Crystal structure of scaffolding protein CheW from thermoanaerobacter tengcongensis

Wang Yao, Liang Shi, Dong-Cai Liang *

National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101, PR China

Received 23 July 2007
Available online 31 July 2007

Abstract

The crystal structure of the scaffolding protein CheW from Thermotoga tengcongensis (TtCheW) is reported with a resolution at 2.2 Å using molecular replacement. Based on the crystal structure TmCheA P4–P5–TtCheW from Thermotoga maritime reported by others, we modeled the TmCheA P4–P5–TtCheW complex and predicted that TtCheW is involved in a hydrophobic interaction with CheA, similar to that for TmCheW. We also found that the conserved motif “NxxGxIxP” from CheW plays an important role in CheA binding. The coincidence of the reported mutation sites related to CheW–MCP binding, and the predicted protein interaction region within the TtCheW molecule, suggest that CheW–MCP binding sites lie in the groove-shaped area between TtCheW and the CheA P4 domain within the assembled model.

© 2007 Elsevier Inc. All rights reserved.

Keywords: TtCheW; CheA; MCP; Binding; Interaction; Crystal structure

Chemotaxis enables bacteria to control their movements in response to gradients of beneficial and toxic chemicals. Responding to temporal changes in chemoeffector concentrations, bacteria control their swimming behavior by switching the rotational sense of their flagellar motors and moving between smooth swimming (CCW) and tumbling swimming (CW) [1].

Two transmembrane protein complexes from bacteria exhibit remarkable sensitivity, through gain and feedback control, to the external environment and internal physiology. The first comprises an assembly of methyl-accepting chemotaxis proteins (MCPs), histidine kinase (CheA), scaffolding protein (CheW), receptor-modification enzymes, and localized phosphatases; the second is the multi-component flagellar motor that propels the cell [2–5]. MCP stimulates the autokinase activity of CheA through binding to repellent ligands; CheA is then autophosphorylated, and in turn phosphorylates the response regulator CheY. Phosphorylated CheY diffuses to the flagellar motor and binds to the switch complex, resulting in a change in the direction of cell movement [6–10].

CheW is an essential component of the first complex. It interacts with both CheA and the cytoplasmic domain of MCP and links their activities. It is required for the formation of stable ternary MCP–CheW–CheA complexes at the poles of cells [11–13]. It could also compete with CheA for MCP binding sites, thereby destroying the ternary MCP–CheA–CheW complexes. Over-expression and deletion of CheW give similar non-chemotaxis phenotypes [11]. Thus, CheW is thought to be a scaffolding protein that transduces the signals generated by MCP to CheA.

Three structures from Thermotoga maritime with 20%, 29%, and 31% sequence similarities to TtCheW have been reported recently. They are a 2.6 Å crystal structure of the TmCheA P3–P4–P5 domain [14], an NMR structure of TmCheW [15], and a 3.5 Å crystal structure of the TmCheA P4–P5–TmCheW complex [16].
Until now, a high-resolution crystal structure of free-state CheW has not been reported. As a result, previous research has failed to consider the conserved motif within CheW that is involved in its interaction with CheA. In addition, there is little information on the conformational changes of CheW after CheA binding, and the MCP binding region within CheW remains unknown.

Here, we report the crystal structure of CheW from *Thermoanaerobacter tengcongensis* at 2.2 Å resolution using molecular replacement. We identified a conserved motif, “NxxGxIxP”, from TtCheW that plays an important role in CheA binding. The binding sites between TtCheW and MCP are also predicted.

**Materials and methods**

*Protein expression and purification.* The coding sequence of the gene TTE0700 (NCBI Accession No. AE013038) was initially amplified by PCR from the *t. tengcongensis* genome [17]. The PCR product was ligated into the Ncol and HindIII restriction sites of the PET-10 vector. The insert was verified by DNA sequencing (Takara Dalian). The resulting plasmid was expressed in *Escherichia coli* strain BL21(DE3) (Novagen). Transformed cells were cultured at 37 °C in LB medium and incubated for 15 h at 20 °C with 0.5 mM IPTG on reaching OD600 = 0.6. Cells were harvested by centrifugation, and resuspended by binding buffer (50 mM Tris–HCl, 0.5 M NaCl, and 20 mM imidazole, pH 8.0). After sonication on ice, the lysate was clarified by centrifugation. The protein was purified using an Ni-affinity column and size exclusion chromatography (Amerham Pharmacia Biotech). After exchange into storage buffer (10 mM Tris–HCl, 5 mM NaCl, pH 8.0), protein was concentrated to 10 mg/ml, frozen in liquid nitrogen and stored at −80 °C until required.

