

Crystal Structure and Comparative Sequence Analysis of GmhA from *Colwellia psychrerythraea* Strain 34H Provides Insight into Functional Similarity with DiaA

Hackwon Do¹, Ji-Sook Yun², Chang Woo Lee^{1,3}, Young Jun Choi², Hye-Yeon Kim⁴, Youn-Jung Kim⁵, Hyun Park^{1,3}, Jeong Ho Chang^{2,*}, and Jun Hyuck Lee^{1,3,*}

The psychrophilic organism *Colwellia psychrerythraea* strain 34H produces extracellular polysaccharide substances to tolerate cold environments. Sedoheptulose 7-phosphate isomerase (GmhA) is essential for producing D-glycero-D-mannoheptose 7-phosphate, a key mediator in the lipopolysaccharide biosynthetic pathway. We determined the crystal structure of GmhA from *C. psychrerythraea* strain 34H (CpsGmhA, UniProtKB code: Q47VU0) at a resolution of 2.8 Å. The tetrameric structure is similar to that of homologous GmhA structures. Interestingly, one of the catalytic residues, glutamate, which has been reported to be critical for the activity of other homologous GmhA enzymes, is replaced by a glutamine residue in the CpsGmhA protein. We also found differences in the conformations of several other catalytic residues. Extensive structural and sequence analyses reveal that CpsGmhA shows high similarity to *Escherichia coli* DnaA initiator-associating protein A (DiaA). Therefore, the CpsGmhA structure reported here may provide insight into the structural and functional correlations between GmhA and DiaA among specific microorganisms.

INTRODUCTION

Psychrophilic organisms are cold-adapted microorganisms that have overcome the constraints associated with living in permanently cold environments (Pikuta et al., 2007; Russell, 1990).

¹Division of Polar Life Sciences, Korea Polar Research Institute, Incheon 406-840, Korea, ²Department of Biology Education, Kyungpook National University, Daegu 702-701, Korea, ³Department of Polar Sciences, University of Science and Technology, Incheon 406-840, Korea, ⁴Protein Structure Group, Korea Basic Science Institute, Chungbuk 363-883, Korea, ⁵Department of Marine Science, Incheon National University, Incheon 406-772, Korea

*Correspondence: junhyucklee@kopri.re.kr (JHL); jhcbio@knu.ac.kr (JHC)

Received 2 July, 2015; revised 1 September, 2015; accepted 15 September, 2015; published online 26 November, 2015

Keywords: *Colwellia psychrerythraea* strain 34H, CpsGmhA, DiaA, psychrophile, sedoheptulose 7-phosphate isomerase

Cold environments can induce molecular changes in an organism, including abnormalities such as low enzymatic activity, protein denaturation, and mis-folding of proteins (Russell, 1998). In order to overcome these deficiencies, several cold-adaptation mechanisms have been described, including cold-active enzymes, an unsaturated fatty-acid membrane, antifreeze proteins, and sugar and/or alcohol compounds such as glycerol, trehalose, and sorbitol (Chattopadhyay, 2006; Feller, 2013; Feller and Gerday, 2003).

One of the most interesting cold-adaptation mechanisms is biofilm formation using extracellular polysaccharide substances (EPSs) that are found in the O-antigen component of lipopolysaccharides (LPS) and cell-bound polysaccharides (Bazaka et al., 2011). The EPSs play important temperature-buffering and cryoprotectant roles in cold-adapted microorganisms (Bazaka et al., 2011; Krembs et al., 2002). The psychrophilic bacterium *Colwellia psychrerythraea* strain 34H is a well-known producer of EPSs, which function as cryoprotectants (Marx et al., 2009). Increased production of EPSs is observed when samples are frozen from -8 to -14°C, which represents a strategy similar to that used by yeasts that increase the production of cryoprotectants (polyols and/or sugar alcohol) in response to osmotic stress (Marx et al., 2009; Shen et al., 1999). Recently, the molecular structure of capsular polysaccharides synthesized from *C. psychrerythraea* strain 34H was determined and characterized. Interestingly, it resembles other known antifreeze (Glyco) proteins and appears to inhibit ice recrystallization (Carillo et al., 2015).

C. psychrerythraea strain 34H was initially isolated from Arctic marine sediments (Huston et al., 2000). The temperature of growth ranges from -12°C to 19°C and its optimal growth temperature is 8°C to 9°C (Wells and Deming, 2006). The results of genomic sequence analyses show that many gene products of *C. psychrerythraea* strain 34H play a role in the production and secretion of EPSs (Methe et al., 2005). These EPSs stabilize an extracellular protease, ColAP, against thermal denaturation (Huston et al., 2004). In addition, EPSs are involved in regulating intracellular biosynthetic reactions to proceed at extremely low temperatures; however, the external conditions that activate the production of EPSs remain unknown (Junge et al., 2006).

