Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of the reduced disulfide-bond isomerase DsbC from *Escherichia coli*

Katarzyna Banaszak,[a]‡ Ingrid Mechin,[b]‡§ Graham Frost[b]¶ and Wojciech Rypniewski[a]*

[a]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12/14, 61-704 Poznan, Poland, and [b]European Molecular Biology Laboratory, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany

‡ These authors contributed equally to this work.
§ Present address: Structural Biology and Molecular Modelling Department, Aventis Pharma SA, 13 Quai Jules Guesdes, BP 14, F-94403 Vitry/Seine CEDEX, France.
¶ Present address: Genome Damage and Stability Centre, University of Sussex, Falmer, Brighton BN1 9RQ, England.

Correspondence e-mail: wojtekr@ibch.poznan.pl

Received 29 June 2004
Accepted 27 July 2004

PDB Reference: DsbC, 1tjd, r1tjdsf.

Disufide-bond isomerase (DsbC) plays a crucial role in folding periplasmically excreted bacterial proteins. The crystal structure of the reduced form of DsbC is presented. The pair of thiol groups from Cys98 and Cys101 that form the reversible disulfide bond in the enzymatic active site are 3.1 Å apart and the electron density clearly shows that the S atoms do not form a covalent bond. The other pair of Cys residues (141 and 163) in DsbC form a disulfide bond. This is different from the previously reported crystal form of DsbC (McCarthy *et al.*, 2000), in which both Cys pairs are oxidized. Specific hydrogen-bond interactions are identified that stabilize the active site in the reactive reduced state with the special participation of hydrogen bonds between the active-site cysteine residues (98 and 101) and threonine residues 94 and 182. The present structure also differs in the orientation of the catalytic domains within the protein dimer. This is evidence of flexibility within the protein that probably plays a role in accommodating the substrates in the cleft between the catalytic domains.

1. Introduction

Disulfide bonds are found in many secreted proteins. They are often required in the folding of proteins into their biologically active conformation and in maintaining their stability. Protein disulfide formation can occur *in vitro* but it is often slow and inefficient. Reactions with molecular oxygen or small disulfide molecules, such as oxidized glutathione, are required for the chemical oxidation of thiol groups to disulfide bonds. Spontaneously formed disulfide bonds often result in protein misfolding. Not surprisingly, we find that disulfide-bond formation is assisted enzymatically *in vivo* (Missiakas & Raina, 1997). Living cells have several classes of proteins that assist in folding.

In eukaryotes, secretory proteins are translocated co-translationally into the lumen of the endoplasmic reticulum, where they rapidly fold and acquire their disulfide bonds before proceeding further along the secretory pathway. The endoplasmic reticulum environment is sufficiently oxidizing to thermodynamically permit the formation of disulfide bonds that are stabilized by the protein conformation, but the rapid rate at which disulfide formation occurs is a consequence of the presence of large amounts of protein disulfide isomerase (PDI). PDI is a 57 kDa protein consisting of five domains, two of which are homologous in sequence to thioredoxin. One of them has been shown by NMR to also be similar in structure to thioredoxin (Kemmiak *et al.*, 1996). Each of the two homologous domains contains the active-site sequence -Cys-

Gly-His-Cys-, similar to the active-site sequence of thioredoxin (-Cys-Gly-Pro-Cys-).

In Gram-negative bacteria such as *Escherichia coli*, extracytoplasmic proteins are translocated into the periplasmic space, where disulfide bonds form rapidly. The isolation of mutants deficient in disulfide-bond formation has led to the characterization of several proteins involved in the process (Missiakas & Raina, 1997). The first to be recognized was DsbA, which is a soluble monomeric periplasmic protein of 21 kDa (Martin *et al.*, 1993). Although it has no apparent sequence homology to thioredoxin, its crystal structure revealed a domain with a very similar fold and an additional helical domain (Martin, 1995). It is therefore a member of the thiol–disulfide oxidoreductase family, which includes thioredoxin, PDI and glutaredoxin. DsbA has a pair of cysteine residues in its active site in the sequence -Cys-Pro-His-Cys- and can reversibly form a disulfide bond. It is believed to be the direct oxidant that 'transfers' its disulfide bond to newly secreted proteins, but it is a poor disulfide isomerase. To fulfil the role of oxidant *in vivo* DsbA has to be recycled, being reoxidized once it has accomplished disulfide-bond formation in the substrate protein. DsbB, a membrane-embedded protein, is believed to recycle the DsbA. How DsbB is itself oxidized and what the ultimate source of oxidizing power is are as yet unknown.

