PROTEINS: Structure, Function, and Bioinformatics 65:1021–1031 (2006)

Crystal Structures of *E. coli* CcmG and its Mutants Reveal Key Roles of the N-Terminal β-Sheet and the Fingerprint Region

Nan Ouyang,[1] Yong-Guang Gao,[2] Hong-Yu Hu,[2]* and Zong-Xiang Xia[1]*

[1]*State Key Laboratory of Bio-organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China*

[2]*Key Laboratory of Proteomics, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China*

ABSTRACT **CcmG, also designated DsbE, functions as a periplasmic protein thiol:disulfide oxidoreductase and is required for cytochrome *c* maturation. Here we report the crystal structures of *Escherichia coli* CcmG and its two mutants, P144A and the N-terminal fifty seven-residue deletion mutant, and two additional deletion mutants were studied by circular dichroism. Structural comparison of *E. coli* CcmG with its deletion mutants reveals that the N-terminal β-sheet is essential for maintaining the folding topology and consequently maintaining the active-site structure of CcmG. Pro144 and Glu145 are key residues of the fingerprint region of CcmG. Pro144 is in cis-configuration, and it makes van der Waals interactions with the active-site disulfide Cys80–Cys83 and forms a C—H···O hydrogen bond with Thr82, helping stabilize the active-site structure. Glu145 forms a salt-bridge and hydrogen-bond network with other residues of the fingerprint region and with Arg158, further stabilizing the active-site structure. The cis-configuration of Pro144 makes the backbone nitrogen and oxygen of Ala143 exposed to solvent, favorable for interacting with binding partners. The key role of cis-Pro144 is verified by the P144A mutant, which contains trans-Ala144 and displays redox property changes. Structural comparison of *E. coli* CcmG with the recently reported structure of CcmG in complex with the N-terminal domain of DsbD reveals that Tyr141 undergoes conformational changes upon binding DsbD. A cis-proline located at the N-terminus of the first β-strand of the ββα motif of the thioredoxin-like domain is a conserved structural feature of the thioredoxin superfamily. Proteins 2006;65:1021–1031.** **© 2006 Wiley-Liss, Inc.**

Key words: ***E. coli* cytochrome c maturation protein; thiol: disulfide oxidoreductase; mutant; crystal structure; circular dichroism; fingerprint region; cis-proline; N-terminal β-sheet; conformational change; redox potential**

INTRODUCTION

CcmG protein, encoded by one of the eight *Escherichia coli* cytochrome *c* maturation genes (*ccmA-ccmH*),[1–3] is especially required for the reducing pathway of cytochrome *c* maturation, which requires the reduction of two cysteine residues before the heme attachment.

CcmG has also been designated DsbE, a member of the disulfide bond (Dsb) protein family.[4,5] A broad variety of periplasmic protein thiol:disulfide oxidoreductases have been identified in *E. coli*, which contribute to the folding pathways of proteins and are involved in the formation and isomerization of disulfide bonds, displaying either an oxidizing, a reducing, or an isomerization activity.[3,4,6] For example, DsbD is a cytoplasmic transmembrane protein, and it catalyzes the reduction of DsbC, DsbE, and DsbG by a cascade of electron transfer reactions.[3,7] Thiol:disulfide oxidoreductases share a C–X–X–C motif as the active site, and the structures of

Abbreviations: CcmG, the protein encoded by *Escherichia coli* cytochrome c maturation gene G; CcmG-*EC*, CcmG from *Escherichia coli* (residues 19–185); CcmG-*BJ*, CcmG from *Bradyrhizobium japonicum*; CD, circular dichroism; Dsb, disulfide bond protein; GSH, the reduced glutathione; GSSG, the oxidized glutathione; *Mtb* DsbE, DsbE protein from *Mycobacterium tuberculosis*; nDsbD, the N-terminal domain of *E. coli* DsbD; nDsbD-CcmG, *E. coli* DsbD (N-terminal domain) in complex with CcmG (residues 43-185); PCR, polymerase chain reaction; rmsd, root-mean-square deviation; TRX, thioredoxin; Δ57CcmG, the *E. coli* CcmG mutant with a deletion of residues 1-57; Δ47CcmG, the *E. coli* CcmG mutant with a deletion of residues 1-47; Δ60CcmG, the *E. coli* CcmG mutant with a deletion of residues 1-60.

Grant sponsor: National Natural Science Foundation of China; Grant number: 20573131; Grant sponsor: The 863 Hi-Tech Program; Grant number: 2002BA711A13; Grant sponsor: Shanghai Commission of Science and Technology; Grant number: 03JC14081.

*Correspondence to: Zong-Xiang Xia, State Key Laboratory of Bio-organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China. E-mail: xiazx@mail.sioc.ac.cn and Hong-Yu Hu, Key Laboratory of Proteomics, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China. E-mail: hyhu@sibs.ac.cn

Received 23 February 2006; Revised 25 June 2006; Accepted 20 July 2006

Published online 3 October 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21184

most of them are characterized by the thioredoxin (TRX)-like fold[8] although they share very low primary sequence homology.[3] Some additional functions, such as a chaperone-like activity independent of the cysteine residue, were also reported for the Dsb proteins.[9–11]

E. coli CcmG is a membrane bound protein, which is anchored to cytoplasmic membrane by a highly hydrophobic N-terminus (residues 5–25) and faces periplasm with its hydrophilic C-terminal TRX-like domain. The TRX-like domain contains a C–X–X–C motif in the active site and a conserved sequence GVXGXPETF that has been identified to be the fingerprint of CcmG,[2] but the role of the fingerprint region has not been reported yet. Based on mutagenic studies it was proposed that each of the active-site cysteines of CcmG was important for attachment of heme to apocytochrome *c*.[2] However, the reaction mechanism is unknown, and the role of CcmG in cytochrome *c* maturation remains controversial.

Both the *E. coli* CcmG protein (residues 19–185) and the N-terminal fifty seven-residue deletion mutant of *E. coli* CcmG (residues 58–185) were overexpressed and purified in our laboratory,[12,13] which are referred to as CcmG-*EC* and Δ57CcmG, respectively, in this article. Here we report the crystal structures of CcmG-*EC*, Δ57CcmG, and the P144A mutant of CcmG-*EC*[†], which reveal key roles of the N-terminal β-sheet and the fingerprint region. During the preparation of this article, the crystal structure of the N-terminal domain of *E. coli* DsbD (nDsbD) in complex with CcmG (residues 43–185) was reported.[14] Here we compare the CcmG-*EC* structure with the nDsbD-CcmG complex structure and demonstrate conformational changes upon binding DsbD. A conserved structural feature of the TRX-superfamily is discussed.

MATERIALS AND METHODS

Site-Directed Mutagenesis

The P144A mutant of CcmG-*EC* was prepared by site-directed mutagenesis. Two polymerase chain reaction (PCR) products were generated using plasmid of CcmG-*EC* as the DNA template. T7 promoter (F) and T7 promoter (R) were used as primers, with oligonucleotides 5′-AAACGTTTCCGCCGCGCCATAGACACCGAGATC-3′ and 5′-TATGGCGCGGCGGAAACGTTTCTTATTGAC-3′ at the desired mutation sites, respectively. The two gene fragments were amplified by PCR using the two T7 promoters to generate a full-length gene product, incorporating the P144A mutation. This PCR product was then digested with *Bam*H I and *Nde* I and ligated into pET-3a expression vector to produce the pETP144A plasmid, which was verified by DNA sequencing and was transformed into *E. coli* cell strain BL21(DE3).

[†]Atomic coordinates have been deposited with the Protein Data Bank: 2B1K (CcmG-*EC*), 2B1L (Δ57CcmG), and 2G0F (P144A).

