Solution structure of the cryptic mannitol-specific phosphotransferase enzyme IIA CmtB from Escherichia coli

Caifang Yu a,b, You Li a,b, Bin Xia a,b,c, Changwen Jin a,b,c,*

a Beijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, China
b College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China
c College of Life Sciences, Peking University, Beijing 100871, China

Received 13 August 2007
Available online 27 August 2007

Abstract

The bacterial phosphoenolpyruvate-dependent sugar phosphotransferase system (PEP-PTS) is essential in the coupled transportation and phosphorylation of various types of carbohydrates. The CmtAB proteins of Escherichia coli are sequentially similar to the mannitol-specific phosphotransferase MtlA. The CmtB protein corresponds to the phosphotransferase enzyme IIA component. Here we report the solution structure of CmtB from E. coli at high resolution by NMR spectroscopy. The results show that CmtB adopts a globular fold consisting of a central mixed five-strand β-sheet flanked by seven helices at both sides. Structural comparison with the IIA domain of MtlA (IIA Mtl) reveals high overall similarity, while notable conformational differences at the active site are observed. The active site pocket of CmtB appears to be wider, and the hydrophobic regions around it is larger compared to IIA Mtl. Further, the essential arginine residue at the active site of IIA Mtl is substituted by a serine in CmtB. Instead, the active pocket of CmtB contains another arginine at a distinct position, suggesting different molecular mechanisms for phosphoryl transfer.

© 2007 Elsevier Inc. All rights reserved.

Keywords: CmtB; Phosphotransferase; Mannitol; NMR; Solution structure
while CmtB shares only 29% identity with the IIA domain of MtlA (IIA\textsuperscript{Mtl}). Previous studies showed that over-expressing the cmtA and cmtB genes using a heterologous promoter can complement some mannitol-negative \textit{E. coli} mutants \cite{5}. However, the native substrate and the exact physiological function of the CmtAB system \textit{in vivo} remain unclear.

Herein, we present the solution structure of \textit{E. coli} CmtB determined by NMR spectroscopy. CmtB adopts an \(\alpha/\beta\) sandwich fold consisting of five \(\beta\)-strands and seven helices. The overall structure of CmtB shows high similarity with MtlA IIA domains, while local structural differences around the active site are observable. Our study provides the structural basis for future investigations on the biological function of the CmtAB system.

Materials and methods

\textit{Cloning, expression, and purification of \textit{E. coli} CmtB.} The cmtB gene from \textit{E. coli} was amplified by PCR, with NdeI and EcoRI digestion sites introduced to N- and C-termini, respectively. PCR products were ligated into pET-28a(+) vector (Novagen). The recombinant plasmid was transformed into \textit{E. coli} strain BL21(DE3) (Invitrogen) for protein expression. Cells were allowed to grow overnight in 25 ml Luria–Bertani (LB) medium at 37 \(^\circ\)C and subsequently transferred into 1 L LB medium. When the A\textsubscript{600} reached around 0.8, cells were centrifuged and resuspended in 200 ml M9 minimal medium at 25 \(^\circ\)C with \(15\text{NH}_4\text{Cl}\) in the presence or absence of \(\text{\textsuperscript{13}C}_6\) glucose for the preparations of \(\text{\textsuperscript{15}C}\)/\(\text{\textsuperscript{15}N}\)-labeled or \(\text{\textsuperscript{15}N}\)-labeled samples, respectively \cite{6}. The protein expression was induced by isopropyl \(\text{\textsuperscript{13}C}_6\)–glucose for the preparations of \(\text{\textsuperscript{13}C}/\text{\textsuperscript{15}N}\)-labeled or \(\text{\textsuperscript{15}N}\)-labeled azide (NaN\textsubscript{3}) and 0.01% 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) were selected as the models for SANE to extend the nuclear Overhauser module of the CYANA program \cite{12}. The 20 lowest-energy structures were selected as the models for SANE to extend the nuclear Overhauser module of the CYANA program \cite{12}. The 20 lowest-energy structures were selected for further refinement by AMBER \cite{15,16}.

\textit{NMR spectroscopy.} All spectra were collected at 25 \(^\circ\)C on Bruker Avance 500-MHz (with cryoprobe) and 600-MHz NMR spectrometers at Beijing Nuclear Magnetic Resonance Center. All experiments were performed in the presence and absence of a 3-s proton pre-saturation period prior to the \(\text{\textsuperscript{15}N}\) excitation pulse.

Results and discussion

\textit{Chemical shift assignments.} \textit{Escherichia coli} CmtB contains a total of 147 amino acid residues. By using triple-resonance experiments, the backbone chemical shift assignments were obtained for all residues except seven prolines (Pro8, 55, 61, 66, 70, 91, and 96). More than 90% of chemical shift assignments for the side chain atoms of \textit{E. coli} CmtB were obtained. A table containing the chemical shift assignments has been deposited into the BioMagRes database under the Accession No. 15126.

