical assay of Ohr revealed that it has peroxiredoxin (Prx) activity and utilizes highly reactive cysteine thiol groups to elicit hydroperoxide reduction (Lesniak *et al.*, 2002). Two protein crystal structures from subfamily III, MPN625 (gi|1673883) from *Mycoplasma pneumoniae* (Choi *et al.*, 2003) and yhfA from *E. coli* (gi|15803870; PDB code 1ml8), have been determined. However, the three-dimensional structures and molecular (biochemical and biophysical) functions of members of salt-shock-induced subfamily I have not been elucidated until now.

In order to obtain the structure and possible inference of the molecular function of a member of subfamily I, we have determined the crystal structure of OsmC from *E. coli* (*Ec*OsmC) and compared it with those of subfamilies II and III.

While this paper was in preparation, another manuscript appeared presenting the same structure (Lesniak *et al.*, 2003). A comparison of the two structures is presented.

Table 2
Crystal parameters and refinement statistics.

Space group	$P2_1$
Unit-cell parameters (Å, °)	a = 49.57, b = 90.39, c = 112.71, β = 93.93
Volume fraction of protein (%)	46.4
V_M (Å^3 Da^{-1})	2.39
Total No. residues	855
Total No. modeled His_6-tag residues	56
Total non-H atoms	6911
No. protein atoms	6332
No. His_6-tag atoms	492
No. water molecules	87
Average temperature factors (Å^2)	
Protein	61.6
His_6-tag	87.0
Solvent	59.2
Resolution range of reflections used (Å)	20.0–2.4
Amplitude cutoff	0.0σ
R factor (%)	22.7
Free R factor (%)	28.7
Stereochemical ideality	
Bond (Å)	0.010
Angle (°)	1.64
Improper (°)	0.88
Dihedral (°)	25.01
Ramachandran plot: residues in (%)	
Most favored and additionally allowed regions	98.9
Generously allowed regions	1.1

2. Materials and methods

2.1. Cloning

Cloning primers (Operon, Alameda, CA, USA) for *OsmC* gene PCR amplification from genomic DNA contained an *Nde*I restriction site in the forward primer (5′- CATATGA-

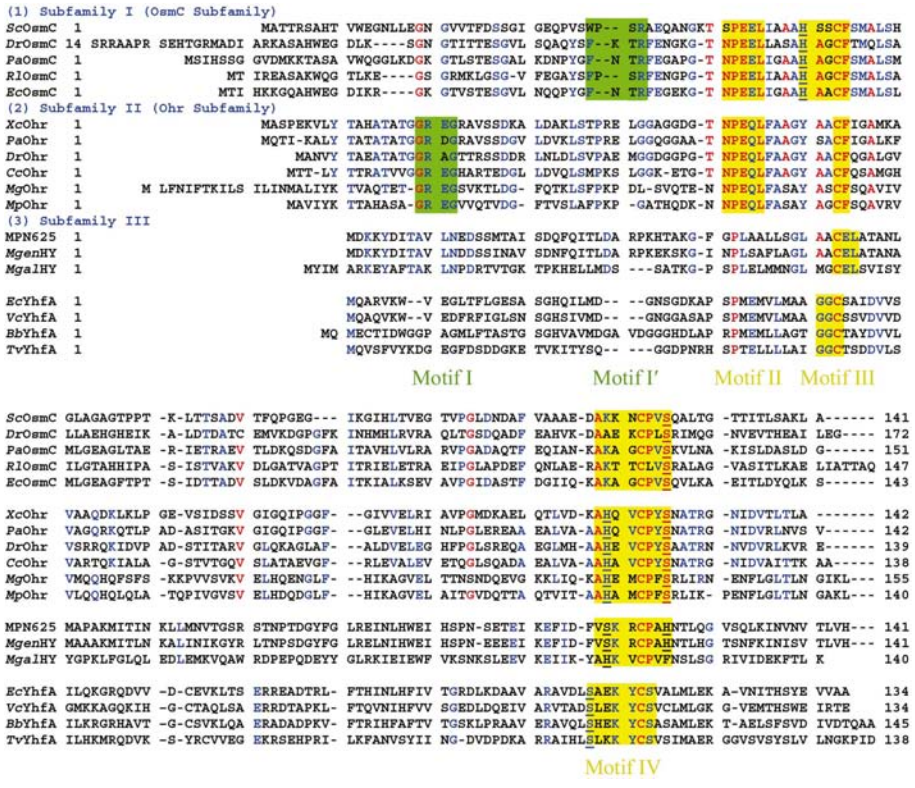

Figure 1
Sequence comparison among OsmC sequence family. OsmC sequence subfamilies are aligned with several homologs. Abbreviations are as follows. (1) OsmC from *Streptomyces coelicolor* (*Sc*OsmC), *Deinococcus radiodurans* (*Dr*OsmC), *Pseudomonas aeruginosa* (*Pa*OsmC), *Rhizobium loti* (*Rl*OsmC) and *E. coli* (*Ec*OsmC). (2) Ohr from *Xanthomonas campestris* (*Xc*Ohr), *Pseudomonas aeruginosa* (*Pa*Ohr), *D. radiodurans* (*Dr*Ohr), *Caulobacter crescentus* (*Cc*Ohr), *Mycoplasma genitalium* (*Mg*Ohr) and *M. pneumoniae* (*Mp*Ohr). (3) MPN625 homologs from *M. genitalium* (*Mgen*HY) and *M. gallisepticum* (*Mgal*HY). (4) YhfA from *E. coli* (*Ec*YhfA) and YhfA homologs from *Vibrio cholerae* (*Vc*YhfA), *Bordetella bronchisepticac* (*Bb*YhfA) and *Thermoplasma volcanium* (*Tv*YhfA). The red characters represent residues highly conserved in OsmC and Ohr subfamilies and the residues in blue represent sequences highly conserved in each subfamily. The yellow and green boxes represent for the conserved motifs in each subfamily primarily based on the OsmC and Ohr subfamilies (Atichartpongkul *et al.*, 2001). Underlined characters represent highly conserved histidine and serine residues neighboring the two conserved cysteines. The amino-acid sequence from residues 139 to 153 is not shown in the case of *Tv*YhfA.