Protein Expression and Purification

The expression and purification of CcmG-*EC* and Δ57CcmG were previously reported.[12,13] The P144A mutant and two additional deletion mutants of *E. coli* CcmG, Δ47CcmG, and Δ60CcmG, in which residues 1–47 and 1–60 were deleted, respectively, were similarly expressed and purified.

Crystallization

Single crystals of CcmG-*EC* were grown by the hanging drop vapor diffusion method at 293 K. The protein solution contained 20 mg mL^{-1} CcmG-*EC*, 25 m*M* Tris-HCl (pH = 8.0), 250 m*M* NaCl, and 1 m*M* EDTA, and the reservoir solution consisted of 0.1*M* HEPES (pH = 7.8), 2.0*M* $(NH_4)_2SO_4$, and 2% (w/v) PEG 4K.[15]

The P144A mutant was crystallized similarly, except that the concentrations of the protein and NaCl in the droplet were 12 mg mL^{-1} and 50 m*M*, respectively, and the pH value of the reservoir solution was 8.0.

Δ57CcmG crystallized as polycrystals at the conditions similar to those used for CcmG-*EC*, but the concentration of NaCl in the protein solution and that of $(NH_4)_2SO_4$ in the reservoir solution decreased to 50 m*M* and 0.6–0.8*M*, respectively. Single crystals of Δ57CcmG were grown using the streak-seeding technique at 293 K, followed by macroseeding. Composition of the reservoir solution used for seeding was similar to that used for vapor diffusion except that the concentration of $(NH_4)_2SO_4$ decreased to 0.4–0.6*M* for streak-seeding and 0.35–0.4*M* for macroseeding.

X-Ray Data Collection

X-ray diffraction data of CcmG-*EC*, Δ57CcmG, and the P144A mutant were collected using a laboratory X-ray source equipped with a CCD detector, to 1.9, 2.6, and 2.2 Å resolution, respectively. An additional data set of Δ57CcmG was collected to 1.9 Å resolution at Beijing Synchrotron Radiation Facility. The X-ray data were processed using the software HKL2000.[16] Crystal data and the data collection statistics are shown in Table I.

Structure Determination and Refinement

The three-dimensional structure of CcmG-*EC* was determined by molecular replacement using the program CNS,[17] and the molecular structure of CcmG-*BJ* (Arg50-Leu193, corresponding to 43-180 of CcmG-*EC*)[18] was used as the search model, in which the segment Gly62-Lys79 containing a seven-residue insertion and a two-residue deletion was omitted. The structure was refined using CNS, and the model building was carried out using the program TURBO-FRODO[19] based on the (2Fo − Fc) and (Fo − Fc) electron density maps as well as the annealing "omit" maps[20] at 2.5 Å resolution. The missing segment Asn55-Gln67 was built during the refinement. The X-ray data were gradually extended up to 1.9 Å resolution. Solvent molecules were added to the model at late stage of the refinement and only those with *B*

TABLE I. Crystal Data and Statistics of Data Collection and Crystallographic Refinement

Data set	CcmG-EC	Δ57CcmG (Data set 2)	P144A
Space group	$P2_12_12_1$	$P2_1$	$P2_12_12_1$
Unit-cell parameter			
a (Å)	35.48	37.69	35.42
b (Å)	48.52	43.35	47.78
c (Å)	84.78	78.99	85.43
β (°)		94.94	
No. of molecules/asymmetric unit	1	2	1
Matthews constant (Å^3/Da)	1.9	2.2	1.9
Resolution (Å)	1.9	1.9	2.2
No. of unique reflection	12,145	19,846	7932
Data completeness (%)	99.7 (97.0)[a]	98.1 (85.2)[a]	99.8 (98.4)[a]
R_{merge} (%)[b]	6.2 (45.6)[a]	5.3 (23.3)[a]	10.4 (59.4)[a]
Reflections with $I > 3\sigma(I)$ (%)	72.2 (31.6)[a]	73.5 (32.8)[a]	66.4 (30.3)[a]
No. of amino acid residues	149	257	149
No. of solvent molecules	65	139	59
R-factor (%)	19.0 (26.0)[a]	19.2 (26.5)[a]	18.5 (28.1)[a]
R_{free}-factor (%)	21.6 (29.4)[a]	23.0 (30.0)[a]	23.8 (35.1)[a]
rmsd			
Bond length (Å)	0.0052	0.0054	0.0054
Bond angle (°)	1.2	1.2	1.2
Dihedrals (°)	23.2	24.0	23.1
Improper (°)	0.76	0.75	0.76
Mean temperature factors (Å^2)			
Main chain	22	24	27
Side chain	26	28	30
Solvent	33	35	36
Mean atomic coordinate error (Å)[c]	0.22	0.22	0.24

[a]The numbers in the parentheses correspond to the data in the highest resolution shell (1.97–1.90 Å for CcmG-*EC* and Δ57CcmG; 2.28–2.20 Å for P144A).
[b]R_{merge} = SUM[ABS(I-⟨I⟩)]/SUM (I).
[c]Obtained from Luzzati plot.

values less than 50 Å^2 were included in the final model. The polypeptide chain was extended towards the N- and C-termini. Finally, dual conformations were established for a few residues and the structure was further refined.

The structure of Δ57CcmG was determined by molecular replacement by applying the program AmoRe[21] of the CCP4 Suite[22] and using the refined structure of CcmG-*EC* (residues 58–184) as the search model. Crystallographic refinement was then carried out using CNS. The X-ray data were gradually extended, using the data sets 1 and 2 successively, up to 1.9 Å resolution.

The P144A mutant structure was determined by the difference Fourier method using CNS, based on the CcmG-*EC* structure in which Pro144 and Ala143 were omitted. One round of simulated annealing and temperature factor refinement led to a (Fo − Fc) difference electron density map at 2.2 Å resolution. Ala143 and trans-Ala144 were fitted to the map and refined well, whereas when cis-Ala144 was tested, it displayed strong positive and negative electron density in the (Fo − Fc) difference electron density map after further refinement.

Circular Dichroism Measurement

Each of the four proteins, CcmG-*EC*, Δ57CcmG, Δ47CcmG, and Δ60CcmG, was diluted with a buffer (20 m*M* Tris-HCl, 25 m*M* NaCl, pH 7.4) to a concentration of ~0.2 mg/mL. The far-UV circular dichroism (CD) spectra were recorded on a JASCO J-715 spectropolarimeter. The pathlength of the cuvette was 0.1 cm, and the spectra were recorded over 190–250 nm with a scan speed of 10 nm/min and a time constant of 0.25 s. The data were further processed for signal averaging, noise reduction, and baseline subtraction. The spectra were presented as mean residue molar ellipticities.