Gram-negative bacteria, including *C. psychrerythraea* strain

34H, contain either LPS or capsular polysaccharides (CPSs) in their outer membrane. Typical LPS is comprised of three covalently linked domains: lipid A, core oligosaccharide, and, in some bacteria, O-antigen. The core oligosaccharide is composed of inner and outer domains; the inner core contains 3-deoxy-D-manno-oct-2-ulosonic acid (Kdo), which is bound to lipid A, and L-glycero-D-manno heptose (heptose), which is linked to 3-deoxy-D-manno-oct-2-ulosonic acid (Brooke and Valvano, 1996; Nikaido, 2003; Raetz and Whitfield, 2002). Heptose is the connecting point for the variable outer-core saccharides that serve as the binding sites for O-antigen to produce complete LPS. Sedoheptulose 7-phosphate isomerase (GmhA), encoded by *gmhA*, catalyzes the conversion of sedoheptulose 7-phosphate (S7P) to D-glycero-D-mannoheptose 7-phosphate (M7P) in the first step of the heptose biosynthesis pathway (Kneidinger et al., 2001). Aberrant GmhA in *Escherichia coli* does not synthesize heptose and produces truncated LPS owing to the lack of a connecting point for more distal moieties (Brooke and Valvano, 1996).

The structure and catalytic mechanisms of GmhA from various species have been characterized because GmhA is considered an important target for antibiotic development (Harmer, 2010; Kim and Shin, 2009; Seetharaman et al., 2006; Taylor et al., 2008). In this study, we determined the crystal structure of GmhA from *C. psychrerythraea* strain 34H (CpsGmhA), which is critical for LPS biosynthesis and activity, and is potentially maintained at low temperatures for production of polysaccharides. In addition, sequence analysis reveals that CpsGmhA has high similarity with *E. coli* DnaA initiator-associating protein A (DiaA). Therefore, characterization of the structure of

CpsGmhA may provide insight into the structural and functional correlations between GmhA and DiaA among microorganisms; additionally, comparisons of these proteins in mesophilic and psychrophilic organisms may help to elucidate the mechanism of low temperature-adapted enzymes.

MATERIALS AND METHODS

Cloning, protein expression, and purification

The gene encoding GmhA protein (UniProtKB code: Q47VU0, amino acids 1-196) was amplified by polymerase chain reaction (PCR) from the genomic DNA of *C. psychrerythraea* strain 34H and subcloned into the pET 28a(+) vector (Novagen, USA). PCR amplification of the gene was performed using primers containing NdeI and XhoI restriction enzyme sites: forward primer 5'-CGATAACATATGCTAGAACAGATCAAAAAC-3' and reverse primer 5'-CGATAACTCGAGTCATGAATCACCTTGCG GGAA-3'. *E. coli* BL21(DE3) cells were transformed with the cloned pET 28a(+) vector, which contains an N-terminal 6-histidine (His) tag and a thrombin cleavage site. The *E. coli* cells were grown at 37°C until the cultures reached an optical density of approximately 0.7 at 600 nm. Heterologous expression of the target protein was induced using 1 mM isopropyl-β-D-thiogalactopyranoside (Bioneer, Korea) at 25°C overnight, and the expressed protein was collected by centrifugation at 7000 rpm for 20 min. The cell pellet was resuspended in lysis buffer (50 mM sodium phosphate [pH 8.0], 300 mM NaCl, and 5 mM imidazole) containing phenylmethane sulfonyl fluoride and disrupted by sonication (Vibra-Cell VCX400; Sonics & Materials, Inc., USA).

Table 1. X-Ray data collection and refinement statistics for CpsGmhA

	CpsGmhA
Data collection	
X-ray source	PLS-5C
Space group	C222
Unit-cell parameters (Å)	$a = 73.3, b = 109.3, c = 44.3$
Wavelength (Å)	0.97951
Resolution Range (Å)	44.3-2.8 (2.95-2.80)
No. of observed reflections	54978 (8311)
No. of unique reflections	4643 (671)
Completeness (%)	99.9 (100)
Redundancy	6.4 (6.5)
R_{merge}^a	0.136 (0.315)
I/σ	13.0 (7.3)
Refinement	
Resolution range	44.3-2.8 (3.53-2.80)
Reflections: working/free	4643/215
R_{cryst}^b	0.210 (0.234)
R_{free}^c	0.280 (0.341)
Ramachandran plot:	
favored/ allowed/disallowed (%) ^d	94.8/5.2/0
R.m.s.d. bonds (Å)	0.001
R.m.s.d. angles (°)	0.381
PDB accession code	5BY2

$$^a R_{\text{merge}} = \sum | \langle I \rangle - I | / \sum \langle I \rangle.$$

$$^b R_{\text{cryst}} = \sum | |F_o| - |F_c| | / \sum |F_o|.$$

^c R_{free} calculated with 5% of all reflections excluded from refinement stages using high-resolution data.

^dThe Ramachandran plot was calculated using MolProbity (<http://molprobity.biochem.duke.edu/>). Values in parentheses indicate the statistics for the highest resolution shells.