A third protein involved in disulfide-bond formation is DsbC. It is a soluble protein homodimer of 2×23 kDa. Each subunit contains a pair of cysteine residues in the sequence -Cys-Gly-Tyr-Cys- at positions 98–101 and cysteine residues at positions 141 and 163. Its function is primarily to catalyse intramolecular disulfide-bond rearrangements. The crystal structure of DsbC contains a thioredoxin motif, similar to that found in DsbA, but the DsbC subunit is twice the size of DsbA and contains a substantial amount of other structure responsible for dimerization and inter-domain flexibility (McCarthy *et al.*, 2000). DsbC is reduced *in vivo* by DsbD.

The common features of these proteins are two cysteine residues that are separated by only two residues and are involved in the thiol–disulfide transformations responsible for their respective activities. A disulfide bond can be formed reversibly between the S atoms of these residues, but only the more N-terminal of the two cysteine residues is exposed and reacts with other thiol-containing molecules. Apart from these common features, the active-site thiol–disulfide chemistry varies greatly between the different members of the thioredoxin family. Understanding how active sites with widely different properties are accommodated into the same overall thioredoxin fold requires in-depth studies of these proteins. Here, we present the structure of the reduced form of DsbC.

2. Materials and methods

DsbC was isolated from an overproducing strain of *E. coli* BL21 (DE3) carrying plasmid pDM801. Overproducing cells were grown overnight to saturation in LB medium and further incubated for 5 h with 0.2 m*M* IPTG in order to induce expression. Harvested cells were subjected to cold osmotic shock in order to release the periplasmic proteins. The released protein was purified by anion-exchange chromatography using Q-Sepharose with a linear gradient of 0–0.4 *M* KCl in 20 m*M* Tris–HCl pH 6.4.

Table 1
Summary of X-ray data collection and refinement.

Values in parentheses are for the highest resolution bin (2.54–2.50 Å).

Beamline	X11
Temperature (K)	100
Wavelength (Å)	0.906
Space group	$C222_1$
Unit-cell parameters (Å)	
a	42.0
b	146.1
c	73.7
$R_{\rm merge}$†	0.080
$R_{\rm work}$	0.203
$R_{\rm free}$	0.267
Mosaicity (°)	0.4
Resolution range (Å)	20–2.50
No. images	248
Oscillation range (°)	0.5
No. raw measurements	93576
No. unique data	8183
Completeness (%)	99.7 (100)
$I/\sigma(I)$	14 (3)
Reflections $> 3\sigma$ (%)	80 (44)
R.m.s.d. bonds (Å)	0.011
R.m.s.d. angles (°)	1.269
Average standard uncertainty in atomic positions (maximum likelihood-based) (Å)	0.19

† $R_{\rm merge} = \sum |I - \langle I \rangle| / \sum I$, where $\langle I \rangle$ is the average intensity for a given measurement and the summation is over all measurements.

Crystals of DsbC were grown by the hanging-drop vapour-diffusion method. The reservoir contained an aqueous solution of 0.2 *M* $LiSO_4$, 0.1 *M* Tris pH 8.4 and 20%(*w*/*v*) PEG 4000. Crystallization drops contained 1.5 µl of DsbC at 10 mg ml^{-1} and 1.5 µl reservoir solution. Crystals grew as thin plates. X-ray diffraction data were collected on the EMBL beamline X11 at the DORIS storage ring, DESY, Hamburg (Table 1). The data were processed and scaled with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

The structure of the reduced form was determined using molecular replacement with the program *AMoRe* (Navaza, 1994) using as the starting model the catalytic domain of the structure of oxidized DsbC (McCarthy *et al.*, 2000; PDB code 1eej). The atomic model was refined with the program *REFMAC*5 (Murshudov *et al.*, 1997) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). The stereochemical parameters of the model have been strongly restrained in order to maintain an acceptable difference between R and $R_{\rm free}$. The model was adjusted and electron-density maps were inspected using the program *TURBO-FRODO* (Roussel & Cambillau, 1991) running on a Red Hat Linux 9 PC. The model was validated using the program *PROCHECK* (Laskowski *et al.*, 1993). Atomic models were superimposed using the program *LSQKAB* (Collaborative Computational Project, Number 4, 1994).