CAATCCATAAGAAAGGTCAGGC) and a *Bam*HI site in the reverse primer (5′- GGATCCTTACGATTTCAACTGGTAATCCAGC). PCR was performed using Deep Vent DNA Polymerase (New England Biolabs, Inc., Beverly, MA, USA) and *E. coli* genomic DNA. The PCR product was cloned into the pCR-BluntII-TOPO vector (Invitrogen) and the *OsmC* gene insert was confirmed by DNA sequencing. The amplified TOPO vector was cut with *Nde*I and *Bam*HI and the gene insert was purified by agarose gel electrophoresis extraction. This insert was cloned in a pSKB3 expression system (a gift from Steve Burley, Rockefeller University, NY, USA) and transformed into B834(DE3)/pSJS1244 (Kim *et al.*, 1998).

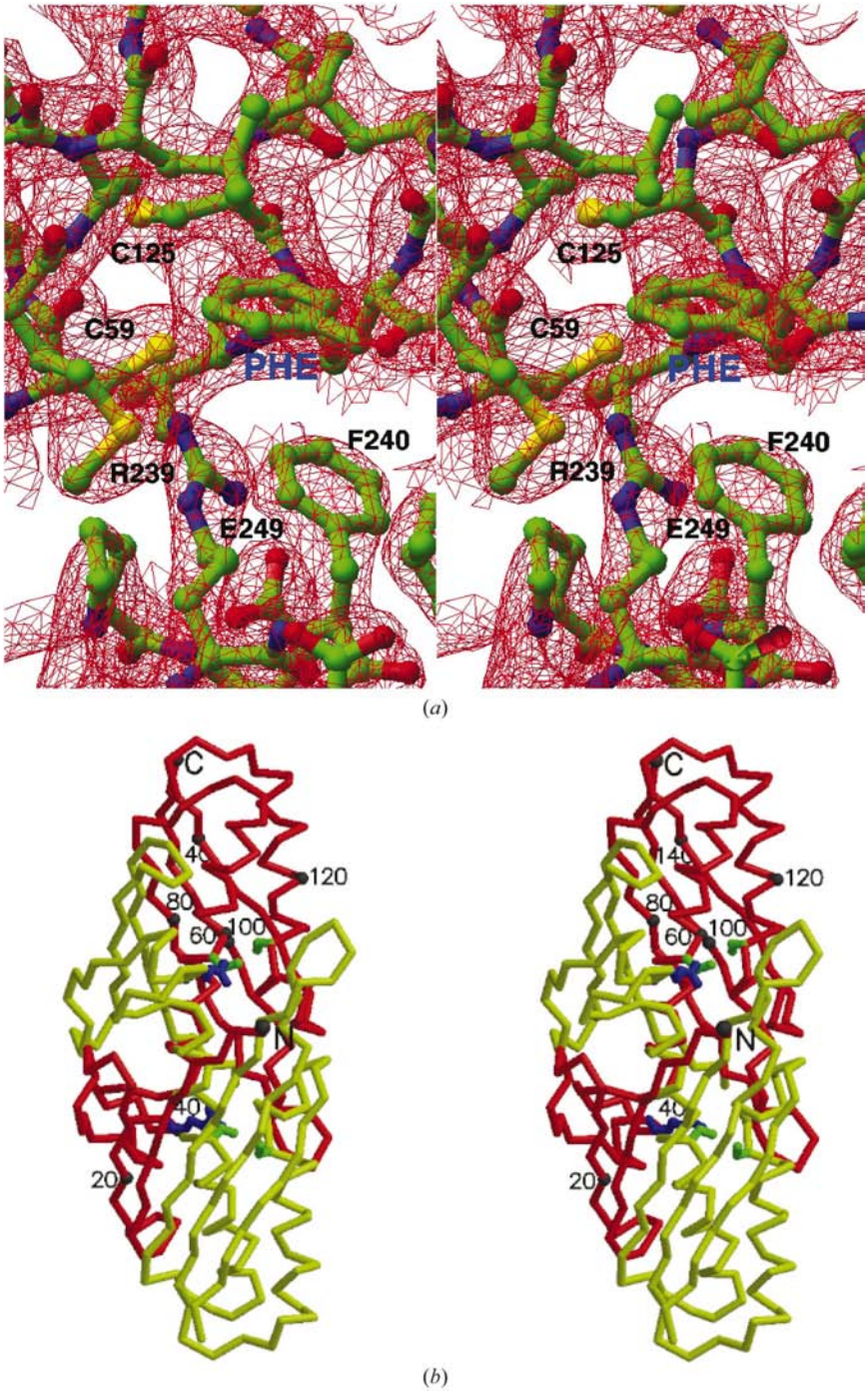

Figure 2
(*a*) Steroview of the putative active site of *Ec*OsmC. The $2F_o - F_c$ map from the final refined phase was calculated using all reflection data between 20 and 2.4 Å. The figure was generated using the program *RIBBONS* (Carson, 1991). The red net represents the electron-density map. The residues are represented by a ball-and-stick model. PHE represents the phenylalanine residue from the His_6-tag. Residues Arg239, Phe240 and Glu249 of one monomer correpsond to Arg39, Phe30 and Glu49 of the other monomer, respectively. Blue represents N atoms, red O atoms and green C atoms. (*b*) A stereo drawing of a C^{α} trace of *Ec*OsmC. Two subunits are represented by a thick line with different colors. Conserved cysteines (Cys59 and Cys125, light green) and Arg39 (blue) are represented by a ball-and-stick model. Every 20th residue is numbered and represented by a dot. The N-terminus (residue Met1) and C-terminus (residue Ser143) are labeled. The figure was generated using *MOLSCRIPT* (Kraulis, 1991).