Table 2. Sequence information of the consensus cladogram

Protein	Species	NCBI reference sequence	
GmhA	<i>Colwellia psychrerythraea</i>	Q47VU0.1	
	<i>Escherichia coli</i>	NP_308276.1	
	<i>Campylobacter jejuni</i>	WP_002858021.1	
	<i>Pseudomonas aeruginosa</i>	WP_016253868.1	
	<i>Burkholderia pseudomallei</i>	WP_011205222.1	
	<i>Vibrio cholera</i>	WP_000284054.1	
	<i>Leptospira alstonii</i>	WP_020772754.1	
	<i>Helicobacter pylori</i>	WP_042636128.1	
	<i>Haemophilus influenza</i>	WP_005663828.1	
	<i>Bordetella bronchiseptica</i>	KFJ54879.1	
	<i>Neisseria meningitidis</i>	WP_002242076.1	
	DiaA	<i>Colwellia psychrerythraea</i>	WP_033091889.1
		<i>Yersinia enterocolitica</i>	WP_005174225.1
		<i>Pseudomonas aeruginosa</i>	WP_025299231.1
		<i>Vibrio cholera</i>	WP_032476931.1
<i>Escherichia coli</i>		WP_004018383.1	
<i>Yersinia pestis</i>		WP_002210146.1	
<i>Franconibacter pulveris</i>		WP_024559311.1	
<i>Salmonella enterica</i>		WP_000893482.1	
<i>Pluralibacter gergoviae</i>		WP_043080845.1	
<i>Erwinia toletana</i>		WP_017801381.1	
<i>Ewingella americana</i>		WP_034786756.1	
<i>Pseudomonas syringae</i>		WP_032629664.1	

Ni-NTA affinity resin (Clontech, USA) pre-equilibrated with lysis buffer was used to purify of the target protein. After washing the column and matrix with a ten-column volume of washing buffer (50 mM sodium phosphate [pH 8.0], 300 mM NaCl, and 35 mM imidazole), the His-tagged protein was eluted using a solution of the same composition but with 300 mM imidazole. Once the His tag was cleaved at 4°C overnight with thrombin (Sigma-Aldrich, USA), the resulting protein has an additional GSHM sequence at the N-terminus, an artifact of the cloning. The cleaved protein was loaded onto a Superdex-200 column (GE Healthcare, USA) equilibrated with 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM dithiothreitol (DTT). The fractions containing CpsGmhA were collected and concentrated to 25.0 mg/ml for crystallization trials.

Crystallization, data collection, and structural determination

The purified CpsGmhA was subjected to an initial crystallization screen using the commercially available crystallization solution kits MCSG I-IV (Microlytic, USA), Wizard Classic I-IV (Emerald Bio, USA), and Classics and Classics II Suite (Qiagen, Germany), following the vapor diffusion method by mixing the protein (0.8 µl) and reservoir (0.8 µl) solutions. Small crystal granules were obtained within two or three days in 0.1 M Tris-HCl (pH 8.5), 0.2 M magnesium chloride, and 30% (v/v) PEG400 or 25% (w/v) PEG3350. The optimum crystallization condition was further refined. Flat hexagonal crystals suitable for X-ray data collection were formed in a hanging drop composed of 1 µl of CpsGmhA (25.0 mg/ml), 20 mM Tris-HCl, 150 mM NaCl, and 1 mM DTT, and 1 µl of reservoir solution in 0.1 M Tris-HCl (pH 7.2), 0.2 M magnesium chloride, and 19% (v/v) PEG400 against 500 µl reservoir. N-paratone was used as a cryoprotect-

tant based on the results of a screening experiment. A native data set was collected at 0.97951 Å at the 5C beamline of the Pohang Light Source (PLS; Korea).

The data were recorded, integrated, and scaled using iMosflm software (Battye et al., 2011). In order to obtain the phase, molecular replacement (MR) was performed using the program MOLREP (Vagin and Teplyakov, 2009), using the homologous structure from *E. coli* DiaA (PDB databank accession number 2YVA; (Keyamura et al., 2007) as the model. A Matthews coefficient of 2.22 and 44.59% solvent indicated that one molecule was present in an asymmetric unit. After refinement up to 2.8 Å with REFMAC5 (Murshudov et al., 2011), the quality of the final structure was validated using the MOLPROBITY server (Chen et al., 2009) and SFCHECK from the CCP4 package (Vaguine et al., 1999), resulting in a final model with an *R*-factor of 21.0% (*R*_{free} 28.0%). The final refinement statistics are presented in Table 1. Figures and the surface and angle data were generated and calculated using PyMOL (DeLano, 2002).

Sequence analysis

The amino acid sequences of GmhA and DiaA from *C. psychrerythraea* strain 34H and other relevant species were obtained from the National Center for Biotechnology Information (NCBI) database (Table 2). Amino acid sequences were aligned using Clustal X (Chenna et al., 2003). Only well-aligned and conserved alignment sites were extracted from each alignment subset using Gblock (Castresana, 2000). Maximum likelihood (ML) analysis was carried out using RAxML Ver. 7.2.8 (Stamatakis, 2006). RAxML was run under the options GTRCAT and protein GAMMA, a complete random starting tree of 1000 bootstrap replicates (Pattengale et al., 2010).