Redox Potential Determination by Fluorescence Measurement

For each of CcmG-*EC* and the P144A mutant, the standard redox potential was determined by measuring the redox state-dependent fluorescence emission from the CcmG/glutathione reaction system.[12,23] The reduced glutathione (GSH) and the oxidized glutathione (GSSG) were purchased from Roche, and the other reagents were of analytical grade. The protein concentration was determined by Bradford method,[24] and 0.5 μ*M* protein was incubated at 30° for 15 h under nitrogen atmosphere in a reaction buffer containing 100 m*M* sodium phosphate (pH 7.0), 1 m*M* EDTA, 1 m*M* GSSG, and GSH at different concentrations in the range of 0–75 m*M*. The fluorescence emission intensities F_{obs} at dif-

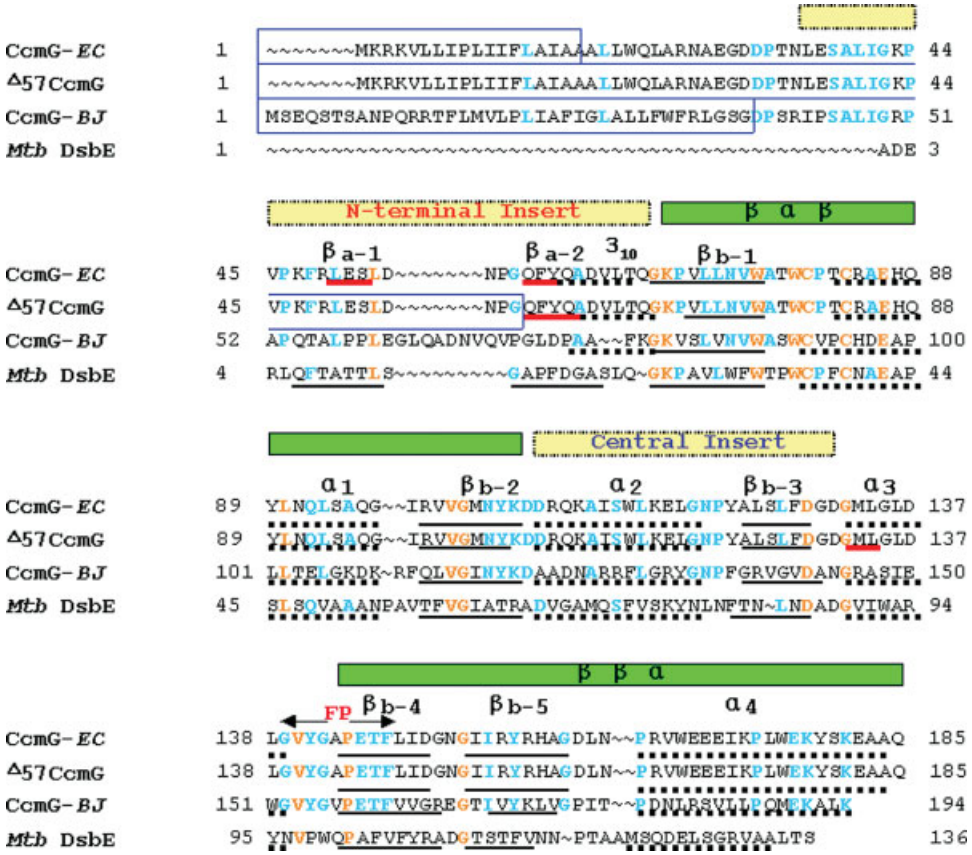

Fig. 1. Sequence alignment of CcmG-*EC* with CcmG-*BJ* and *Mtb* DsbE. Residues identical in all the three proteins are shown in orange, and residues identical in CcmG-*EC* and either CcmG-*BJ* or *Mtb* DsbE are shown in blue. The helices and β-sheets are underlined as dotted lines and solid lines, respectively, and β_a is underlined in red. The secondary structures are labeled for CcmG-*EC*. The βαβ motif and ββα motif as well as the N-terminal insert and the central insert are indicated. The fingerprint region of CcmG-*EC* is labeled as FP. The deleted N-terminal regions of CcmG-*EC*, Δ57CcmG, and CcmG-*BJ* are boxed.

ferent GSH/GSSG ratio were recorded at 350 nm on a Hitachi F-4500FK fluorescence spectrophotometer with an excitation wavelength of 280 nm. The fraction R of the reduced protein was calculated according to Eq. (1) where F_{ox} and F_{red} are the fluorescence intensities of completely oxidized and reduced proteins, respectively. The R values were plotted versus the logarithm of ($[GSH]^2/[GSSG]$) where [GSH] and [GSSG] are the concentrations of GSH and GSSG after equilibrium, respectively, and they are calculated according to Eqs. (2) and (3), where $[GSH]_0$, $[GSSG]_0$, and $[Protein]_0$ are initial concentrations of GSH, GSSG, and the protein, respectively.

$$R = (F_{obs} - F_{ox})/(F_{red} - F_{ox}) \quad (1)$$

$$[GSH] = [GSH]_0 - 2R[Protein]_0 \quad (2)$$

$$[GSSG] = [GSSG]_0 - R[Protein]_0 \quad (3)$$

The equilibrium constant K_{eq} was obtained after nonlinear regression according to Eq. (4). The standard redox potential $E_{0(protein)}$ was then calculated from the Nernst Eq. (5), where $E_{0(GSH/GSSG)} = -0.205$ V.[23]

$$R = ([GSH]^2/[GSSG])/(K_{eq} + [GSH]^2/[GSSG]) \quad (4)$$

$$E_{0(protein)} = E_{0(GSH/GSSG)} - (RT/nF) \ln K_{eq} \quad (5)$$

RESULTS

Quality of the Structures

Both structures of CcmG-*EC* and Δ57CcmG were refined at 1.9 Å resolution, and the P144A mutant structure at 2.2 Å resolution. The R-factor and R_{free} are 19.0% and 21.6% for CcmG-*EC*, 19.2% and 23.0% for Δ57CcmG, and 18.5% and 28.1% for P144A, respectively. 93.5%, 95.9%, and 89.6% of the residues are in the most favored regions, and 6.5%, 4.1%, and 10.4% of the residues in the additional allowed regions of the Ramachandran plot, for CcmG-*EC*, Δ57CcmG, and the P144A mutant, respectively, calculated using the pro-