2.2. Protein expression, purification and crystallization

A selenomethionine derivative of the protein was expressed in a methionine auxotroph, *E. coli* strain B834(DE3)/pSJS1244, grown in M9 medium supplied with selenomethionine (Kim *et al.*, 1998). The cloned *OsmC* gene containing an N-terminal His_6-tag was affinity purified from the soluble fraction using Talon metal-affinity resin (Clonetech, Palo Alto, CA, USA). The target protein was bound onto a 5 ml HiTrap Q column (Pharmacia Biotech, Uppsala, Sweden) and eluted with 150 m*M* NaCl. Since the His_6-tag was not cleaved, the 25-residue tag MGSSHHHHHHDYDIPTTENLYFQGH remained fused to the N-terminus of the protein. Using the sparse-matrix crystal-screening method (Jancarik & Kim, 1991), rod-shaped crystals appeared at room temperature. The optimized crystallization condition was 1 µl protein solution (9 mg ml^{-1}) in 50 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl and 5 m*M* DTT, mixed with 1 µl reservoir solution containing 0.2 *M* magnesium formate and 20% PEG 3350.

2.3. Data collection and reduction

X-ray diffraction data sets were collected at one wavelength corresponding to the selenium absorbance peak (λ = 0.9793 Å) at the Macromolecular Crystallography Facility beamline 5.0.2 at the Advanced Light Source at Lawrence Berkeley National Laboratory using an Area Detector System Co. (Poway, CA, USA) Quantum 4 CCD detector placed 200 mm from the sample. X-ray diffraction data were processed and scaled using *DENZO* and *SCALEPACK* from the *HKL* program

Table 3
Structural comparison of OsmC sequence family structures.

All the values are from the results of structural comparison performed using the Combinatorial Extension (CE) Method (http://cl.sdsc.edu/ce.html; Shindyalov & Bourne, 1998).

	R.m.s. deviation (Å)	Aligned amino acids	*Z* score	Sequence identity (%)
OsmC/Ohr	2.2	128	6.1	21.1
OsmC/MPN625	2.4	129	5.0	17.1
OsmC/yhfA	2.3	118	5.6	15.3
Ohr/MPN625	2.6	133	6.0	15.8
Ohr/yhfA	3.3	125	5.3	8.8
MPN625/yhfA	1.7	128	6.0	17.2

suite (Otwinowski & Minor, 1997). Data statistics are summarized in Table 1.

2.4. Structure determination and refinement

The program *SOLVE* (Terwilliger & Berendzen, 1999) was used to locate the selenium sites in the crystal and to calculate initial phases at 2.4 Å resolution. The partial models were built based on an initial map using the program *O* (Jones *et al.*, 1991). The presence of six molecules in the asymmetric unit was found based on partial models and this was used to find the non-crystallographic symmetry (NCS) matrices. Sixfold NCS density averaging was carried out using the *DM* program in the *CCP*4 package (Dodson *et al.*, 1997). The improved density map clearly revealed the presence of three homodimers in the asymmetric unit. The model was then refined using the program *CNS* (Brünger *et al.*, 1998), with 10% of the data randomly chosen for free *R*-factor cross-validation. The NCS restraints were used for the entire refinement steps and were released for the final cycle of refinement. In one monomer model we could not observe residues 41–43, but we could see all 143 residues in the other five monomer models. Four out of the six monomer models display a His_6-tag and the lengths of the His_6-tags observed are eight, ten, 19 and 19 residues, respectively. The refinement statistics are shown in Table 2.

3. Results and discussion

3.1. Quality of the model

Five of the six final models in the asymmetric unit contain all 143 amino-acid residues. One monomer model is missing residues 41–43. Most of the residues are well defined by electron density in the refined models of *Ec*OsmC, including some of the His_6-tag residues (Fig. 2*a*). The final model has been refined to 2.4 Å resolution to a crystallographic *R* factor of 22.7%. The root-mean-square (r.m.s.) deviations from ideal stereochemistry are 0.010 Å for bond lengths, 1.64° for bond angles and 0.88° for improper angles. The averaged *B* factors for main-chain atoms and side-chain atoms are 57.5 and 64.7 Å^2, respectively. The averaged *B* factor of all modeled His_6-tag residues is 87.0 Å^2. Table 2 summarizes the refinement statistics as well as the model-quality parameters. The mean positional error in atomic coordinates for the refined model is estimated to be ∼0.3 Å by a Luzzati plot (Luzzati, 1952). All residues lie in the allowed region of the Ramachandran plot produced with *PROCHECK* (Laskowski *et al.*, 1993). The C^{α} r.m.s. deviations among the three dimers in the asymmetric unit are in the range 0.4–0.8 Å.

3.2. Overall structure

The *Ec*OsmC monomer has approximate dimensions of 50 × 35 × 30 Å with two domains (Fig. 2*b*). The N-terminal half is composed of three β-strands forming a β-sheet and the C-terminal half is a mixed α–β structure composed of a two-layer αβ-sandwich in the CATH classification (Orengo *et al.*, 1997). Two identical monomers form a cylindrically shaped dimer in which six helices are located on the inside and two six-stranded β-sheets formed by domain swapping wrap around these helices (Fig. 2*b*).

3.3. Overall structural comparison among subfamilies

The sequence alignment based on tertiary structures indicates that the secondary-structures are well aligned with each other except for one helix (H1) and several loops (Fig. 3). Helix H1 is unique to the *Ec*OsmC structure. Loops L2, L4 and L6 form the walls of the active-site pocket and vary in size and sequence, as shown in the sequence alignment and Fig. 4.