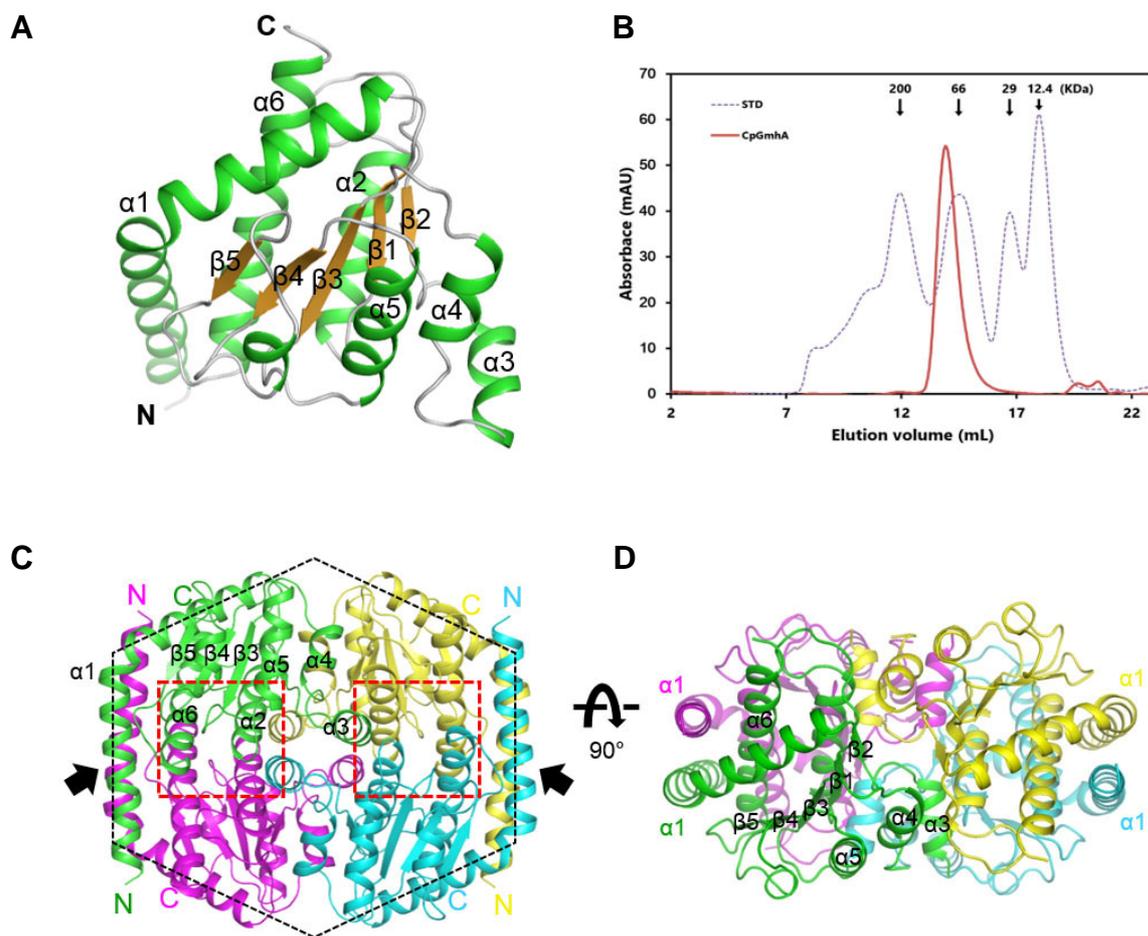


Fig. 1. Overall structure of *CpsGmhA*. (A) The overall structure of the monomeric *CpsGmhA* is shown as a cartoon representation. The α -helices, β -sheets, and loops are colored in green, brown, and grey, respectively. (B) Size exclusion chromatogram of *CpsGmhA* for the measurement of molecular weight. The estimated molecular weight is 84 kDa based on the standard (STD), whereas the calculated molecular weight of the monomer is 20.9 kDa. (C) The overall structure of tetrameric *CpsGmhA* viewed from the top. The four monomers are colored green, yellow, cyan, and magenta, respectively. The active sites are marked with red dotted lines. The black arrows indicate the clamp helical regions. The black dashed hexagon indicates the overall shape by interactions of the α_1 helices. (D) Side view of the *CpsGmhA* tetramer by 90° rotation over an X-axis from Fig. 1C.

Data deposition

The atomic coordinates and experimental structure have been deposited in the Protein Data Bank (accession codes 5BY2).

RESULTS

Overall structure

The GmhA from *C. psychrerythraea* strain 34H (*CpsGmhA*), including the additional serine-histidine residues located before the N-terminus as an artifact of cloning, was crystallized and the structure was determined by X-ray crystallography at a resolution of 2.8 Å. The monomeric structure is composed of five central parallel β -sheets surrounded by six α -helices (Fig. 1A). Most residues fit well in the electron density except for residues Tyr93, Gln94, Gln193, Gly194, Asp195, and Ser196. The final model was refined to R and R_{free} values of 21.0% and 28.0%, respectively (Table 1). The overall topology of *CpsGmhA* is very similar to that of the homologous structures from various species, with an average root mean square deviation (r.m.s.d.) of 1.3 Å in monomeric $C\alpha$ (Harmer, 2010; Kim and Shin, 2009; Seetharaman et al., 2006; Taylor et al., 2008).