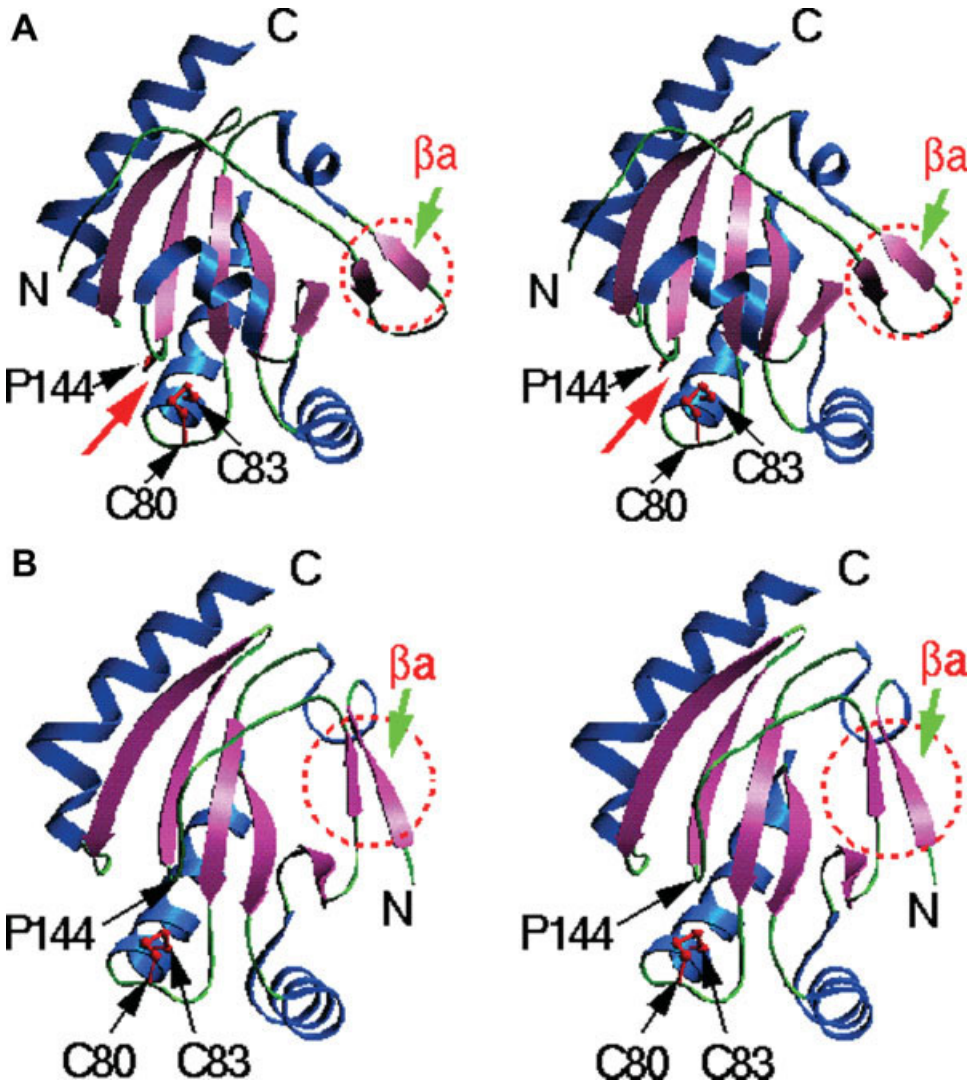

Fig. 2. Ribbon diagrams of (**A**) CcmG-*EC*. Helices and β-sheets are shown in blue and pink, respectively. The active-site cleft is indicated with a red arrow. Disulfide C80-C83 and side chain of P144 are shown in red. β_a is circled. (**B**) Δ57CcmG. Helices and β-sheets are shown in blue and pink, respectively. Disulfide C80-C83 is shown in red. The Cα position of P144 is indicated. β_a is circled. These diagrams were prepared using the program SETOR.[39]

gram PROCHECK.[25] The refinement statistics are shown in Table I.

Overall Structure of CcmG-*EC*

The amino acid sequence of CcmG-*EC* is shown in Figure 1, which is aligned with the sequences of the CcmG homologue from *Bradyrhizobium japonicum* (CcmG-*BJ*) and the DsbE protein from *Mycobacterium tuberculosis* (*Mtb* DsbE). The polypeptide chain of CcmG-*EC* consists of a TRX-like fold and two inserts, an N-terminal insert, and a central insert (Fig. 1). The side chains of five residues, Val71, Arg84, Ser114, Trp115, Trp176, exhibit dual conformations, most of which are located at the molecular surface.

In the final structure model of CcmG-*EC* each molecule contains 149 amino acid residues (Leu36-Ala184). The N-terminal segment Ala19-Asn35 is disordered, and the C-terminal Gln185 displayed poor electron density and could not be located.

The crystal structure of CcmG-*EC* shows that each molecule contains two domains, an N-terminal domain (Leu36-Gln67), and a C-terminal TRX-like domain (Gly68-Ala184). The secondary structures consist of four α-helices, one 3_{10} helix and two β-sheets, as shown in Figures 1 and 2 (a). β_a is a two-stranded antiparallel β-sheet, which, along with the 3_{10} helix, forms the main body of the N-terminal domain. β_b is a five-stranded β-sheet: the middle strand β_{b-1} starts from Gly68 and forms a big β-bulge at Lys69 and Pro70; β_{b-2}, β_{b-3} are on one side and β_{b-4}, β_{b-5} on the other side; β_{b-1}, β_{b-2},

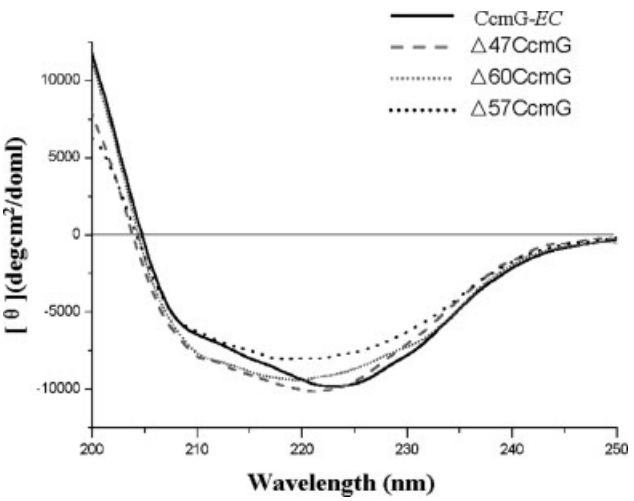

Fig. 3. Far-UV CD spectra of CcmG-*EC*, Δ57CcmG, Δ47CcmG, and Δ60CcmG.

β_{b-3} are parallel to each other and β_{b-1}, β_{b-4}, β_{b-5} are antiparallel to each other. The typical TRX-fold[8] is characterized by a βαβ motif and a ββα motif linked by a short helix to form a four-stranded β-sheet surrounded by three helices. In the CcmG-*EC* structure, the βαβ (β_{b-1}, α_1, β_{b-2}) motif and ββα (β_{b-4}, β_{b-5} and α_4) motif are linked by α_2, β_{b-3}, and α_3, forming the TRX-like domain. In this domain α_3 corresponds to the helical linker of the typical TRX-fold; α_2 and β_{b-3} constitute an additional linking region.

The overall structure of CcmG-*EC* is similar to the known structures of CcmG-*BJ*[18] and *Mtb* DsbE,[26] with the root-mean-square deviations (rmsd) of 2.2 Å and 2.1 Å for 144 and 130 Cα atoms, respectively. CcmG-*EC* structure contains a longer N-terminus than the other two structures. An acidic environment of the active site and a groove formed by two inserts of the TRX-like fold were reported to be two unique structural features of CcmG-*BJ* required for the biological activity,[18] however, neither of them is observed in the CcmG-*EC* structure. The three acidic residues Asp97, Glu158, Glu98 of CcmG-*BJ* correspond to Ala85, Glu145, and Glu86 of CcmG-*EC*. Ala85 is a hydrophobic residue. Glu86 constitutes a part of the wall of the active-site cleft, but the shortest distance from its carboxyl group to the active-site disulfide is 4.9 Å, and Glu86 is far from the binding interface (~10 Å), as shown in the nDsbD-CcmG complex structure. Glu145 and another acidic residue Asp162 point away from the disulfide, with distances of 10 and 8.5 Å, respectively, and their carboxyl groups are not involved in binding nDsbD (6.6 and 5.3 Å, respectively).

Overall Structure of Δ57CcmG

Each Δ57CcmG molecule consists of 128 amino acid residues (Gln58–Gln185). There are two β-sheets in the Δ57CcmG molecule, as shown in Figures 1 and 2(b). Like in CcmG-*EC*, β_b is a β-sheet consisting of five β-strands. However, the other β-sheet, β_a, is very different from that in CcmG-*EC*. In the CcmG-*EC* structure β_a is an anti-parallel β-sheet and both β-strands of β_a are formed by the N-terminal residues, but in the Δ57CcmG structure β_a is composed of two parallel β-strands, β_{a-1} (Gln58-Ala62) from the N-terminal region and β_{a-2} (Gly132-Leu134) from the residues between the βαβ motif and the ββα motif of the TRX-like domain. As shown in Figure 1, β_{a-1} of Δ57CcmG corresponds to β_{a-2} of CcmG-*EC* but the former is a little longer, and the segment following this β-strand is a one-turn α helix in Δ57CcmG and a 3_{10} helix in CcmG-*EC*. In addition, the β_{b-1} strand following this helix is correspondingly shorter in Δ57CcmG than in CcmG-*EC*. α_3 of CcmG-*EC* does not exist in Δ57CcmG, and the N-terminal part of it (Gly132-Leu134) turns out to be β_{a-2} in Δ57CcmG. The different folding topology gives rise to striking differences in the tertiary structure between Δ57CcmG and CcmG-*EC*, especially in the regions of Gln58-Gln67 and Leu127-Glu145.