3. Results

3.1. Structural overview

The asymmetric unit contains one monomer. The biological dimer consists of two monomers related by a crystallographic twofold axis. The final model of the DsbC structure consists of 215 amino-acid residues of the monomer and 71 water molecules. The electron density for the structure is well defined, with the exception of residues 1, 11 and 58. In addition, the following side chains are not visible in the electron-density map: 1, 6–16, 28, 44 and 58 in the N-terminal domain and 87, 126, 143 and 216 in the C-terminal domain. The Ramachandran plot generated by *PROCHECK* has 92% of residues in the most favoured regions and 8% of residues in the additionally allowed regions. The DsbC dimer is V-shaped, where each arm of the V is a monomer (Fig. 1*a*). Each monomer consists of two domains, a smaller N-terminal domain and a larger C-terminal domain, connected by a flexible linker region.

3.2. Secondary structure

3.2.1. N-terminal dimerization domain. The dimerization domain consists of residues 1–61. Residues 4–10 form helix α_1, followed by six β-strands that are responsible for the dimerization, especially strands β_4 and β_5 (Fig. 1*b*). The β_4-strand of one monomer interacts in an antiparallel manner *via* hydrogen bonds with the β_5-strand of another monomer, holding symmetrical monomers together. Hence, there are two antiparallel β-sheets in the small domains of the dimer, each consisting of strands β_1, β_2, β_3 and β_4 of one monomer and β'_5 and β'_6 of the symmetrical monomer. Strands β_4 and β_5 are almost perpendicular to each other owing to the presence of a *cis*-proline (Pro50) between them.

3.2.2. Linker helix. The dimerization domain is followed by helix α_2, composed of residues 61–73, followed by a single turn of a 3_{10}-helix (residues 74–77) that links the dimerization domain and the catalytic domain. In the dimer the two linker helices point in opposite directions, thus keeping the catalytic domains separated by 21–24 Å. Thus, a large cleft with a hydrophobic inner surface is created.

3.2.3. Catalytic domain. The larger C-terminal domain is composed of a mixed five-stranded β-sheet (β_7–β_{11}) surrounded by five α-helices (α_3–α_7) (Fig. 1*c*). Two subdomains can be distinguished: the thioredoxin domain, which follows the general extended thioredoxin fold, and the helical subdomain composed of helices α_4 and α_5. Just before the N-terminal of the β_9-strand there is the second of the structure's *cis*-prolines (Pro183). The turn induced in the protein chain by the *cis*-proline enables the β_9 and β_{10} strands to interact with the other strands making the C-terminal domain's β-sheet.