Since one molecule of *CpsGmhA* was observed in the asymmetric unit, we performed a size exclusion chromatographic analysis. The result indicates that *CpsGmhA* also exists in tetrameric form in solution, which is in agreement with the results of previous studies (Harmer, 2010; Taylor et al., 2008) (Fig. 1B).

The bulk of helices (α_1 - α_4 and α_6) form a tetramer with a broad surface range. Half of the long N-terminal α_1 helix interacts with the same region of the neighboring monomers via the hydrophobic residues Met1, Leu2, Ile5, Phe9, Ile13, Ile17, and Ile27, resulting in a four-helix frame around the tetramer (Figs. 1C and 1D). The four α_1 -helices in the tetramer form a hexagon tilted at an angle of 120°, which could be important for structural stabilization, because the four α_1 -helices surrounding the tetramer have broad contacting surfaces that serve as clamps on each of the two subunits (Fig. 1C). Extensive interactions with neighboring molecules support the order of the α_3 helix, such that the overall shape has a closed conformation in

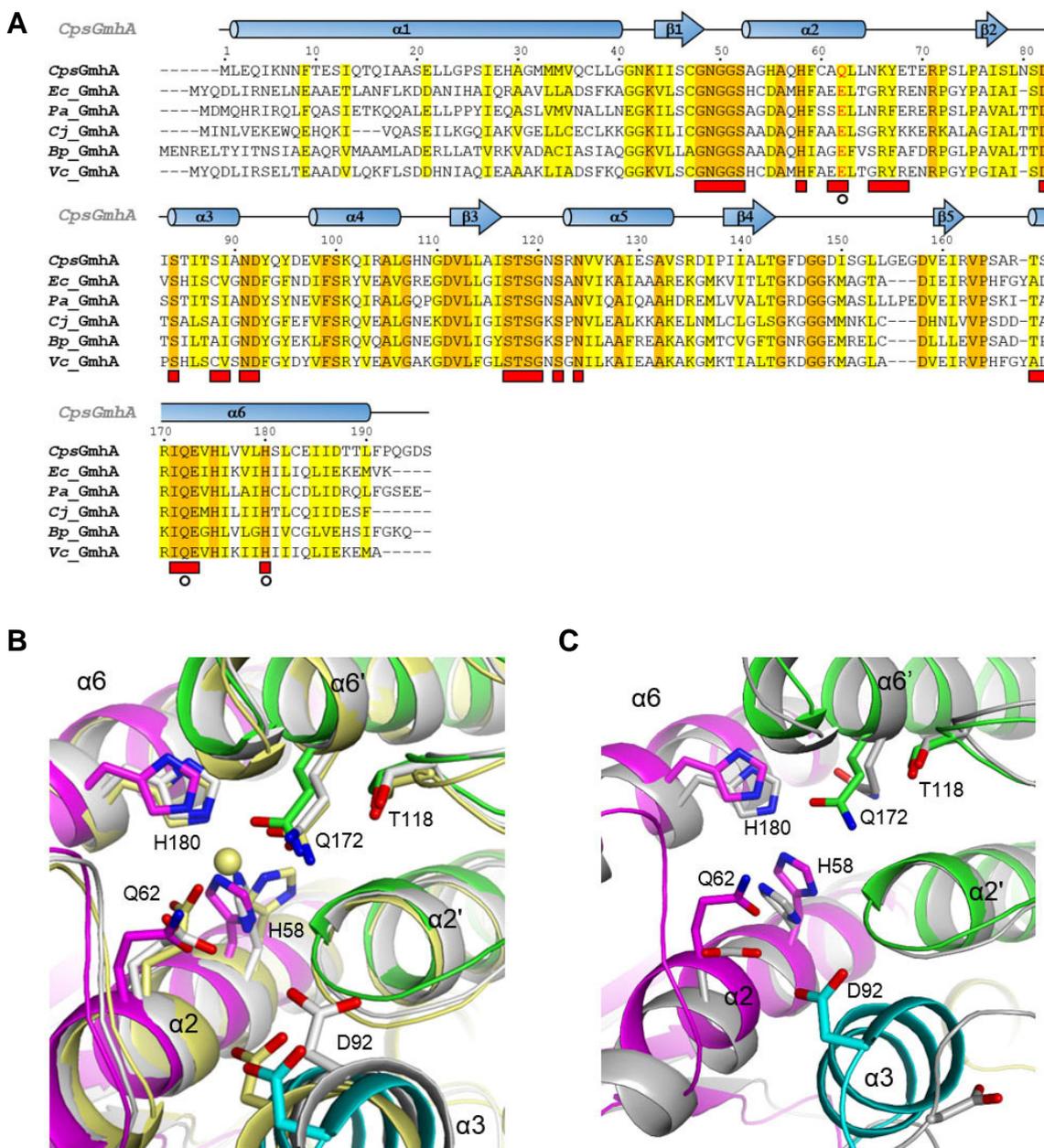


Fig. 2. Active site of *CpsGmhA*. (A) Structure-based multiple sequence alignment of *CpsGmhA* with other homologs. Conserved residues are presented in either orange or yellow backgrounds depending on their degree of conservation. The residues composed of the active site are indicated with red rectangles under the sequences. The catalytic residues are marked