\textit{Solution structure of CmtB.} The solution structure of CmtB from \textit{E. coli} was determined at high resolution using a total of 5984 NOE restraints, 114 dihedral angle restraints and 13 hydrogen bond restraints. A superimposition of the 20 lowest-energy structures is shown in Fig. 1A and B, respectively. The structural statistics is shown in Table 1. The coordinates of CmtB have been deposited in the Protein Data Bank (PDB) under the Accession No. 2OQ3.

\textit{Protein CmtB is a globular protein that folds into an \(\alpha/\beta\) sandwich consisting of a five-stranded \(\beta\)-sheet flanked by seven helices. The central mixed \(\beta\)-sheet is formed by five \(\beta\)-strands (\(\beta1\) (Ile12-Ile15), \(\beta2\) (Ile58-Ala60), \(\beta3\) (Val63-Ala64), \(\beta4\) (Gly78-Val88), and \(\beta5\) (Ile97-Ser104)) connected in the order of \(\beta2\)–\(\beta3\)–\(\beta5\)–\(\beta4\)–\(\beta1\). The four strands \(\beta2\)–\(\beta3\)–\(\beta5\)–\(\beta4\) are anti-parallel, while \(\beta1\) is oriented parallel to \(\beta4\). Five \(\alpha\)-helices (\(\alpha1\) (Trp21-Asp34), \(\alpha2\) (Glu40-Asn52), \(\alpha3\) (Ala108-Leu122), \(\alpha4\) (Glu125-Leu132), and \(\alpha5\) (Glu137-Ser145)) are packed around the protein core, with \(\alpha1\), \(\alpha2\) on one side of the \(\beta\)-sheet and \(\alpha3\)–\(\alpha5\) on the other. Another two short \(\alpha\) helices \(\alpha1\) and \(\alpha2\) were formed by residues Leu3-Tyr6 and Pro70-Cys72, respectively.

Conformational flexibility

The protein adopts an overall highly rigid conformation. The root mean square deviation (RMSD) for backbone atoms is 0.40 \(\pm\) 0.09 \AA for all residues and is 0.31 \(\pm\) 0.07 \AA for regular secondary structures. The structural rigidity is also reflected by the high and unified heteronuclear \(\{\text{\textsuperscript{1}H}\}\text{\textsuperscript{15}N}\)-NOE values. Apart from the N- and C-termini, the only regions that show relative internal flexibility are the two loops Asn53-Tyr57 and Tyr89-Pro96, with larger
than average RMSD and reduced heteronuclear $^1$H-$^{15}$N-NOE values (Fig. 1C).

<table>
<thead>
<tr>
<th>NMR restraints</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance restraints</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total unambiguous NOE</td>
<td>4810</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intra-residue</td>
<td>1782</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential (</td>
<td>i−j</td>
<td>= 1)</td>
<td>1160</td>
<td></td>
</tr>
<tr>
<td>Medium-range (1 &lt;</td>
<td>i−j</td>
<td>&lt; 5)</td>
<td>747</td>
<td></td>
</tr>
<tr>
<td>Long-range (</td>
<td>i−j</td>
<td>≥ 5)</td>
<td>1121</td>
<td></td>
</tr>
<tr>
<td>Total ambiguous NOE</td>
<td>1138</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrogen bonds</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total dihedral angle restraints</td>
<td>114</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varphi$</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi$</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Structural statistics

| Mean AMBER energy (kcal/mol) | $-7113.05 \pm 13.84$ |  |  |  |
| Restraint violations |  |  |  |  |
| Distance (>0.3 Å) | 1 |  |  |  |
| Dihedral angle (>5°) | 0 |  |  |  |

RMSD from mean structure (Å)

| Heavy | 0.88 ± 0.09 |  |  |  |
| Backbone | 0.40 ± 0.09 |  |  |  |
| Secondary structure heavy | 0.72 ± 0.07 |  |  |  |
| Secondary structure backbone | 0.31 ± 0.07 |  |  |  |

Ramachandran statistics (%)

| Residues in most favored regions | 87.0 |  |  |  |
| Residues in additional allowed regions | 12.4 |  |  |  |
| Residues in generously allowed regions | 0.5 |  |  |  |
| Residues in disallowed regions | 0.1 |  |  |  |

### The active site

Sequence comparison showed that residue His67 of CmtB is the putative histidine that is to be phosphorylated by HPr. His67 locates in the loop connecting the C-terminal end of strand $\beta_3$ to the $\beta_10$ helix $\eta_2$ (Pro70-Cys72). As shown in Fig. 2B, the active site pocket of CmtB is formed by helices $\alpha_2$, $\alpha_3$, the short $\beta$-hairpin formed by $\beta_2$ and $\beta_3$, and the two loop segments Asn53-Tyr57 and Met65-Arg69. Residue His67 is situated in the center of the active pocket, surrounded by a patch of hydrophobic residue, such as Pro55, Leu59, Met65, Pro70, Ile112, Ile115, as well as Tyr56 and Tyr57 with bulky aromatic side chains. Another histidine residue His111, which is also highly conserved, locates proximal to His67. A positively charged residue Arg69 at the active site pocket is both sequentially and spatially close to His67.