Figure 3
Sequence alignment based on tertiary structures. Sequence alignment of *Ec*OsmC, Ohr, MPN625 and yhfA based on three-dimensional structure comparisons using the Combinatorial Extension (CE) Method (http://cl.sdsc.edu/ce.html; Shindyalov & Bourne, 1998). The secondary structure derived from *Ec*OsmC is shown above the sequence. Blue characters represent amino-acid residues belonging to α-helix, green 3_{10}-helix and red β-strand. The yellow-shaded regions refer to the conserved motifs as in Fig. 1. GR*X*G (motif I) and F*XX*R (motif I′) are motifs unique to the Ohr and *Ec*OsmC subfamilies, respectively, based on structural comparison and sequence alignment (Fig. 1). Underlined characters represent highly conserved histidine and serine residues neighboring the two conserved cysteines. The sequence numbers refer to that of *Ec*OsmC.

The averaged r.m.s. deviations of C^{α} atoms among the subfamilies are listed in Table 3. The structural comparison indicates that the overall structures are quite similar to each other, although the pairwise sequence identity only ranges between 8.8 and 21.1%.

A sequence alignment of the OsmC and Ohr subfamilies suggests the presence of four conserved regions (Völker *et al.*, 1998; Atichartpongkul *et al.*, 2001; Fig. 1): (i) a glycine residue near the amino-terminus, (ii) an NPEQ/E*X*L motif, (iii) a CF motif and (iv) an A*XXX*CP*X*S motif. The OsmC family structures reveal that the latter three motifs comprise the active site in all three subfamilies. Interestingly, the active site is located at the interface of two subunits (Fig. 5): two conserved cysteine residues (Cys59 and Cys125) from motif III and IV from one subunit are located at the bottom of the active sites and the NPEQ/E*X*L motif is contributed by the other subunit.

All the structures from the OsmC sequence family revealed the presence of homodimers. The extensive and tight interaction between monomers is found in all four structures representing the three subfamilies of the OsmC sequence family. The approximate molecular surface area buried by the dimers of *Ec*OsmC, Ohr, MNP625 and yhfA are 2070, 2210, 2010 and 1360 Å^2 per monomer, respectively. The relatively small buried molecular surface area of yhfA is a consequence of the short N-terminal loop and the absence of three residues (Ala77, Asp78 and Thr79) that contact with the other subunit in other subfamily structures (Fig. 4).

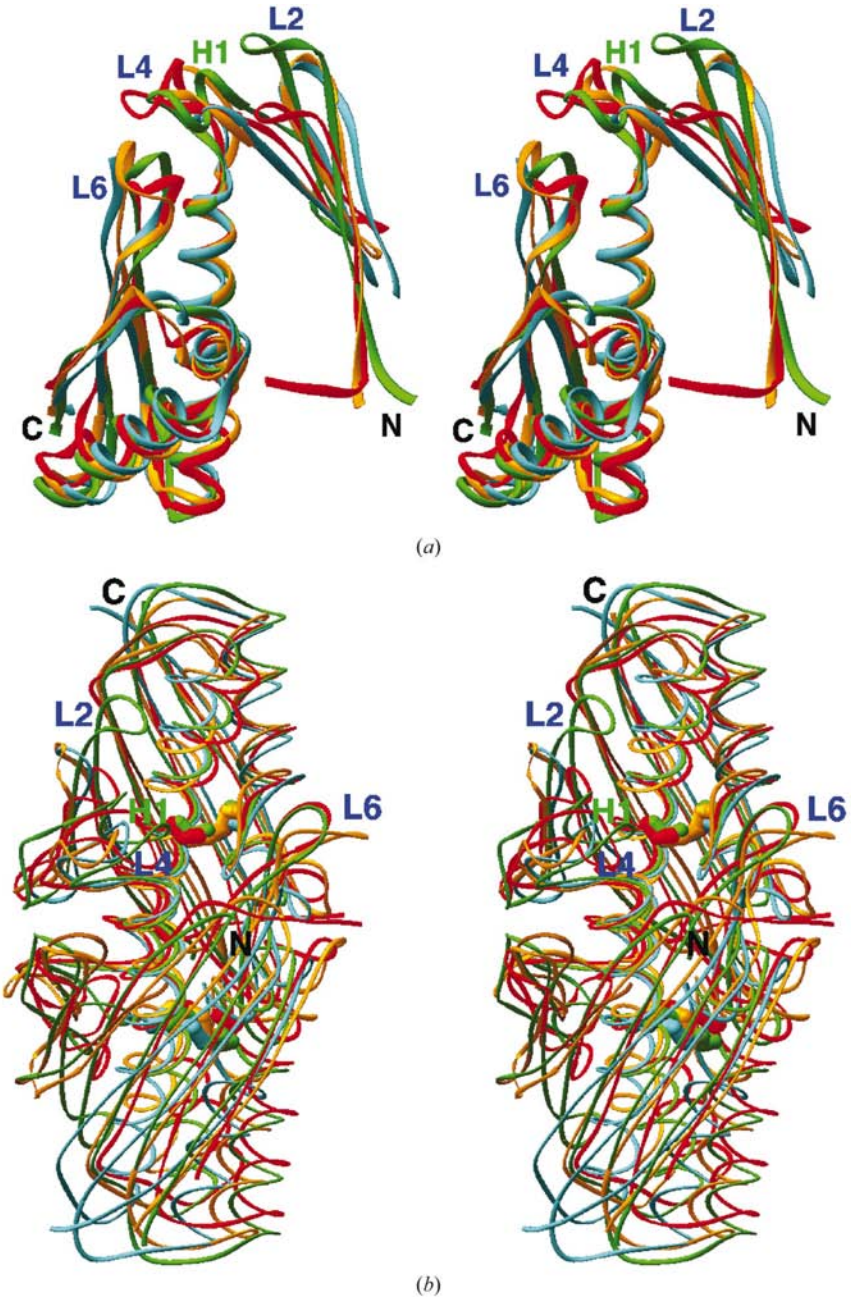

Figure 4
Structural comparison. The stereo drawing was generated using the program *RIBBONS* (Carson, 1991). *Ec*OsmC is colored green, Ohr (PDB code 1n2f) red, MPN625 (1lql) yellow and yhfA (1ml8) cyan. The two conserved cysteine residues in the active sites are represented by a ball-and-stick model with corresponding color for each structure. The three loops (L2, L4 and L4) that confine the active-site pocket are labeled. The N- and C-termini of one subunit of *Ec*OsmC are labeled. (*a*) Monomer, (*b*) dimer. The figure was drawn with the operational matrix obtained from (*a*).