with open circles. The secondary structure was generated based on the structure of *CpsGmhA*. Sequences were aligned using Clustal X 2.1 (Chenna et al., 2003) and photographs were generated using the program ESPript (Gouet et al., 1999). (B) Overlay of the catalytic residues of *CpsGmhA* with those of *BpGmhA* (PDB id: 2X3Y) and *CjGmhA* (PDB id: 1TK9). The colors of *CpsGmhA* are the same as those described for Fig. 1C. *BpGmhA* and *CjGmhA* are colored in yellow and grey, respectively. The yellow sphere indicates the zinc ion present in the *BpGmhA* structure. Oxygen and nitrogen atoms are shown in red and blue, respectively. (C) Overlay of the catalytic residues of *CpsGmhA* with *EcGmhA* (PDB id: 212W), which is shown in grey.

comparison with homologous structures (see below) (Harmer, 2010; Taylor et al., 2008).

Active site

The active site of GmhA is mainly composed of four α -helices, which are each composed of two equivalent helices ($\alpha 2$ - $\alpha 6$ and

$\alpha 2'$ - $\alpha 6'$) from two subunits. In addition, the conserved residues on the $\alpha 3$ helix and the three loops ($\beta 1$ - $\alpha 2$, $\alpha 2$ - $\beta 2$, and $\beta 3$ - $\alpha 5$) support the conformation of the active site (Fig. 2A). The E-Q-H residues, which are located at the $\alpha 2$, $\alpha 6'$, and $\alpha 6$ helices, respectively, are strongly involved in catalytic reactions. The E-Q-H residues are highly conserved in the GmhA family; however,