*Crystallization.* X-ray data collection, structure determination and refinement. The initial crystallization conditions for TtCheW were obtained from the PEG/ION screen (Hampton Research), using the hanging-drop, vapor-diffusion method. All chemicals used for crystallization were purchased from Sigma Chemical Co. Optimized crystals were grown from drops comprising a mixture of equal volumes of protein (10 mg/ml) and reservoir solution, containing 0.1 M acetate buffer, pH 5.4, 150 mM glycerol. Diffraction data were collected at 100 K using Rigaku R-AXIS (10 mg/ml) and reservoir solution, containing 0.1 M acetate buffer, pH 5.4, and 103 water molecules were also incorporated. There are two long loops formed by residues 28–41 and 98–116 (loop 1 and loop 2), which could be traced from clear electron density maps. Loop 1 interacts with CheA directly [16], while the function of loop 2 is unknown. Data collection and refinement statistics are listed in Table 1.

The TtCheW molecule comprises two β-barrel domains, each of which forms a five β-sheet around an internal hydrophobic core (Fig. 1B). The topology map is shown in Fig. 1C. The chain starts from a β-hairpin (β1 and β2) in domain 1 and then goes into domain 2 to form three consecutive β-sheets (β3, β4, β5) and a hairpin (β6 and β7). Finally, the chain returns to domain 1 and forms the last three β-sheets (β8, β9, β10). This topology is equivalent to two consecutive β-barrels with an exchange of a β-hairpin in the middle. According to MSDfold analysis [26], the secondary structure of the TtCheW molecule belongs to the CheW family of OB-fold, while the first β-barrel belongs to the SH3 SCOP family. The characteristic of the CheW family is a tandem repeat of two swapped domains, one with a canonical OB-fold topology and the other with a circular permutation.

Hydrophobic residues 8F, 16G, 135L, and 136F, and 43G, 54V, 55I, 75I, 76I, 85G, 86L, and 88V, which play key roles in internal hydrophobic core formation, are buried in domain 1 and domain 2, respectively. The two β-barrel domains are joined by two irregular turns at residues 19–22, and 91, which are hydrophobic except residue 21. Thus, the exchange of β strands creates a hydrophobic surface.

Based on size exclusion chromatography, the functional unit of TtCheW is a monomer. However, we observed a “dimer” molecule owing to crystal packing instead of function related by twofold non-crystallographic symmetry in an asymmetric unit. The root mean square deviation (r.m.s.d.) between all Cα atoms of the two molecules is 0.92 Å. The surface area is 9266 Å², 8208 Å², and 15940 Å² for molecule A, molecule B, and the complex, respectively, and the total buried surface area is about 1533.18 Å² for the “dimer”; about 9% of the surface area of each molecule is buried. A number of surface hydrogen bonds, salt bridges, and hydrophobic interactions have been identified within the “dimer”. Residues 19I, 29H, 123D, 71D, 25V, 94V, 70K, 73R, 27K, and 22V of molecule A and 4K, 25V, 48R, 73R, 94V, 23H, 71D, 91V, 94V, and 92H of molecule B are involved in hydrogen bond formation. The hydrophobic interactions between residues

**Results and discussion**

**Overall structure**

The crystal structure of scaffolding protein CheW from *T. tengcongensis* (TtCheW) was determined with a resolution at 2.2 Å using the 3.5 Å crystal structure of CheW from *T. maritima* (PDB code: 2CH4) as a molecular replacement model. The refined model contains 300 of 330 residues with two molecules related by twofold non-crystallographic symmetry in an asymmetric unit (Fig. 1A). A1, B1–9 (including six His-tags), and 10 residues of the C-terminal of both molecules were not built since their electron densities was not clearly visible, probably because the C-terminals extended into the solvent and seemed more flexible. Three glycerols and 103 water molecules were also incorporated. There are two long loops formed by residues 28–41 and 98–116 (loop 1 and loop 2), which could be traced from clear electron density maps. Loop 1 interacts with CheA directly [16], while the function of loop 2 is unknown. Data collection and refinement statistics are listed in Table 1.

The TtCheW molecule comprises two β-barrel domains, each of which forms a five β-sheet around an internal hydrophobic core (Fig. 1B). The topology map is shown in Fig. 1C. The chain starts from a β-hairpin (β1 and β2) in domain 1 and then goes into domain 2 to form three consecutive β-sheets (β3, β4, β5) and a hairpin (β6 and β7). Finally, the chain returns to domain 1 and forms the last three β-sheets (β8, β9, β10). This topology is equivalent to two consecutive β-barrels with an exchange of a β-hairpin in the middle. According to MSDfold analysis [26], the secondary structure of the TtCheW molecule belongs to the CheW family of OB-fold, while the first β-barrel belongs to the SH3 SCOP family. The characteristic of the CheW family is a tandem repeat of two swapped domains, one with a canonical OB-fold topology and the other with a circular permutation.