3.4. Structural comparison of the active sites of *Ec*OsmC, subfamily II and subfamily III

When the active-site residues of the four structures were compared, those of *Ec*OsmC and Ohr were surprisingly similar despite their belonging to different subfamilies. Specifically, the arginine (Arg18) residue neighboring the active cysteine (Cys60) in Ohr (Lesniak *et al.*, 2002) is also structurally conserved (Arg39) in *Ec*OsmC (Fig. 6). Biochemical studies of Ohr revealed that Arg18 is important for removing reactive oxygen species (ROS; Lesniak *et al.*, 2002). The role of a positively charged residue in stabilizing the reactive thiol is known to be important for Prx activity (Choi *et al.*, 1998). Therefore, structural similarity of the active site of *Ec*OsmC with Ohr strongly suggests that one of the biochemical functions of *Ec*OsmC may also be to reduce peroxide generated during stress. Although the side chain containing the functional group in the active site is positioned similarly in both structures, it comes from residues located at different positions in the two structures. For example, the functional Arg18 of Ohr comes from the glycine motif (motif I) of Ohr, but the equivalently located Arg39 of *Ec*OsmC comes from the F*XX*R motif (motif I′) of *Ec*OsmC (Figs. 1 and 3).

The highly conserved histidine and serine side chains in motif IV are clustered near the two conserved cysteines in three of the structures, the exception being yhfA, although the functional reason for this motif is not clear (Fig. 6*b*). In the OsmC subfamily, this histidine residue comes from motif II rather than motif IV as found in other subfamilies (Figs. 1 and 3). YhfA only has a conserved serine residue (Ser109) in this region.

The environment of the cysteine located at the active site of MPN625 and yhfA is quite different from that in *Ec*OsmC and Ohr. Firstly, the neighboring residues of the active cysteine of MPN625 and yhfA (Cys52 and Cys47 based on the Ohr structure, respectively) have no arginine next to them. Secondly, two conserved cysteines of these two proteins form a disulfide bond, unlike the thiol form in *Ec*OsmC and Ohr (Figs. 1 and 4*b*). Functional studies of Ohr showed that DTT could regenerate the thiol form during the catalytic cycle (Lesniak *et al.*, 2002). The crystallization conditions of *Ec*OsmC, Ohr and MPN625 contained 5, 3 and 1 m*M* DTT, respectively (the crystallization conditions for yhfA are not available). Therefore, one of the reasons for the oxidized form of MPN625 in the crystal structure could be the result of the small amount of DTT in the crystallization conditions or because of a requirement for a different reducing agent. We have previously proposed that the function of MPN625 could be that of a structural or regulatory protein in the signal transduction pathway of the stress response, based on the absence of prominent characteristics at the active site (Choi *et al.*, 2003). However, the different environment of the active site of MPN625 does not exclude the possibility of a Prx function. The reactive thiol is, if formed, still stable in a hydrophobic environment as shown in the reduced thioredoxin structure (Dyson *et al.*, 1989; Jeng *et al.*, 1994). The oxidized forms of MPN625 and yhfA indicate that the structural differences between the reduced and oxidized form among the OsmC sequence family members are very small (Fig. 4*b*). The Prx family generally shows a large conformational change from a dimer (reduced) to a decamer (oxidized) during a catalysis cycle (Wood *et al.*, 2002). However, OsmC sequence family structures did not show this type of quaternary conformational change.

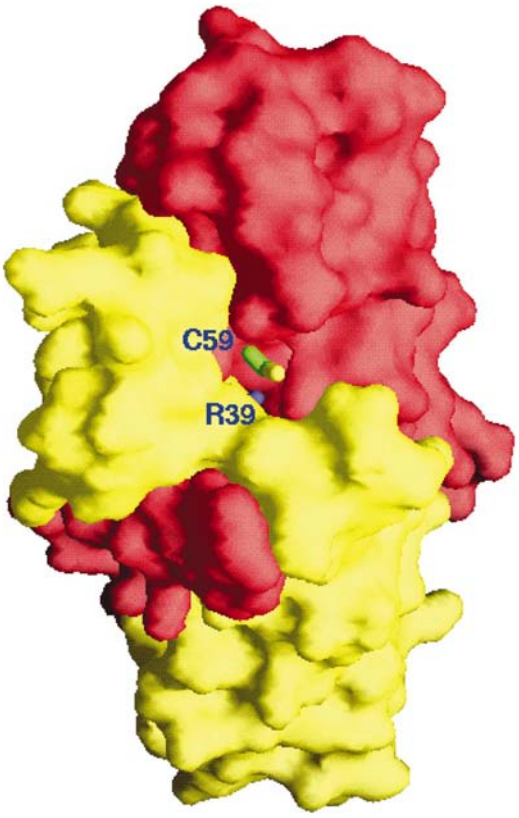

Figure 5
Molecular surface of *Ec*OsmC. One subunit is colored yellow and the other is colored scarlet. The figure shows that the dimer interface confines the active-site pocket. Yellow color represents S atoms, blue nitrogen and green carbon. The molecule was rotated to show a better view of an active site. The figure was drawn with the program *GRASP* (Nicholls *et al.*, 1991).