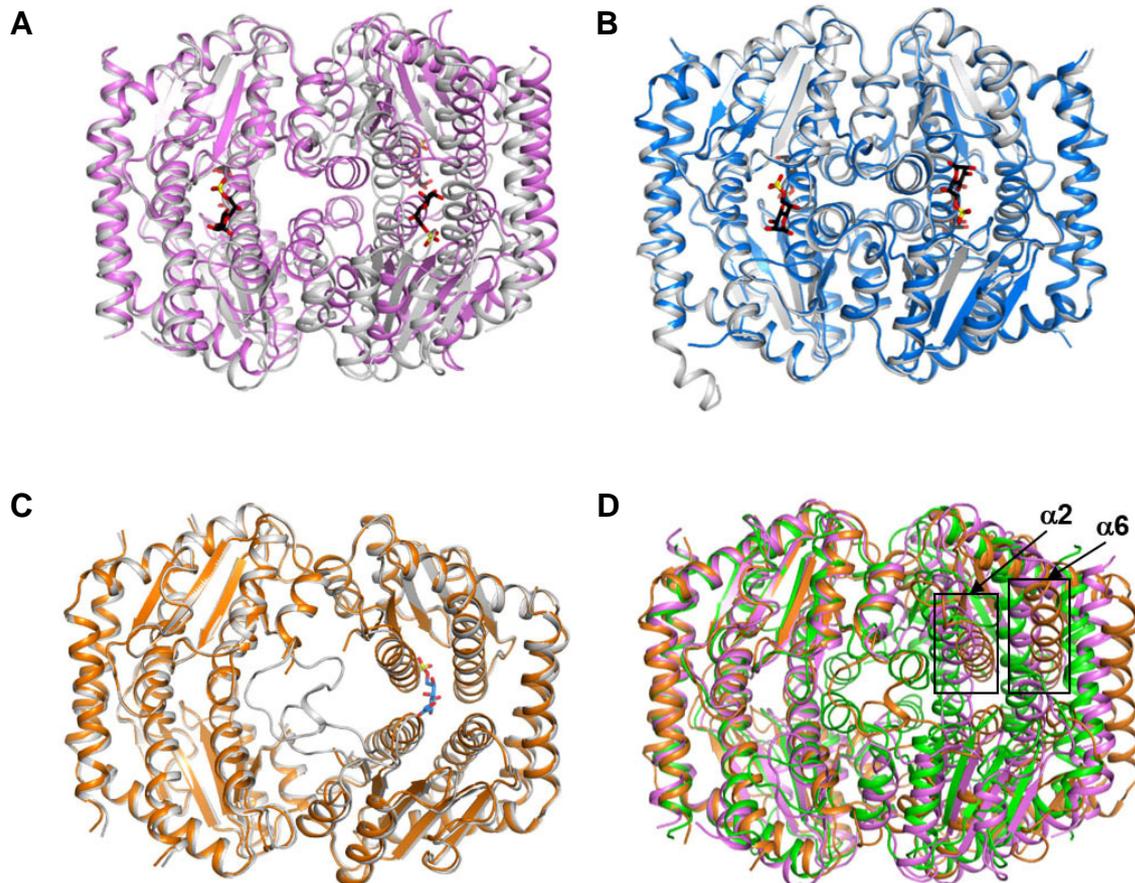


Fig. 3. Comparison of GmhA homologous structures. (A) Overlay of the apo and M7P product-bound GmhA structures from *P. aeruginosa*. The *PaGmhA*-M7P complex (PDB id: 1X92) and apo *PaGmhA* (PDB id: 3BJZ) are shown in pink and grey, respectively. The M7P molecules are shown in black with the phosphorous and oxygen atoms in yellow and red, respectively. (B) Overlay of the apo and M7P product-bound GmhA structures from *B. pseudomallei*. The *BpGmhA*-M7P complex (PDB id: 2XBL) and apo *BpGmhA* (PDB id: 2X3Y) are shown in blue and grey, respectively. (C) Overlay of the apo and substrate S7P-bound GmhA structures from *E. coli*. The *EcGmhA*-S7P complex (PDB id: 2I22) and apo *EcGmhA* (PDB id: 2I2W) are shown in orange and grey, respectively. The S7P molecules are shown in blue. (D) Overlay of the apo structures of *CpsGmhA* (green), *PaGmhA* (pink) and *EcGmhA* (orange). The tilted $\alpha 2$ and $\alpha 6$ helices of *EcGmhA* are presented in the black squares.

the glutamate residue is replaced by glutamine at the equivalent position in *CpsGmhA* (Fig. 2A). Previous mutagenesis studies have shown no GmhA activity in EQ mutants (Harmer, 2010; Taylor et al., 2008), suggesting that *CpsGmhA* might be more closely related to DiaA this inference is supported by the bioinformatic analysis (see below).

The tetrameric structures of GmhA can be classified into two distinct conformations, open or closed, depending on whether an additional $\alpha 3'$ helix is present in the central region. The closed conformation may allow for the ingress and egress of a substrate and product in solution due to the extension of the solvent accessible area of the active site (Harmer, 2010). Since the conformation of *CpsGmhA* is closed, we compared the positions of the catalytic residues to those in other GmhAs with closed conformation from *Burkholderia pseudomallei* and *Campylobacter jejuni* (Fig. 2B). Most residues are oriented in a similar manner, except for the Glu63 residue of *C. jejuni* GmhA (*CjGmhA*), which is equivalent to the Gln62 residue in *CpsGmhA*. However, in comparison with those of the *E. coli*

GmhA (*EcGmhA*), conformations of the catalytic residues are not well overlaid (Fig. 2C). In particular, the conformation of the Asp94 residue, which is equivalent to Asp92 in *CpsGmhA*, is quite different in that it faces the opposite side to Asp92. In addition, the positions of the Glu65 and Gln172 residues (equivalent to Gln62 and Gln172 of *CpsGmhA*, respectively) are different. Therefore, the conformations of the catalytic residues are affected by whether or not the overall structure is an open or closed conformation.

Comparison with homologous GmhA structures

Conformational differences between the open and closed may be because of the binding of the product to the active site. For example, the open-conformation structure of *PaGmhA* is converted to a closed structure by M7P binding (Taylor et al., 2008) (Fig. 3A). M7P binding drives the $\alpha 4$ helix to move away from the central cavity, resulting in a disordered $\alpha 3$ helix as a result of its interaction with the $\alpha 2$ of the other subunit. However, the *B. pseudomallei* and *E. coli* GmhA structures are not changed