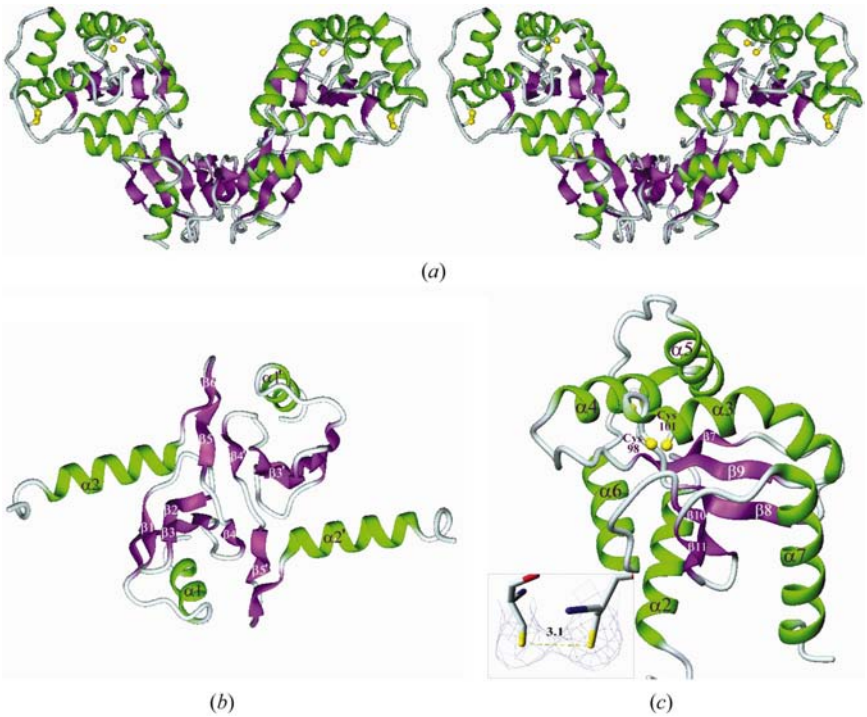

Figure 1
(*a*) The overall DsbC structure viewed normal to the twofold axis. The active-site cysteine residues, Cys98 and Cys101, are shown as a ball-and-stick model close to the cleft between the subunits. The other pair of cysteine residues, Cys141 and Cys163, are visible forming a disulfide bridge on the outer side of the molecule. (*b*) Secondary-structure elements of the dimerization domains and their interactions viewed along the twofold axis in the direction into the substrate-binding cleft. (*c*) Secondary-structure elements of the DsbC catalytic domain viewed from the inside of the substrate-binding cleft. The other pair of cysteine residues are obscured by helix α4. The inset shows a close-up view of the reduced Cys98 and Cys101 with the $(2F_o - F_c)$ electron-density map. All figures in this paper were generated using the program *DINO* (Philippsen, 2002).

3.3. DsbC active site

The active sites in DsbC are located on the inner side of the V-shaped dimer. The active site contains two key cysteine residues, 98 and 101, located at the N-terminal end of helix α_3. The cysteines are located in a small cavity and are surrounded by Tyr97, Gly99, Tyr100, Thr182 and Pro183, which create a hydrophobic environment, and the side chain of Arg125 (Fig. 2*a*). The active site and its surroundings are well defined in the electron-density map, especially the two cysteine residues. Interestingly, the active-site cysteines are in the reduced form; there is no disulfide bond between the S atoms. The distance between them is 3.1 Å (as opposed to the ideal disulfide-bond distance of 2.03 Å) and they are clearly separate in the electron density (inset in Fig. 1*c*). The Cys98 S^{γ} atom is accessible to the outside environment, while the Cys101 residue is buried inside the protein.

The conclusions of this work depend on the detailed structure of the active site, especially the distance between the S atoms of Cys98 and Cys101. The observed distance of 3.1 Å was obtained after refinement carried out without bonding restraints

between the two S atoms. The distance indicated that there was no covalent bond and that the cysteine residues were in the reduced state. The difference map, calculated with ($F_o - F_c$) coefficients, showed no significant features in the vicinity of the S atoms. In order to verify this result, 18 cycles of refinement were carried out under a disulfide-bonding restraint between the S atoms. The restraint caused the atoms to move within 2.1 Å of each other, while a negative peak appeared at the level of 5 r.m.s. in the ($F_o - F_c$) map between the two atoms. Corresponding positive peaks were observed on the opposite sides of the atomic centres. Upon releasing the bonding restraint and further refinement the S atoms moved back to their original non-bonded positions and the difference map no longer showed significant peaks in the vicinity. The difference between the observed sulfur–sulfur distance of 3.1 Å and the ideal bonding distance (2.03 Å) is also significant by the measure of the estimated standard uncertainty (e.s.u.) in atomic positions (Table 1). The e.s.u. value of 0.19 Å is the average for the whole structure. In the case of the relatively dense and well ordered S atoms the errors in atomic positions are expected to be smaller.

In the DsbC monomer there is another pair of cysteine residues (Cys141 and Cys163) located on the outer side of each arm of the V-shaped dimer in the helical subdomain. Interestingly, this pair is oxidized, forming a disulfide bond. The disulfide bond is partly exposed to the solvent.

3.4. Comparison with the oxidized DsbC structure

The oxidized form of DsbC has been reported previously (McCarthy *et al.*, 2000). The asymmetric unit of the oxidized structure contains a homodimer (chains *A* and *B*), but in the present model the monomers that make up the dimer are related by crystallographic symmetry.