3.5. Comparison of active-site pocket

A detailed look at the molecular surface and charge distribution of each structure revealed differences around the entrance to the active-site pocket, although the overall structure of the three subfamilies are similar (Fig. 7). The differences are thought to be related to the substrate specificity, assuming a Prx function. Since the entrance to the active site of Ohr is mostly surrounded by hydrophobic side chains, Ohr prefers the metabolism of hydrophobic rather than inorganic hydroperoxides (Lesniak *et al.*, 2002). The *Ec*OsmC structure shows that the entrance to the active site has more aromatic residues than in the other two subfamily structures. Interestingly, an aromatic residue (phenylalanine) from the uncleaved His_6-tag is bound to the entrance of the active-site pocket in the *Ec*OsmC structure (Fig. 2*a*). This is reminiscent of the crystal structure of peroxiredoxin 5 (Prx 5), in which a benzoate molecule was bound around the active site in the crystal structure (Declercq *et al.*, 2001). Prx 5 has an aromatic residue at the entrance of the active site that conferred a preference for aromatic hydroperoxides as substrates (Declercq *et al.*, 2001). Similarly, *Ec*OsmC may have a possible preference for aromatic hydroperoxides as substrates as suggested by the crystal structure. The active sites of *Ec*OsmC and Ohr have a diameter of ∼4 Å owing to the location of aromatic amino acids around the active-site pockets: Phe36, Phe40 and Phe93 in the case of *Ec*OsmC, and Phe67 and Phe95 (each from a different subunit) in the case of Ohr. The active-site pockets of MPN625 and yhfA are apparently wider and more open. The diameters of the putative active-site pockets of MPN625 and yhfA are ∼10 Å (Fig. 7). This data, together with the different charged distributions next to the active sites (Fig. 7), would indicate that if MPN625 and yhfA have Prx activity, their substrate could be different from those that bind to *Ec*OsmC and Ohr.

3.6. The comparison between OsmC sequence family and peroxiredoxin

The structural comparison within the OsmC sequence family confirmed that members of this family may have Prx

activity and contain a unique fold and molecular characteristics that are distinct from the structure of the first reported Prx family (Pfam00578; Choi *et al.*, 1998), which is highly conserved in eukaryotes and prokaryotes. The crystal structure of hORF6 revealed an extended shape with a ten-stranded β-sheet (in dimeric state) in the center that is distinct from the barrel shape found in the OsmC sequence family structures. Unlike the OsmC sequence family, the two cysteines of the 2-Cys Prx family come from different subunits, which results in the large quaternary structural changes that occur during a catalysis cycle. The molecular surface area of ~1700 Å^2 buried by the dimer in hORF6 (Choi *et al.*, 1998) is smaller than that of the averaged value of the OsmC sequence family (~1910 Å^2). In both Prx families, the interface of the dimer-contact region confines the active-site pocket (Fig. 5).

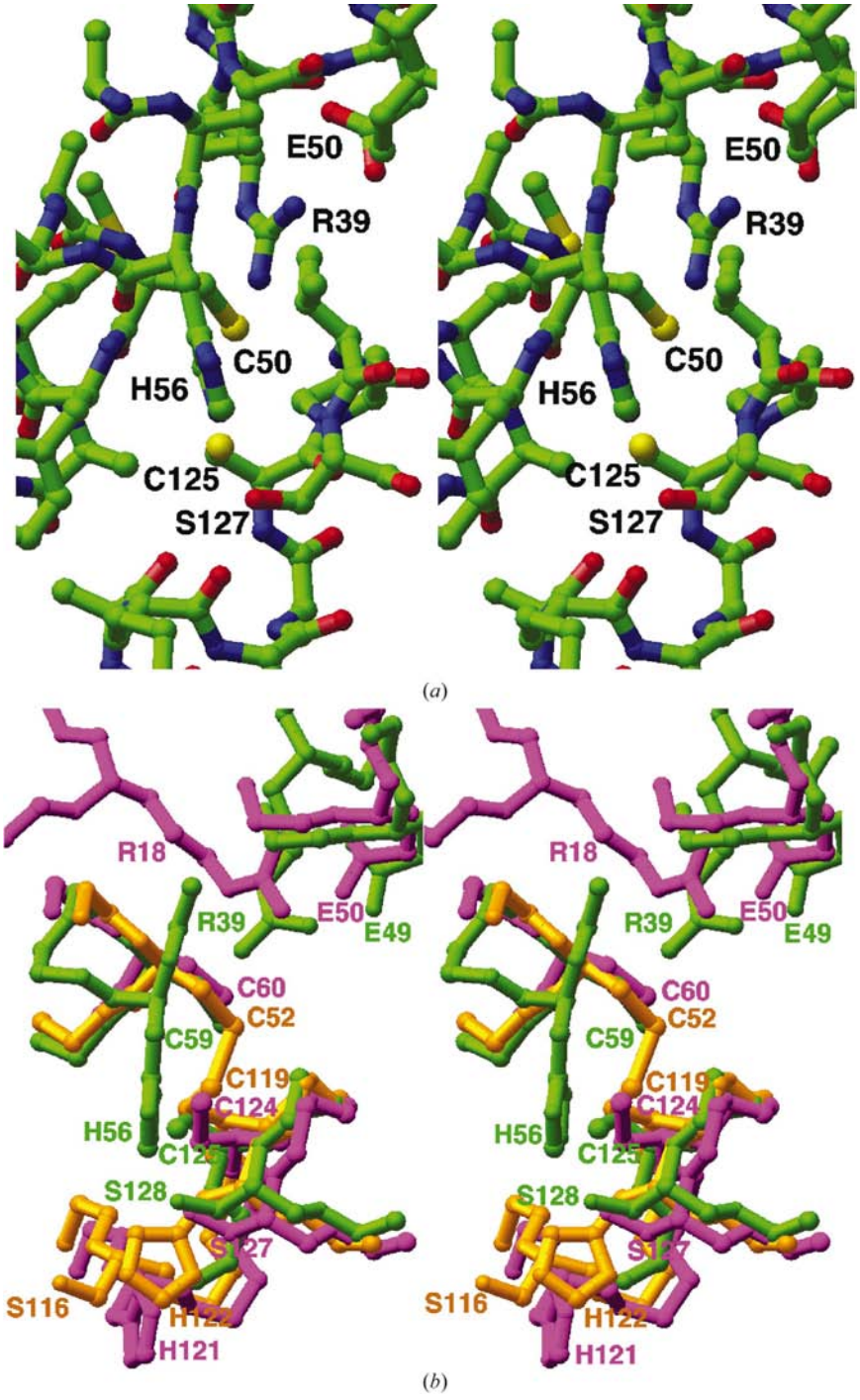

Figure 6
(*a*) Active-site residues of *Ec*OsmC. The residues around the active sites are labeled and represented by a ball-and-stick model. The view is from opposite the entrance to the active site. (*b*) Comparison of active-site residues. Green color represents *Ec*OsmC, pink Ohr and yellow MPN625. The average distances between the S^{δ} atoms of the two active-site cysteines of *Ec*OsmC, Ohr, MPN625 and yhfA (not shown in this figure) are 3.78, 3.67, 2.07 and 2.05 Å, respectively.