Hydrophobic residues 8F, 16G, 135L, and 136F, and 43G, 54V, 55I, 75I, 76I, 85G, 86L, and 88V, which play key roles in internal hydrophobic core formation, are buried in domain 1 and domain 2, respectively. The two β-barrel domains are joined by two irregular turns at residues 19–22, and 91, which are hydrophobic except residue 21. Thus, the exchange of β strands creates a hydrophobic surface.

Based on size exclusion chromatography, the functional unit of TtCheW is a monomer. However, we observed a “dimer” molecule owing to crystal packing instead of function related by twofold non-crystallographic symmetry in an asymmetric unit. The root mean square deviation (r.m.s.d.) between all Cα atoms of the two molecules is 0.92 Å. The surface area is 9266 Å², 8208 Å², and 15940 Å² for molecule A, molecule B, and the complex, respectively, and the total buried surface area is about 1533.18 Å² for the “dimer”; about 9% of the surface area of each molecule is buried. A number of surface hydrogen bonds, salt bridges, and hydrophobic interactions have been identified within the “dimer”. Residues 19I, 29H, 123D, 71D, 25V, 94V, 70K, 73R, 27K, and 22V of molecule A and 4K, 25V, 48R, 73R, 94V, 23H, 71D, 91V, 94V, and 92H of molecule B are involved in hydrogen bond formation. The hydrophobic interactions between residues
19I, 24E, 25V, 27K, and 48R of molecule A, and 19I, 92H, 94V, 96P, and 123D of molecule B are also effective in “dimer” stabilization. The conformation of some loops (between \( b_3 \) and \( b_4 \), \( b_5 \) and \( b_6 \), \( b_9 \) and \( b_{10} \)) are quite different between the two molecules, probably due to environmental differences of crystal packing between the two molecules and the flexibility of the loops.

**Homologous structure superposition**

TtCheW and the P5 domain of CheA from *T. maritima* (TmCheA, PDB code: 1B3Q) share the highest three-dimensional structural similarity (Z-score 13.5, r.m.s.d. for all C\(_\alpha\) atoms is 2.3 Å, sequence identity 20%) based on the DALI server results [27]. However, they are still quite different in three regions (Fig. 2A). First, there is relatively long and flexible loop within loop 1 region of TtCheW, while there is a \( \beta \)-sheet insertion within the same region of the TmCheA P5 domain. Second, the conformation of loop 2 is dramatically different between them, and the position of the inserted \( 3_{10} \)-helix also differs. The third difference lies in the C-terminal. Unlike the long, extended loop within TtCheW, the loop is much shorter, and embraces a \( 3_{10} \)-helix within the TmCheA P5 domain.

Three-dimensional structure superposition between TtCheW and the free state of TmCheW (PDB code: 1K0S; Z-score 12.1, r.m.s.d. for all C\(_\alpha\) atoms is 2.9 Å, sequence identity 29%) infers that they almost coincide with each other at conserved sites. However, they also differ dramatically at loop 2 and the C-terminal region as the sequence differs (Fig. 2B). For example, there is a long loop extended to the solvent within the C-terminal (residues 140–146) of TtCheW, while within TmCheW there is an \( \alpha \)-helix instead.

We also compared TtCheW and TmCheW on binding to CheA (PDB code: 2CH4). \( \beta_9 \) (residues 116–122) comprises \( \beta \)-sheets in both TtCheW and TmCheW in the free state (PDB code: 1B3Q), but in TmCheW in the CheA-binding state the conformation of \( \beta_9 \) is markedly altered, into a loop structure (Fig. 2C). The mechanism of this remains unclear.

**TtCheW–CheA interactions**

To investigate the interaction between TtCheW and CheA, we created a model of the complex by substituting the TmCheW from previously published model TmCheA P4–P5–TmCheW (PDB code: 2CH4) for TtCheW. As
Fig. 3 shows, loop 1 of TtCheW and loop 2 of the TmCheA P5 domain interact. We then used Ligplot software [28] to explore the CheW–CheA P5 domain interaction, and concluded that TtCheW shows a similar pattern of interaction to TmCheW in the TmCheA P4–P5–TmCheW structure. These interactions involve residues 34I, 35P, 41I, 44I, 46N, 49G, 51I, and 53P (shown in Fig. 3), which are similar between the two complexes except for the C-terminal-most residues.