Comparison of the present model and the previously determined model reveals large shifts in the overall shape of the dimer. The core of the dimer is formed by the interacting dimerization domains. Superposition of the two dimerization domains of the reduced DsbC onto the corresponding residues of the oxidized structure gives an r.m.s.d. of 0.9 Å for 60 of the 61 C^{α} atoms of chain *A* and 55 of the 61 C^{α} atoms of chain *B* (the omitted residues deviated by more than 3 r.m.s.). The catalytic domains are shifted and twisted compared with the oxidized structure about the hinge region located in helix α_2. This results in differences of up to 15 Å compared with the *A* chain of the oxidized structure and up to 9 Å compared with the *B* chain of the same structure (Figs. 3*a* and 3*b*). Helix α_2 acts as a hinge between the dimerization domain and the catalytic domain and has an effect on the relative distance between the catalytic domains in the dimer, making the cleft between them variable and allowing it to accommodate substrates of various sizes or various substrate conformers. There are no major rearrangements within the catalytic domains. Superpositions of the catalytic domains of the reduced structure and the corresponding parts of chains *A* and *B* of the oxidized structure both give an r.m.s.d. of 0.6 Å for 139 C^{α} atoms.

3.5. The reduced and oxidized active site

The major and the most interesting differences between the compared models are found in their active sites. As mentioned before, the reduced state is the active form of the enzyme because it is able to reduce existing disulfide bonds in the substrate. No reducing agents were added to the preparation during protein expression, purification and crystallization to prevent spontaneous oxidization. Therefore, local factors

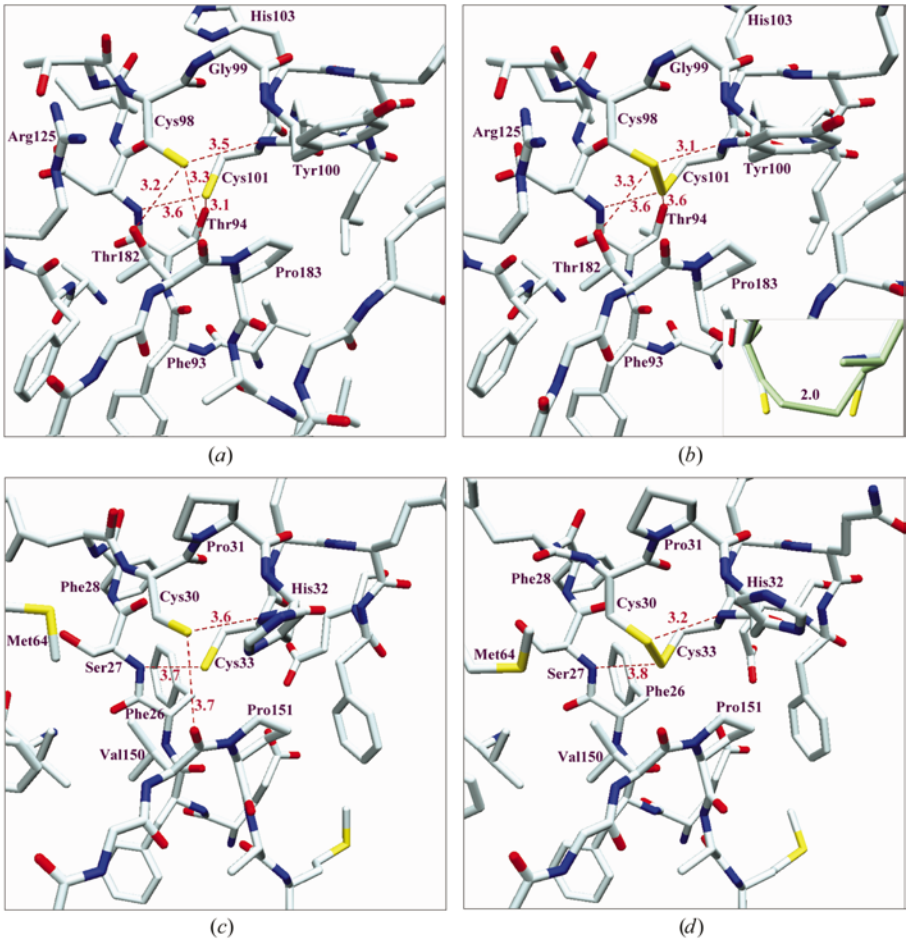

Figure 2
A view of the hydrogen-bond interactions within the active sites (shown as dashed lines). Distances for these interactions are shown in Å. (*a*) Reduced DsbC; (*b*) oxidized DsbC and a close-up view of the superposition of the reduced and oxidized (green) S atoms; (*c*) reduced DsbA; (*d*) oxidized DsbA.

within the active site should be considered responsible for the stabilization of the reactive thiol group.