3.7. Possible molecular function of OsmC

Since OsmC is induced by salt stress, its molecular function as a Prx is not clear at first glance. However, OsmC-induced cells shows high viability during oxidative stress: an OsmC mutant exhibited a higher sensitivity to *t*-butanol in the exponential growth phase and to H_2O_2 and *t*-butanol in the stationary phase (Conter *et al.*, 2001). The relationship between osmotic pressure and the generation of reactive oxygen species (ROS) is not known in *E. coli*. However, in plants, during water stress brought about by salt stress, reduction of chloroplast stromal volume and generation of ROS result in the inhibition of photosynthesis (Price & Hendry, 1991). Therefore, the generation of ROS by osmotic pressure is a hazardous situation in plants (Price & Hendry, 1991). A similar phenomenon was reported during osmotic shock of *E. coli* cells in NaCl or sucrose medium, resulting in a large decrease in the cytoplasmic volume and in the inhibition of growth, of the electron-transfer chain and of four different types of sugar-transport systems (Houssin *et al.*, 1991). Alhough direct evidence has not yet been reported for the production of ROS in these situations, the malfunction of cellular machinery during osmotic pressure might be the source of production of ROS in *E. coli* as well as in plants. Therefore, one of the molecular functions of OsmC may be Prx activity working as a scavenger for specific ROS.

4. Conclusions

The structure of *Ec*OsmC provides a structural view of an OsmC subfamily I member. Our study has confirmed that the OsmC sequence family may have a Prx function. The MPN625 structure in an oxidized state clearly showed that the conformational

change during a catalysis cycle might be very small in the OsmC sequence family compared with the drastic quaternary structural change shown in the first Prx family (Wood *et al.*, 2002; Choi *et al.*, 2003). Even though the substrates of the proteins are not known, the different entrance surface features in the subfamilies suggest different substrate specificities. The genome sequences of *Mycobacterium genitalium* and *M. pneumoniae* revealed a lack of known genes involved in peroxide metabolism (Mongkolsuk *et al.*, 1998) except for the OsmC sequence family. Since overcoming oxidative stress induced by the host is essential for survival against the host defense system, the OsmC sequence family may be a good target for a new class of antibacterial drugs.

During the preparation of this paper, another manuscript appeared presenting the same structure with experimentally proven Prx function (PDB code 1qwi; Lesniak *et al.*, 2003). The r.m.s. deviations of 278 C^{α} atoms between the structural models (dimers) of the two *Ec*OsmCs are in the range 0.4–0.8 Å and reveal no significant difference, although our structural models contain an uncleaved His_6-tag model.