Structure Comparison of CcmG-*EC* With its Deletion Mutants by CD Studies

Figure 3 shows the far-UV CD spectra of CcmG-*EC* and its three deletion mutants, Δ57CcmG, Δ47CcmG, and Δ60CcmG. CcmG-*EC* exhibits a negative peak at 223 nm and a shoulder at 208 nm, while Δ57CcmG gives a weaker negative peak around 218 nm, suggesting a decrease of the content of the helices in Δ57CcmG compared with that in CcmG-*EC*, which is consistent with the crystal structures of them. CcmG-*EC* and Δ47CcmG present similar far-UV CD spectra, and the spectra of Δ57CcmG and Δ60CcmG are similar to each other, indicating similar secondary structures of each pair. It implies that the deletion of residues 48–57, where β_{a-1} is located, results in remarkable changes in the secondary structures and the folding topology.

Structures of Active Site and the Fingerprint Region

Figure 2(a) shows an active site cleft of the CcmG-*EC* structure. The main chain conformation of Cys80-Pro81-Thr82-Cys83 is a β-turn, and Cys80 and Cys83 form an active-site disulfide bridge [Fig. 4(a)], indicating that CcmG-*EC* is in oxidized state. Cys80 is exposed to solvent whereas Cys83 is buried. The disulfide is located at bottom of the cleft, capped by Pro144. Pro144 is in cis-configuration [Fig. 4(a)] and is located at the N-terminus of the first β-strand of the ββα motif. The cis-peptide and the side chain of Pro144 make extensive van der Waals interactions with the active-site disulfide, as shown in Figure 4(a), and the side-chain atom CD of Pro144 makes a C—H···O hydrogen bond with the hydroxyl oxygen atom OG1 of Thr82 (3.2 Å), which is shown in Figure 4(b). These interactions help stabilize the active-site disulfide. In addition, the cis-configuration of Pro144 makes the backbone nitrogen and oxygen of

Ala143 exposed to solvent, favorable for interacting with its binding partner. This is verified by the nDsbD-CcmG complex structure in which Ala143 of CcmG forms a pair of backbone hydrogen bonds with Cys109 of nDsbD.[14]

cis-Pro144 is located in a region with conserved sequence GVXGXPETF (139–147), which has been identified to be the fingerprint for CcmG subfamily.[2,3] This region forms the N-terminal half of β_{b-4} and a loop between α_3 and β_{b-4}, the main chain of which makes two turns of $\sim 90^\circ$ at Tyr141 and cis-Pro144. Glu145 is a conserved residue in the fingerprint region and is adjacent to cis-Pro144, which forms a salt-bridge and hydrogen-bond network with the surrounding residues. As shown in Figure 4(b), the OE1 and OE2 atoms of Glu145 make salt bridge and hydrogen bond interactions with NE and NH1 of Arg158 and form hydrogen bonds with the main-chain nitrogen atoms of two fingerprint residues, Tyr141 and Gly142. Arg158 NH2 is hydrogen bonded to the backbone oxygen atoms of both Gly139 and Glu37. In addition, atoms CB, CG, and CD of Arg158 make van der Waals interactions with the phenyl ring of another conserved residue Phe147 of the fingerprint region. These interactions help maintain cis-Pro144 at an appropriate position to stabilize the active-site structure.

Tyr141 Changes Conformation Upon Binding nDsbD

In the fingerprint region residues Thr146-Phe147 are buried, and Gly139-Glu145 are exposed to solvent. Among the latter part Tyr141 makes hydrogen-bond interactions with nDsbD.[14] As shown in Figure 4(c), when CcmG binds nDsbD, the side chain of Tyr141 of CcmG undergoes conformational changes to form hydrogen bonds with the hydroxyl group of Ser9 and the backbone oxygen of Phe11 of nDsbD. To avoid close contacts with Tyr141, Arg158 of the CcmG also changes conformation so that its NH1 atom forms a hydrogen bond with the main-chain oxygen of Leu138, whereas the conformation of Glu145 does not change obviously.

The N-Terminal Fifty Seven-Residue Deletion Destroys the Active-Site Structure

Figure 4(d) shows the active-site structure of Δ57CcmG compared with that of CcmG-*EC*. The segment Cys80-Pro81-Thr82-Cys83 does not show obvious changes in its position and conformation, and Cys80-Cys83 of Δ57CcmG displays strong electron density characteristic of a disulfide bridge, indicating that it is in oxidized state. It is noteworthy that the fingerprint region is located in the segment Leu127-Glu145 that shows striking differences between Δ57CcmG and CcmG-*EC*, as described above. Pro144 remains in cis-configuration, however, its Cα atom moves by 1.6 Å, and Cβ and Cγ atoms by ~3 Å, compared with those in the CcmG-*EC* structure. The distances form cis-Pro144 to the Cys80–Cys83 disulfide in Δ57CcmG are consequently longer than that in CcmG-*EC*, generally larger than 4.4 Å except for a distance of 4.0 Å between the sulfur atom of Cys80 and the Cβ atom of Pro144, which suggests that the van der Waals interactions between the disulfide and cis-Pro144 are destroyed in this deletion mutant. The C—H···O hydrogen bond between Pro144 and Thr82 does not exist in Δ57CcmG either. Resulting from the conformational changes of the backbone of the fingerprint region, the amide nitrogen of Ala143 of Δ57CcmG is no longer solvent-accessible and the carbonyl oxygen of Ala143 moves upward by 3.5 Å, so that Ala143 would not form a pair of hydrogen bonds with nDsbD, suggested by the superposition of CcmG-*EC* with the nDsbD-CcmG complex. Compared with CcmG-*EC,* the Cα atom of Tyr141 in Δ57CcmG shows a movement of 5.5 Å and its side chain changes orientation [Fig. 4(d)], so that it would not interact with nDsbD without major changes in main-chain conformation in this region. In the Δ57CcmG structure the salt-bridge and hydrogen-bond network involving Glu145 is disrupted except that the Glu145 OE1 atom forms salt bridges with the Arg158 NE and NH2 atoms. It is suggested that the deletion of the N-terminal β-strand results in the structure changes in the fingerprint region and consequently destabilize the active-site structure.

The P144A Mutant Displays a Cis- to Trans-Configuration Change and Redox Property Changes

The overall structure of the P144A mutant is basically the same as the wild-type protein CcmG-*EC*. Compared with CcmG-*EC*, the local conformations at the active site Cys80-Pro81-Thr82-Cys83 and in the fingerprint region do not display obvious changes. In the P144A mutant structure the segment Tyr141-Glu145 shows an upward movement, and the segment Cys80-Glu86 moves accordingly, so that the relative positions between them do not change significantly. When cis-Pro144 is mutated to alanine, Ala144 adopts trans-configuration, exhibiting a cis- to trans-configuration change at the mutation site. As shown in Figure 5, the Ala144 Cβ atom is in van der Waals contact with the sulfur atom of Cys83, however, the van der Waals interactions are less extensive than those in the wild-type protein because of the configuration change of the peptide and the replacement of the five-member ring by one atom, suggesting that the active-site disulfide is less stable in P144A than in CcmG-*EC*. In the P144A mutant Cβ atom of Ala144, instead of the CD atom of Pro144 in Ccmg-*EC*, forms a C—H···O hydrogen bond with the OG1 atom of Thr82 (3.1 Å). Compared with the wild-type protein, the carbonyl group of Ala143 rotates by $\sim 60^\circ$ and becomes less solvent-accessible and unfavorable for interacting with the binding partner. The superposition of the P144A mutant with the nDsbD-CcmG complex demonstrates that the Ala143 oxygen of the mutant is 4.7 Å away from the Cys109 nitrogen of nDsbD, beyond the hydrogen bonding distance, while the hydrogen bond between the am-