A comparison between the reduced and oxidized active site is shown in Figs. 2(*a*) and 2(*b*), respectively. The largest changes are in the position of the S atoms of the key cysteine residues: each S atom in the two thiol groups is shifted by about 0.6 Å compared with the disulfide form. The hydrogen-bond network in the active site is affected by the oxidation state of the thiol groups. The oxidized Cys98 S atom forms a strong hydrogen bond[1] with the Cys101 amide (3.1 Å), while the thiol S atom in the present structure is located further away (3.5 Å) (Figs. 2*a* and 2*b*). Cys98 makes two other short contacts: as a proton donor with carbonyl O atom of Thr182 (3.3 Å) and with the hydroxyl group of Thr182 (3.2 Å), where the Cys probably acts as a proton acceptor. The reduction of the active site is also accompanied by a shift of the Cys101 S atom, which then forms a short hydrogen bond (3.2 Å) with the O atom of the Thr94 side chain (compared with 3.6 Å in the oxidized form).

3.6. Comparison with DsbA oxidase

DsbA catalyses the formation of substrate disulfide bonds. Hence, the active form of the enzyme has the two cysteines joined in a disulfide bond as opposed to the DsbC isomerase, which is active in the reduced state. The differences between DsbC and DsbA are expected to reveal the factors responsible for the stabilization of the active sites of the two enzymes. The DsbA active site reveals a hydrogen-bonding interaction between the S atom of the more N-terminal cysteine (Cys30) and the N atom of the neighbouring cysteine (Cys33) that is shorter in the oxidized form (3.2 Å; Guddat *et al.*, 1998; PDB code 1a2j) than in the reduced form (3.6 Å; Guddat *et al.*, 1998; PDB code 1a2l), similar to DsbC structures (Fig. 2). Notably, the Cys30 thiol group is able to form only one other hydrogen bond (in DsbC there are two bonds) because the DsbC threonine (Thr182) is replaced by valine (Val150) in DsbA. Similarly, the shorter of the two hydrogen bonds between the DsbC thiol group that is closer to the C-terminus and Thr94 O^{γ} does not occur in the DsbA structure because the threonine position is occupied by a phenylalanine residue (Phe26).

4. Discussion

DsbC is a member of the thioredoxin fold superfamily with the characteristic active-site motif -Cys-*X*-*X*-Cys-. The V-shaped homodimer with the N-terminal dimerization domains and the C-terminal thioredoxin domains has a biological function (Martin *et al.*, 2001). The enzyme is active in its reduced form and capable of reduction, rearrangement and reformation of substrate disulfide bonds. In order to understand the activity of the enzyme, it is essential to pinpoint the interactions responsible for the stability of the reduced state.

The DsbC active site is positioned in an extensive hydrogen-bond network. The S^{γ} atom of Cys98 forms two hydrogen bonds with Thr182 that keep it away from the other cysteine residue. Similarly, Cys101 S^{γ} is pulled away from Cys98 S^{γ} by a short hydrogen interaction with the side chain of Thr94 (compared with the oxidized structure). The hydrogen bonds presumably stabilize the thiol groups.

Even though DsbC and DsbA proteins contain characteristic active-site -Cys-*X*-*X*-Cys- sequences and belong to the thioredoxin superfamily, they play different roles in the periplasm and consequently are active in different redox forms. It is suggested that the presence of the threonine residues (Thr94 and Thr182) that form strong hydrogen bonds with the DsbC thiol groups are responsible for the reduced-state stabilization. Owing to the positions of the threonine residues they can form no other hydrogen-bonding interactions. The role of the Thr residues in stabilizing the reduced form seems probable because in the disulfide-active form of DsbA the threonine residues are replaced by Val and Phe, which are incapable of hydrogen-bonding the thiol groups. This might, in turn, favour the disulfide in the active site of DsbA.