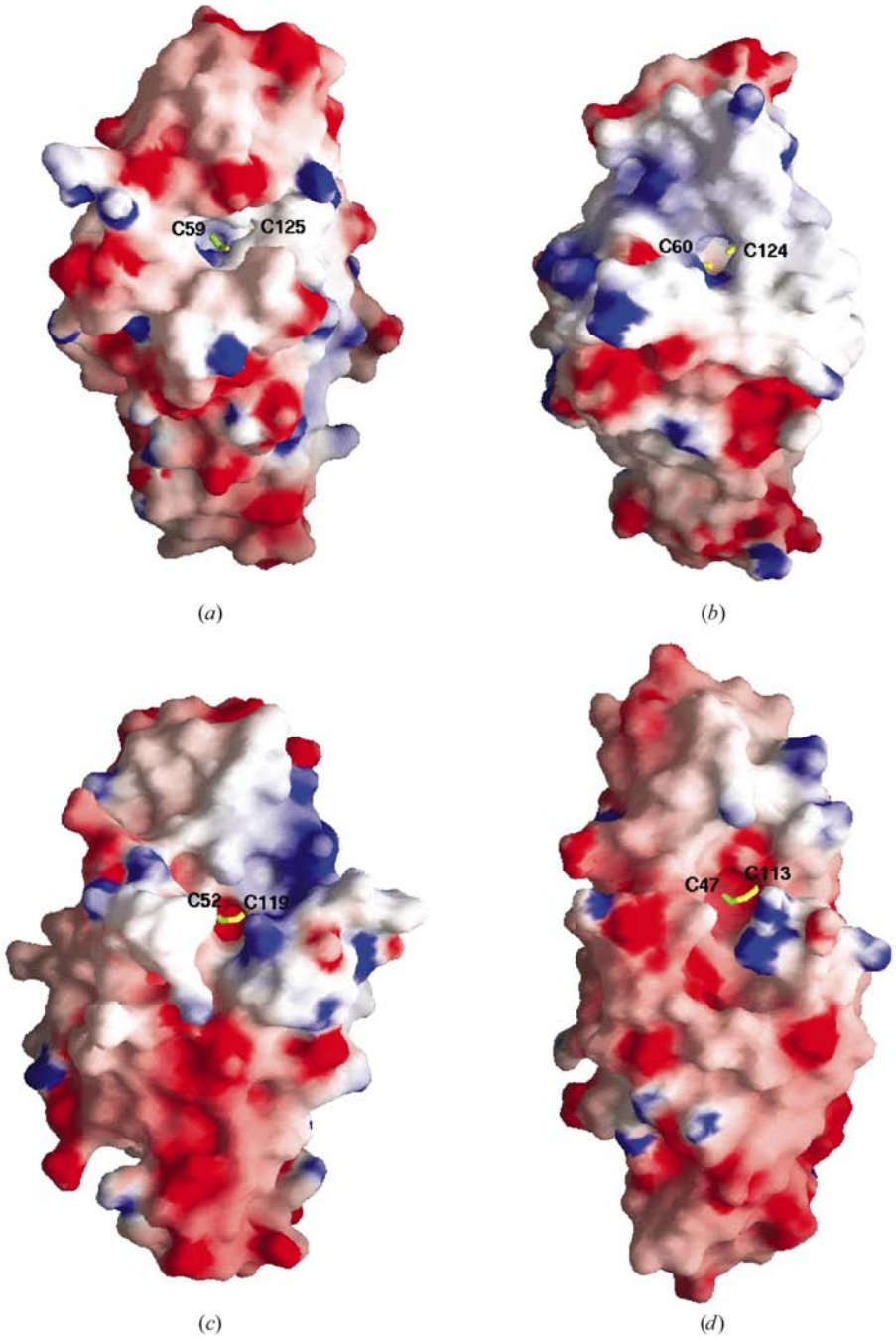

Figure 7
Electrostatic surface potential of *Ec*OsmC, Ohr, MPN625 and yhfA. Electrostatic surface potentials of *Ec*OsmC (*a*), Ohr (*b*), MPN625 (*c*) and yhfA (*d*) were drawn with the program *GRASP* (red, negative; blue, positive; white, uncharged; Nicholls *et al.*, 1991). The conserved cysteine residues are labeled and represented by a ball-and-stick model. Yellow color represents S atoms and green represents C atoms. Each molecule was rotated to show a better view of the active site.

We thank Dr Keith Henderson (Advanced Light Source, Lawrence Berkeley National Laboratory) for assistance during data collection. We are also grateful to Dr Igor Grigoriev for target selection, Barbara Gold for cloning and Bruno A. Martinez and Marlene Henriquez for protein expression and cell growth. The work described here was supported by the National Institutes of Health GM 62412.

References

Atichartpongkul, S., Loprasert, S., Vattanaviboon, P., Whangsuk, W., Helmann, J. D. & Mongkolsuk, S. (2001). *Microbiology*, **147**, 1775–1782.
Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Howe, K. L. & Sonnhammer, E. L. (2000). *Nucleic Acids Res.* **30**, 276–280.
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.
Carson, M. (1991). *J. Appl. Cryst.* **24**, 958–961.
Choi, H. J., Kang, S. W., Yang, C. H., Rhee, S. G. & Ryu, S. E. (1998). *Nature Struct. Biol.* **5**, 400–406.
Choi, I. G., Shin, D. H., Brandsen, J., Jancarik, J., Busso, D., Yokota, H., Kim, R. & Kim, S.-H. (2003). *J. Struct. Funct. Genomics*, **4**, 31–34.
Conter, A., Gangneux, C., Suzanne, M. & Gutierrez, C. (2001). *Res. Microbiol.* **152**, 17–26.
Declercq, J. P., Evrard, C., Clippe, A., Stricht, D. V., Bernard, A. & Knoops, B. (2001). *J. Mol. Biol.* **311**, 751–759.
Dodson, E. J., Winn, M. D. & Ralph, A. C. (1997). *Methods Enzymol.* **277**, 620–633.
Dyson, H. J., Holmgren, A. & Wright, P. E. (1989). *Biochemistry*, **28**, 7074–7087.
Hengge-Aronis, R. (1996). *Mol. Microbiol.* **21**, 887–893.

Houssin, C., Eynard, N., Shechter, E. & Ghazi, A. (1991). *Biochim. Biophys. Acta*, **1056**, 76–84.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jeng, M. F., Campbell, A. P., Begley, T., Holmgren, A., Case, D. A., Wright, P. E. & Dyson, H. J. (1994). *Structure*, **2**, 853–868.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Jung, J. U., Gutierrez, C., Martin, F., Ardourel, M. & Villarejo, M. (1990). *J. Biol. Chem.* **265**, 10574–10581.
Kaasen, I., Falkenberg, P., Styrvold, O. B. & Strom, A. R. (1992). *J. Bacteriol.* **174**, 889–898.
Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S.-H. (1998). *Biotechnol. Lett.* **20**, 207–210.
Kolter, R., Siegele, D. A. & Tormo, A. (1993). *Annu. Rev. Microbiol.* **28**, 855–874.
Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
Lange, R. & Hengge-Aronis, R. (1991). *Mol. Microbiol.* **28**, 49–59.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lesniak, J., Barton, W. A. & Nikolov, D. B. (2002). *EMBO J.* **21**, 6649–6659.
Lesniak, J., Barton, W. A. & Nikolov, D. B. (2003). *Protein Sci.* **12**, 2838–2843.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Mongkolsuk, S., Praituan, W., Loprasert, S., Fuangthong, M. & Chamnongpol, S. (1998). *J. Bacteriol.* **180**, 2636–2643.
Nicholls, A., Sharp, K. A. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Orengo, C. A., Michie, A. D., Jones, S., Jones, D. T., Swindells, M. B. & Thornton, J. M. (1997). *Structure*, **5**, 1093–1108.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Price, A. H. & Hendry, G. A. F. (1991). *Plant Cell Environ.* **14**, 477–484.
Shindyalov, I. N. & Bourne, P. E. (1998). *Protein Eng.* **11**, 739–747.
Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.
Völker, U., Andersen, K. K., Antelmann, H., Devine, K. M. & Hecker, M. (1998). *J. Bacteriol.* **180**, 4212–4218.
Wood, Z. A., Poole, L. B., Hantgan, R. R. & Karplus, P. A. (2002). *Biochemistry*, **41**, 5493–5504.
Yim, H. H. & Villarejo, M. (1992). *J. Bacteriol.* **174**, 3637–3644.