### Structural comparison with IIA$^{Mtl}$

The crystal structure of the mannitol-specific phosphotransferase Enzyme IIA domain from *E. coli* (IIA$^{Mtl}$) has been reported [20]. Structural comparison between *E. coli* CmtB and IIA$^{Mtl}$ reveal overall similarity, as shown in Fig. 2A. The RMSD of the 140 aligned $C_\alpha$ atoms between the two structures is 2.6 Å by DaliLite [21].

However, local structural differences are observed, especially near the active site pocket. According to the structural based sequence alignment, the loop that connects $\alpha_2$ and $\beta_2$ (Asn53-Tyr57 in CmtB) is two residue longer than that in IIA$^{Mtl}$. In IIA$^{Mtl}$, this loop consists
only of a three-residue sequence Thr-Pro-Thr. On the other hand, the loop in CmtB consists of residues Asn53-Gly54-Pro55-Tyr56-Tyr57. Due to the additional residues, helix $a_2$ is more tilted away from the active site pocket in CmtB compared to IIA Mtl, resulting in a relatively wider and shallower pocket (Fig. 2B and C). However, the bulky aromatic side chains of Tyr56 and Tyr57 that situate close to the active site in CmtB appear to partially shield the active pocket.

In addition, a positively charged residue Arg49 on helix $a_2$ in the crystal structure of IIA Mtl is in proximity with the active histidine. This residue is highly conserved in IIA Mtl throughout Gram-negative and Gram-positive bacteria species, and is considered to be critical for the function of mannitol-specific phosphotransferase [20,22,23]. In CmtB, however, this position is substituted by a serine. Instead, an Arg69 is present in the loop that also accommodates the active His67 in CmtB, and is spatially close to the active site (Fig. 2B and C). It is possible that Arg69 of CmtB plays similar roles as Arg49 in IIA Mtl during the phosphoryl transfer reactions despite their distinct locations.

Furthermore, the surface electrostatic potential of CmtB is different from IIA Mtl. The surface of IIA Mtl is comprised of almost equally distributed areas of positive and negative charges, while CmtB shows an overall acidic surface with only a few scattering positively charged spots (Fig. 3).

The protein–protein interaction surface

Previous studies of established that IIA Mtl uses a common surface for interaction with both HPr and IIB Mtl domain during the two sequential phosphoryl transfer reaction. This surface is comprised of the regions surrounding the active site pocket, including the C-terminal end of helix $a_2$, the short anti-parallel $\beta$-hairpin formed by $\beta2$ and $\beta3$, the loop containing the histidine to be phosphorylated, the loop connecting $\beta4$ and $\beta5$, and helix $a3$ [20,22,23]. In CmtB, the two loops (Asn53-Tyr57 and Tyr89-Pro96) that show higher conformational flexibility than other parts of the structure are all located in the protein–protein interaction surface. The internal motions may play a role in facilitating interactions.

Further, it has been shown that the interactions between IIA Mtl and both protein partners were found to be mainly hydrophobic [20,22,23]. An examination of the electrostatic potential of the two proteins again reveals noticeable difference. In IIA Mtl, the interaction surface contains a hydrophobic patch with surrounding charged spots, while the hydrophobic patch on the corresponding surface of CmtB appears to be larger (Fig. 3). This difference may affect their binding affinity with HPr and their downstream IIB domains in the phosphoryl transfer reactions in vivo.
The active sites are shown in dashed circles.

Summary

The structure of *E. coli* CmtB shows overall similarity with the mannitol-specific phosphotransferase IIA domain, suggesting a generally similar mechanism for phosphoryl transfer. However, the local structural differences observed at the active site pocket may underlie the differences in biological functions of the two systems in *vivo*. The widening of the active site pocket and the relatively larger hydrophobic area observed in CmtB may affect its binding specificity and affinity, as well as its efficiency in phosphoryl transfer. The distinct position of the arginine residue in the active site pocket of CmtB (Arg69) from IIAMtl (Arg49) suggest that the CmtAB system may have a slightly modified mechanism for phosphoryl transfer, and the conformations of reaction intermediates may be different from the MtlA system. Further investigations are expected to elucidate the physiological function of the CmtAB system and to correlate structure and function.

Acknowledgments

All NMR experiments were carried out at the Beijing NMR Center, Peking University. We thank Yunfei Hu, Xinxin Zhang and Jichen Hu for useful discussions. This research was supported by Grant 30325010 from the National Natural Science Foundation of China and Grant 2006AA02A323 from the National High Technology Research and Development Program of China to C.J.

References