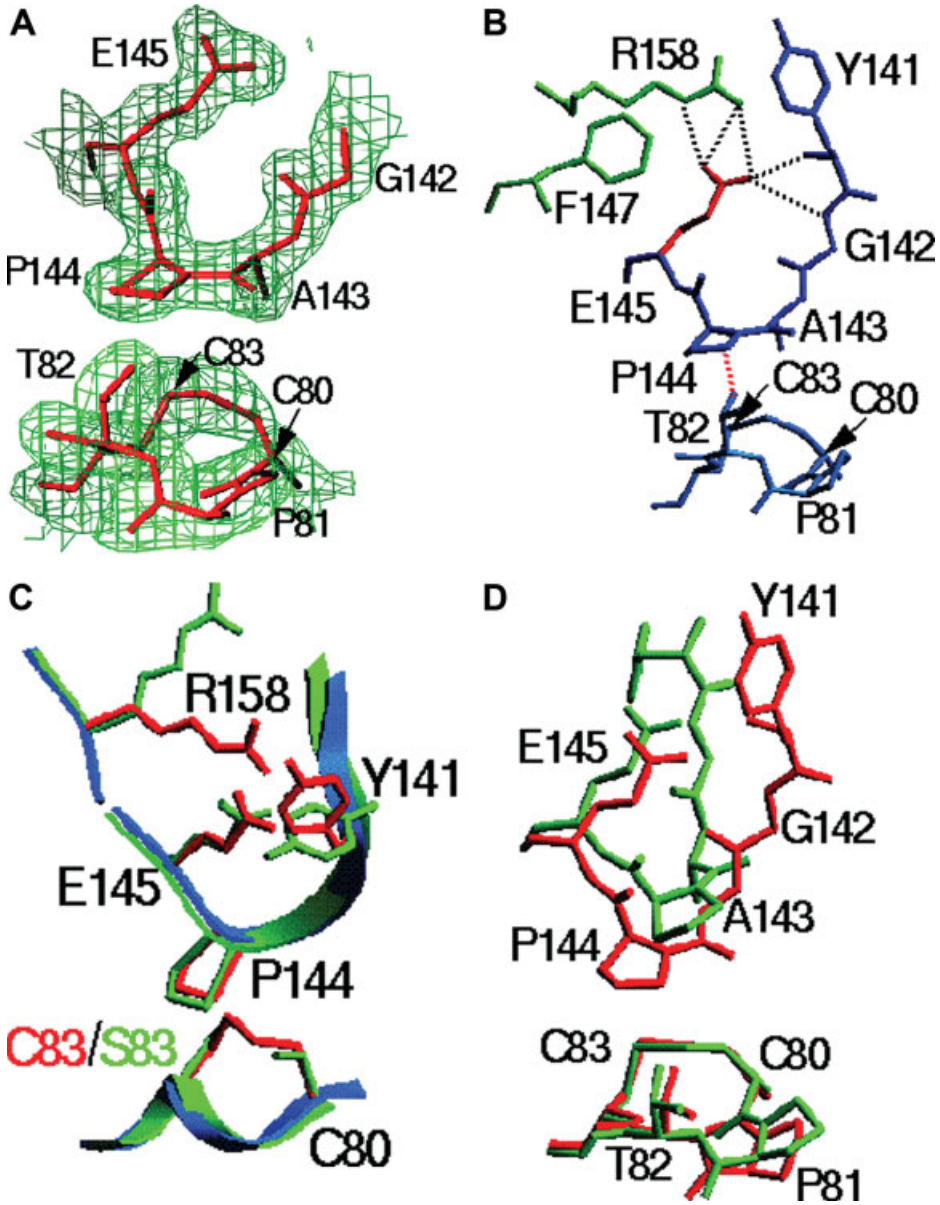

Figure 4.

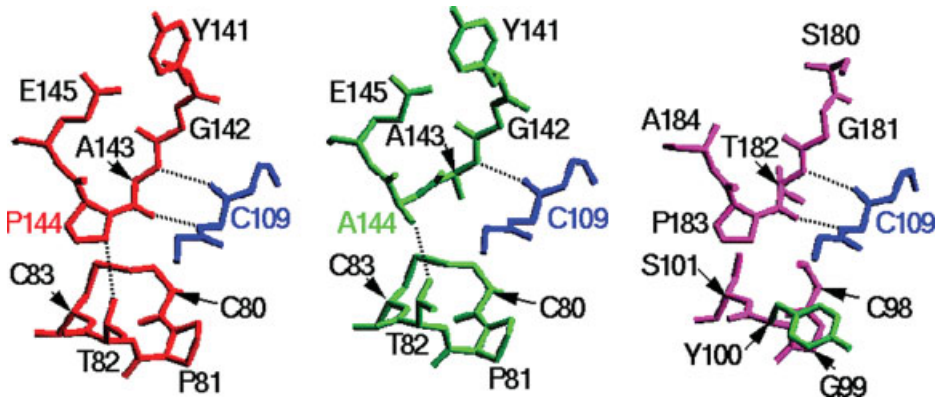

Figure 5.

ide nitrogen of Ala143 of the mutant and the carbonyl oxygen of Ser109 of nDsbD still remains (Fig. 5).

The reduced CcmG displays stronger fluorescence emission than the oxidized CcmG since the intrinsic fluorescence of tryptophan is quenched by the active-site disulfide nearby.[12,23] As shown in Figure 6, the P144A mutant is easier to be reduced by GSH than CcmG-*EC*, leading to an increase of the relative amount of the reduced state. The equilibrium constants K_{eq} for the CcmG/glutathione system are $1.14 \times 10^{-1}M$ and $1.80 \times 10^{-2}M$, and the standard redox potentials are -0.178 and -0.155 V, for the wild-type CcmG-*EC* and the P144A mutant, respectively. The redox potential change indicates a decrease of the reduction activity resulting from the mutation, which agrees well with the less stability of the active-site disulfide suggested by the weaker van der Waals interactions between the mutation site and the disulfide, as shown in the P144A mutant structure. Experimentally, the changes in the redox potential support the role of the fingerprint region in stabilizing the active-site disulfide.

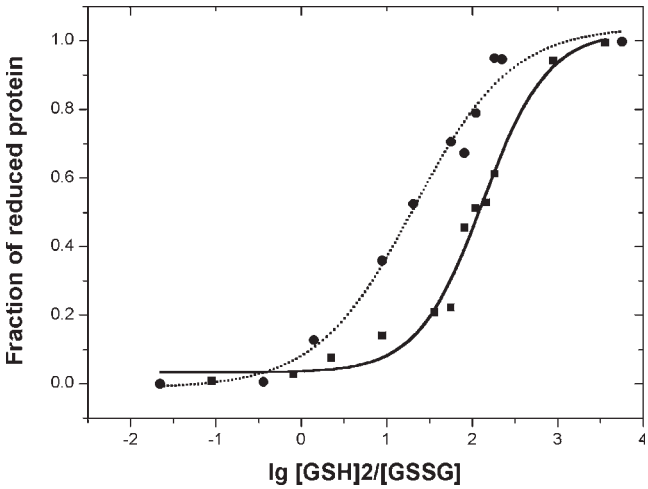

Fig. 6. Redox equilibrium of CcmG-*EC* (squares) and the P144A mutant (circles) with glutathione.