In both complexes, four conserved residues, 35P, 49G, 51I and 53P, of CheW participate in hydrophobic interactions with the CheA P5 domain. Conserved residue Asn46 is involved in a hydrogen interaction with Gly659 of CheA in the original complex, while it shows a hydrophobic interaction in the assembled complex. Based on these results, the conserved motif "NxxGxIxP" of CheW plays an essential role in CheW–CheA binding.

**TtCheW–MCP interactions**

A series of mutation experiments have been performed by others on CheW from *E. coli* to determine the interactions between CheW and MCP, and CheW site-mutants V22I, R48C, D89N, and V94M are predominant sites affecting CheW–MCP interactions [10]. E24K, G49S, G85S, V91I, and V94I are suppressive in MCP binding [10], and I19T also diminishes binding [29]. In the stereo view of the assembled complex model, there is a groove between TtCheW and the P4 domain of CheA (Fig. 4) and all of the mutation sites above are located in this groove (dashed ellipse indicated in Fig. 4).

To predict potential protein interaction regions within the TtCheW molecule, we used PPI-Pred [30] software. From higher to lower possibility, we obtained three potential protein binding regions—I, II, and III. Each region contains about thirty residues. Region I is consistent with the groove, while region II is consistent with TmCheW–CheA binding.

---

**Table 1**

<table>
<thead>
<tr>
<th>Data collection and refinement statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Space group</strong></td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
</tr>
<tr>
<td>β (°)</td>
</tr>
<tr>
<td><strong>Resolution (Å)</strong></td>
</tr>
<tr>
<td><strong>Rmerge (%)</strong></td>
</tr>
<tr>
<td><strong>I/σI</strong></td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
</tr>
<tr>
<td><strong>Redundancy</strong></td>
</tr>
</tbody>
</table>

**Refinement**

| Resolution (Å) | 19.8–2.20 |
| Wavelength (Å) | 1.5418 |
| No. of reflections | 14,589 |
| **Rwork/Rfree (%)** | 19.4/23.3 |
| No. of atoms | 2469 |
| Protein | 2348 |
| Glycerol | 3 |
| Water | 103 |
| **Average B-factors (Å²)** |
| Protein (A, B) | 19.63, 23.77 |
| Glycerol | 31.39 |
| Water | 25.13 |
| **r.m.s.d.** |
| Bond lengths (Å) | 0.014 |
| Bond angles (°) | 1.450 |

**Fig. 2.** Homologous structure superposition. (A) Superposition between TtCheW and the P5 domain of TmCheA (PDB code: 1B3Q). The main difference between the two lies in loop 1 and loop 2. Purple: TtCheW; blue: P5 domain of TmCheA. (B) Superposition between TtCheW and the free state of TmCheW (PDB code: 1KOS). They differ dramatically at their C-terminal regions. There is a long loop extending to solvent within the C-terminal of TtCheW, while within TmCheW there is an α-helix instead. Purple: TtCheW; cyan: TmCheW. (C) Superposition between TtCheW and TmCheW in the CheA-binding state (PDB code: 2CH4). β9 comprises β-sheets in both TtCheW and TmCheW in the free state (PDB code: 1B3Q), but in TmCheW in the CheA-binding state the conformation of β9 is markedly altered to a loop. Purple: TtCheW; yellow: TmCheW in the CheA-binding state. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)
According to the coincidence of the reported mutation sites related to CheW–MCP binding, and the most likely predicted protein interaction region, the CheW–MCP binding sites may lie in the groove-shaped area of the assembled complex, which is also the interface between the two molecules within the crystal structure of TtCheW.

As residue Gly49 lies in the corner of the loop between $\beta 4$ and $\beta 5$, it is more disordered and suitable for protein binding. According to our research on CheW–CheA interactions and mutation analysis performed by other groups [10], residue Gly49 is incorporated in both CheA and MCP binding.

In conclusion, we have determined the crystal structure of TtCheW at 2.2 Å resolution. Combined with previous research, we observed a conformational change of CheW after CheA binding. We have also indicated the conserved sites of TtCheW involved in CheA binding and their patterns of interaction. The location of the TtCheW–MCP binding area is also proposed, though this needs to be confirmed from the crystal structure of the TtCheW–MCP complex.
Acknowledgments

This work was supported by National Basic Research Program (also called the 973 Program) Nos. 2002CB713801 and 2006CB806501, National Project No. 2006CB090102, and the National Natural Science Foundation of China No. 30600102. We thank Professor Run-Sheng Chen of the Institute of Biophysics of CAS, Beijing for providing the *T. tengcongensis* strain. We thank EMBL for providing pETM-10 vector. We also thank Dr. Xiao-xue Yan for help with experimental skills.

References