Furthermore, the common stabilizing interaction for the cysteine thiol group in the thioredoxin-fold redox proteins is

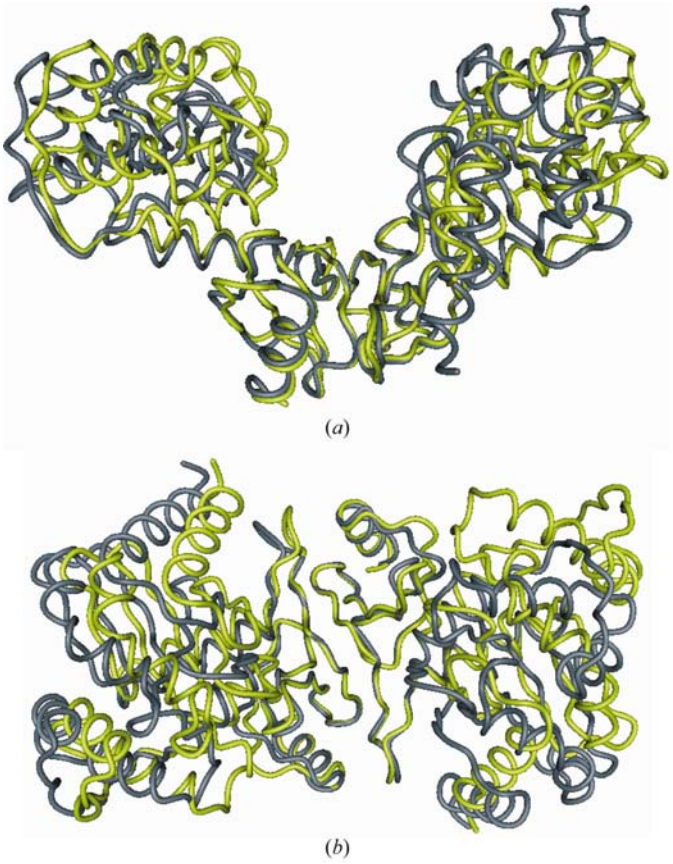

Figure 3
Two perpendicular views of the superimposition of backbone C^{α} traces of the oxidized and reduced DsbC crystal structures. The reduced DsbC structure is shown in green and the oxidized DsbC structure is shown in grey. (*a*) Viewed normal to the molecular twofold axis; (*b*) viewed down the molecular twofold axis.

[1] The hydrogen-bonding distance between a thiol group and nitrogen or oxygen is usually 3.3–3.6 Å (Gregoret *et al.*, 1991).

an interaction with the helix dipole. The partial positive charge of the helix dipole favours maintaining the active-site cysteine in the reduced form. It was shown that a cysteine at the N-terminus of any helix is highly favoured as a result of interaction with the helix dipole (Kortemme & Creighton, 1995) and leads to a decrease in the cysteine pK_a compared with the normal value.

The stabilization of the reactive thiol group at the active site is not sufficient for disulfide-bond isomerase activity (Martin *et al.*, 2001). An important role is also played by the flexible linker helices responsible for the homodimer's shape that define the broad uncharged cleft between the subunits that can accommodate a misfolded substrate protein. Thus, the cleft may loosely bind and partially shield a substrate protein to enable the correct course of reaction. In this way, DsbC combines both chaperone and disulfide-bond isomerase activity.

References

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

Gregoret, L. M., Rader, S. D., Fletterick, R. J. & Cohen, F. E. (1991). *Proteins*, **9**, 99–107.

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

Kemmiak, J., Darby, N. J., Dijkstra, K., Nilges, M. & Creighton, T. E. (1996). *Biochemistry*, **35**, 7684–7691.

Kortemme, T. & Creighton, T. E. (1995). *J. Mol. Biol.* **253**, 799–812.

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

McCarthy, A. A., Haebel, P. W., Torronen, A., Rybin, V., Baker, E. N. & Metcalf, P. (2000). *Nature Struct. Biol.* **7**, 196–199.

Martin, J. L. (1995). *Structure*, **3**, 245–250.

Martin, J. L., Bardwell, J. C. A. & Kuriyan, J. (1993). *Nature (London)*, **365**, 464–468.

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

Missiakas, D. & Raina, S. (1997). *J. Bacteriol.* **179**, 2465–2471.

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

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

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

Philippsen, A. (2002). *Visualizing Structural Biology*, http://www.dino3d.org.

Roussel, A. & Cambillau, C. (1991). *Silicon Graphics Geometry Partners Directory*, p. 86. Mountain View, CA, USA: Silicon Graphics.