DISCUSSION

N-Terminal β-Sheet is Essential for Maintaining Folding Topology and Active-Site Structure

The crystal structure of CcmG-*EC* consists of a TRX-like fold (Gly68-Lys106 and Gly132-Gln185) as well as an N-terminal insert (Ala36-Gln67, containing a two-stranded β-sheet and a 3_{10} helix) and a central insert (Asp107-Asp131) (Fig. 1). The structure comparison of Δ57CcmG with CcmG-*EC* in combination with CD spectra of the three deletion mutants of CcmG indicates that deletion of either one or two β-strands of the N-terminal β-sheet (βa) gives rise to remarkable changes in secondary structure, which suggests that the N-terminal β-sheet is essential for maintaining the CcmG folding topology and consequently stabilizing the active-site structure and maintaining the biological activity of CcmG-*EC*. This conclusion well agrees with the recently reported experimental results: the deletion mutant Δ42CcmG (with an N-terminal deletion of 42 residues) contained the intact β_a-sheet and it stably bound nDsbD,[14] whereas the deletions of ΔLeu24-Thr66 and ΔAsp31-Gln67 made the two mutants unable to complement cytochrome *c* maturation, in both of which the whole β_a sheet was deleted.[27]

The Fingerprint Residue Cis-Pro144 and Glu145 Play Key Roles in Stabilizing the Active-Site Structure

The conserved sequence GVXGXPETF(139-147) was previously identified to be the fingerprint region of the CcmG subfamily, however, its role was unclear. The CcmG-*EC* structure revels that the fingerprint residue Pro144 is in cis-configuration, and the cis-Pro144 helps stabilize the active-site structure by making extensive van der Waals interactions with the active-site disulfide and forming a C-H···O hydrogen bond with Thr82. The cis-configuration of Pro144 also makes its adjacent residue Ala143 form a pair of main-chain hydrogen bonds with nDsbD. These structural features demonstrate that cis-Pro144 plays key roles in stabilizing the active-site structure of CcmG and favoring the interaction with the binding partner. This conclusion is verified by the structure and the redox property of the P144A mutant. The fingerprint residue Glu145 forms a salt-bridge and hydrogen-bond network, helping maintain the appropriate position of cis-Pro144 and further stabilize the active-site structure. Tyr141 is also an important fingerprint residue, which exhibits conformational change and forms interprotein hydrogen bonds when CcmG binds nDsbD.

A Conserved Structural Feature of the TRX-Superfamily

The three-dimensional structures of some other members of the Dsb protein family, such as DsbA,[28] DsbC,[29] DsbG,[30] and the C-terminal domain of DsbD[31] were

Fig. 4. Active site and fingerprint region of CcmG-*EC* and Δ57CcmG. (**A**) Electron density of segments C80-C83 and G142-E145 of CcmG-*EC*, contoured at 1.0 σ. (**B**) Hydrogen-bond network in the fingerprint region of CcmG-*EC*. Hydrogen bonds are shown as dotted lines. The C−H···O hydrogen bond is shown in red. (**C**) Y141 and R158 of CcmG-*EC*, superimposed with those in the nDsbD-CcmG complex. The main chain and side chains of CcmG-*EC* are shown in blue and red, respectively, and those of nDsbD-CcmG are shown in green. (**D**) The segments C80-C83 and Y141-E145 of Δ57CcmG, superimposed with those of CcmG-*EC*. CcmG-*EC* and Δ57CcmG are shown in red and green, respectively. The residues of CcmG-*EC* are labeled. Diagram (A) was prepared using the program FRODO-TURBO, and diagrams (B–D) were prepared using the program SETOR.

Fig. 5. The active-site disulfide and the cis-proline loop of CcmG-*EC* in comparison with those of the P144A mutant and of DsbC in the DsbC-nDsbD complex. CcmG-*EC* is shown in red in the left panel, P144A in green in the central panel, DsbC in pink in the right pannel (the side-chain of Y100 in green for clarity), and nDsbD in blue in the three panels. The C−H···O hydrogen bond and interprotein hydrogen bonds are shown as dotted lines. This diagram was prepared using the program SETOR.

reported. Each of them contains a TRX-like domain, and the superposition of the TRX-like domain of CcmG-*EC* with those of the other members of the Dsb protein family give rmsd values in the range of 2.5–2.9 Å for 92 Cα atoms. The folding topology of DsbA is very different from those of the other members of the family since a helical domain is inserted between the βαβ motif and the ββα motif of the TRX-like domain in DsbA.[28] However, around the active site C–X–X–C and the capped cis-proline the structures of CcmG-*EC* and DsbA are similar to each other. The cis-peptide and the Pro151 side chain of DsbA make van der Waals interactions with the active-site disulfide but no C—H···O hydrogen bond exists between them. The P151A mutant of DsbA[32] shows a cis- to trans-configuration change at the mutation site, similar to the change observed in the P144A mutant of CcmG. But unlike the latter, the backbone of the Gly149-Ala152 loop in the DsbA P151A mutant structure shows a global rearrangement, which results in a complete loss of the van der Waals interactions with the active site disulfide. In addition, the side chain of His32 shows a great conformational change in the P151A mutant of DsbA. It was reported that P151A mutant showed lower protein stability and enzymatic activity than the wild-type DsbA. Recently, other mutational alterations of DsbA were reported to result in accumulation of enzymatic reaction intermediates.[33]

In the nDsbD-CcmG complex structure there is a pair of backbone hydrogen bonds between Ala143 of CcmG and Cys109 of nDsbD, as described above. Ala143 is the residue preceding cis-Pro144 in the sequence. We examined the known structures of complexes of other members of the TRX-superfamily and found out that these structural features are conserved in these complexes. For example, in the DsbC-nDsbD complex structure[34] cis-Pro183 of DsbC is close to the active site Cys98, and Thr182 forms a pair of backbone hydrogen bonds with Cys109 of nDsbD. When the nDsbD molecules in the two complexes nDsbD-CcmG and DsbC-nDsbD superimpose with each other, the CcmG/DsbC backbone of the loop containing the active-site disulfide (Ala77-Ser83 of CcmG) and the middle part of the fingerprint region (Tyr141-Glu145 of CcmG) in the two structures also superimpose quite well, with rmsd of less than 1.5 Å. The rmsd values between the Cα atoms of cis-Pro144, Ala143, Cys80 of CcmG and those of the corresponding residues cis-Pro183, Thr182, Cys98 of DsbC are in the range of 0.5–0.7 Å, and those between the Cα atom of Ser83 (mutated from Cys83) of CcmG and that of Ser101 (mutated form Cys101) of DsbC is 1.2 Å. The similarity in the active-site disulfide, the cis-proline loop, and the pair of hydrogen bonds with nDsbD, between the two complexes, are shown in Figure 5.

We recently reported the crystal structure of human κ class glutathione transferase, in which cis-Pro184 is close to the active site Ser16, and Leu183 forms a pair of hydrogen bonds with the oxidized product of its substrate glutathione, and these structural features are conserved in various glutathione transferase complexes.[35]

Furthermore, when we examined the structures of other members of the Dsb protein family as well as many other proteins belonging to the TRX-superfamily, such as thioredoxin,[36] glutaredoxin,[37] and disulfide protein isomerase,[38] it was found out that each of them contains a cis-proline near the active site, and the backbone nitrogen and oxygen atoms of the residue preceding the cis-proline are exposed to solvent, favorable for forming complexes by making a pair of hydrogen bonds to the binding partners.

All of these cis-proline residues are located near the active site and at a conserved position relative to the TRX-like fold, i.e. at the N-terminus of the first β-strand of the ββα motif of the TRX-like fold, which is a conserved structural feature of the TRX-superfamily and appears essential for interactions with the binding partner.

CONCLUSIONS

The crystal structures of *E. coli* CcmG and its two mutants reveal that the N-terminal β-sheet is essential for maintaining the folding topology of CcmG and consequently maintaining the active-site structure, and the fingerprint region, especially cis-Pro144 and Glu145, plays key roles in stabilizing the active-site structure. The structural comparison of *E. coli* CcmG with the nDsbD-CcmG complex demonstrates a conformational change of Tyr141 upon binding DsbD. A cis-proline residue located at the first β-strand of the ββα motif of the TRX-like fold is found to be a conserved structural feature of the TRX-superfamily, which appears essential for interacting with the binding partners.

ACKNOWLEDGMENTS

We are grateful to Mr. Wei-Yan Chen and Dr. Jie Li for their technical assistances.

REFERENCES

1. Thony-Meyer L, Fischer F, Kunzler P, Ritz D, Hennecke H. *Escherichia coli* genes required for cytochrome *c* maturation. J Bacteriol 1995;177:4321–4326.
2. Fabianek RA, Hennecke H, Thony-Meyer L. The active-site cysteines of the periplasmic thioredoxin-like protein CcmG of *Escherichia coli* are important but not essential for cytochrome *c* maturation in vivo. J Bacteriol 1998;180:1947–1950.
3. Fabianek RA, Hennecke H, Thony-Meyer L. Periplasmic protein thiol:disulfide oxidoreductases of *Escherichia coli*. FEMS Microbiol Rev 2000;24:303–316.
4. Missiakas D, Reina S. Protein folding in the bacterial periplasm. J Bacteriol 1997;179:2465–2471.
5. Tan JT, Bardwell JCA. Key players involved in bacterial disulfide-bond formation. ChemBioChem 2004;5:1479–1487.
6. Debarbieux L, Beckwith J. Electron avenue: pathways of disulfide bond formation and isomerization. Cell 1999;99:117–119.
7. Goulding CW, Awaya MR, Parseghian A, Lim V, Eisenberg D, Missiakas D. Thiol-disulfide exchange in an immunoglobulin-like fold: structure of the N-terminal domain of DsbD. Biochemistry 2002;41:6920–6927.
8. Martin JL. Thioredoxin—a fold for all reasons. Structure 1995;3:245–250.

9. Jacon-Debuisson F, Pinkner J, Xu Z, Striker R, Padmanhaban A, Hultgren SJ. PapD chaperone function in pilus biogenesis depends on oxidant and chaperone-like activities of DsbA. Proc Natl Acad Sci USA 1994;91:11552–11556.
10. Sauvonnet N, Pugsley AP. The requirement for DsbA in pullulanase secretion is independent of disulfide bond formation in the enzyme. Mol Microbiol 1998;27:661–667.
11. Chen J, Song J, Zhang S, Wang Y, Cui D, Wang S. Chaperone activity of DsbC. J Biol Chem 1999;274:19601–19605.
12. Li Q, Hu H-Y, Wang W-Q, Xu G-J. Structural and redox properties of the leaderless DsbE (CcmG) protein: both active-site cysteines of the reduced form are involved in its function in the *Escherichia coli* periplasm. Biol Chem 2001;382:1679–1686.
13. Li Q, Hu H-Y, Xu G-J. Biochemical characterization of the thioredoxin domain of *Escherichia coli* DsbE protein reveals a weak reductant. Biochem Biophys Res Commun 2001;283:849–853.
14. Stirnimann CU, Rozhkova A, Grauschopf U, Grutter MG, Glockshuber R, Capitani G. Structure basis and kinetics of DsbD-dependent cytochrome *c* maturation. Structure 2005;13: 985–993.
15. Ouyang N, Chen W-Y, Li Q, Gao Y-G, Hu H-Y, Xia Z-X. Crystallization and preliminary crystallographic studies of *Escherichia coli* CcmG/DsbE Protein. Acta Crystallogr D Biol Crystallogr 2003;59:1674, 1675.
16. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation methods. Methods Enzymol 1997;276:307–326.
17. Brunger AT, Adams PD, Clore GM, Delano WL, Gros P, Grosse-Kunstleve RW, Jinag J-S, Kuszewski J, Nilges N, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR systems (CNS): a new software system for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998;54:905–921.
18. Edeling MA, Guddat LW, Fabianek RA, Thony-Meyer L, Martin JL. Structure of CcmG/DsbE at 1.14 Å resolution: high-fidelity reducing activity in an indiscriminately oxidizing environment. Structure 2002;10:973–979.
19. Roussel A, Cambillau C. TURBO-FRODO: silicon graphics partner geometry dictionary. Mountain View, CA: Silicon Graphics; 1991.
20. Hodel A, Kim S-H, Brunger AT. Model bias in macromolecular crystal structures. Acta Crystallogr A 1992;48:851–858.
21. Navaza J. AMoRe: an automated package for molecular replacement. Acta Crystallogr A 1994;50:157–163.
22. Collaborative computational project, Number 4. Acta Crystallogr D Biol Crystallogr 1994;50:760–763.
23. Wunderlich M, Glockshuber R. Redox properties of protein disulfide isomerase (DsbA) from *Escherichia coli.* Protein Sci 1993; 2:717–726.
24. Bradford MA. Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976;72:248–254.
25. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM. Stereochemical quality of protein structure coordinates. Proteins Struct Funct Genet 1992;12:345–364.
26. Goulding CW, Apostol MI, Gleiter S, Parseghian A, Bardwell J, Gennaro M, Eisenberg D. Gram-positive DsbE protein function differently from Gram-negative DsbE homologs. J Biol Chem 2004; 279:3516–3524.
27. Edeling MA, Ahuja U, Heras B, Thony-Meyer L, Martin JL. The acidic nature of the CcmG redox-active center is important for cytochrome *c* maturation in *Escherichia coli.* J Bacteriol 2004; 186:4030–4033.
28. Martin JL, Bardwell JC, Kuriyan J. Crystal structure of the DsbA protein required for disulphide bond formation in vivo. Nature 1993;365:464–468.
29. McCarthy AA, Haebel PW, Torronen A, Rybin V, Baker EN, Metcalf P. Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli.* Nat Struct Biol 2000;7:196–199.
30. Heras B, Edeling MA, Schirra HJ, Raina S, Martin JL. Crystal structures of the DsbG disulfide isomerase reveal an unstable disulfide. Proc Natl Acad Sci USA 2004;101:8876–8881.
31. Kim JH, Kim SJ, Jeong DG, Son JH, Ryu SE. Crystal structure of DsbDγ reveals the mechanism of redox potential shift and substrate specificity. FEBS Lett 2003;543:164–169.
32. Charbonnier J-B, Belin P, Moutiez M, Stura EA, Quemeneur E. On the role of the *cis*-proline residue in the active site of DsbA. Protein Sci 1999;8:96–105.
33. Kadokura H, Nichols L, Beckwith J. Mutational alterations of the key *cis* proline residue that cause accumulation of enzymatic reaction intermediates of DsbA, a member of the thioredoxin superfamily. J Bacteriol 2005;187:1519–1522.
34. Haebel PW, Goldstone D, Katzen F, Beckwith J, Metcalf P. The disulfide bond isomerase DsbC is activated by an immunoglobulin-fold thiol oxidoreductase: crystal structure of the DsbC-DsbDα complex. EMBO J 2002;18:4774–4784.
35. Li J, Xia Z, Ding J. Thioredoxin-like domain of human κ class glutathione transferase reveals sequence homology and structure similarity to the θ class enzyme. Protein Sci 2005;14:2361–2369.
36. Katti SK, LeMaster DM, Eklund H. Crystal strucutre of Thioredoxin from *Escherichia coli* at 1.68 Å resolution. J Mol Biol 1990; 212:167–184.
37. Sodano P, Xia TH, Bushweller JH, Bjornberg O, Holmgren A, Billeter M, Wuthrich K. Sequence-specific 1H NMR assignments and determination of the three-dimensional structure of reduced *Escherichia coli* glutaredoxin. J Mol Biol 1991;221:1311–1324.
38. Kemmink J, Darby NJ, Dijkstra K, Nilges M, Creighton TE. Structure determination of the N-terminal thioredoxin-like domain of protein disulfide isomerase using multidimensional heteronuclear $^{13}C/^{15}N$ NMR spectroscopy. Biochemistry 1996; 35:7684–7691.
39. Evans SV. SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. J Mol Graph 1993;11: 134–138.