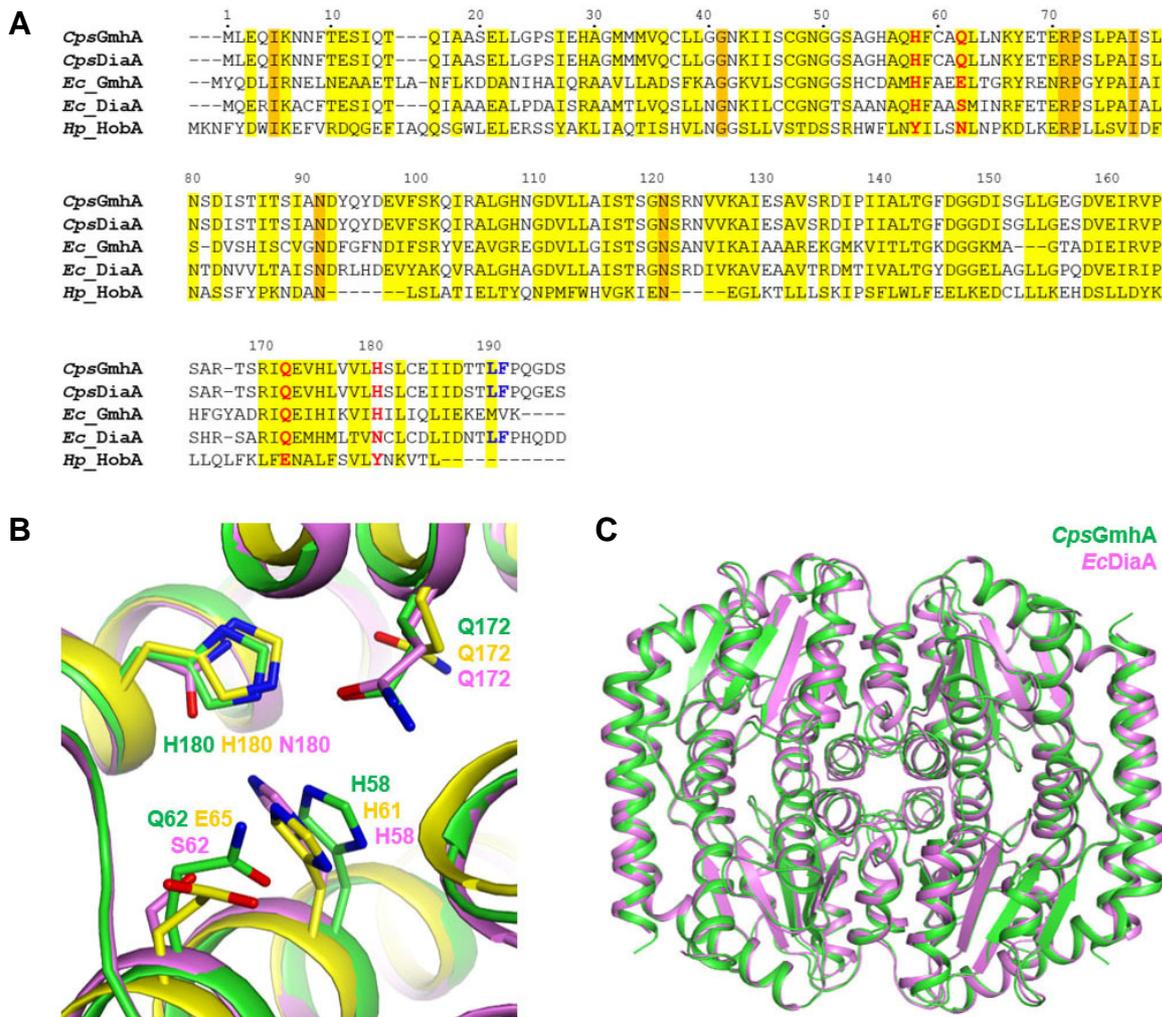


Fig. 4. Comparison of GmhA and DiaA structures. (A) Multiple sequence alignment of *CpsGmhA*, *CpsDiaA*, *EcGmhA*, *EcDiaA*, and *HpHobA*. Conserved residues are presented in either orange or yellow background according to the degree of conservation. The catalytic residues are indicated in red. The residues known to be involved in DnaA binding are presented in blue. (B) Overlay of the catalytic residues of *CpsGmhA* (green), *EcGmhA* (yellow), and *EcDiaA* (pink). Each of the residue numbers are indicated by their corresponding colors. (C) Overlay of the overall structures of *CpsGmhA* and *EcDiaA*. *CpsGmhA* and *EcDiaA* are shown in green and pink, respectively.

upon product or substrate binding (Harmer, 2010; Taylor et al., 2008) (Figs. 3B and 3C). In addition, the open conformations of *EcGmhA* and *PaGmhA* are slightly different in that the $\alpha 2$ and $\alpha 6$ helices in *EcGmhA* are tilted outwards at 25° and 15° , respectively (Fig. 3D). Therefore, the two pairs of subunits in *EcGmhA* are not symmetric, and the three different GmhA structures exist in two open states and one closed state (Fig. 3D). Taken together, no correlations are currently evident between the open and closed conformation and the structures.

Similarity between *CpsGmhA* and *EcDiaA*

We searched for structural homologs of *CpsGmhA* using the DALI server system and found that DiaA from *E. coli* (*EcDiaA*) has high similarity to GmhA, with a Z-score of 32.7 (Holm and Rosenstrom, 2010). DiaA is essential for the accurate initiation and precise timing of chromosomal replication during the cell cycle (Ishida et al., 2004). Thus, DiaA directly and specifically

interacts with DnaA independent of nucleotides or DNA (Ishida et al., 2004; Keyamura et al., 2007). Moreover, the structure of DiaA shows similarity with the DnaA-binding protein HobA from *Helicobacter pylori* (*HpHobA*) (Natrajan et al., 2007). Based on this structural information, we compared the protein sequences and found that the catalytic glutamate residues are variable, with Gln (*CpsGmhA*), Ser (*EcDiaA*), and Asn (*HpHobA*) being observed at catalytic sites (Fig. 4A). In addition, the other catalytic residues are varied and HobA has no identical residues. Given the difference in the catalytic residues of *HpHobA*, we compared the conformations of the catalytic residues among *CpsGmhA*, *EcGmhA*, and *EcDiaA* (Fig. 4B). The conformation of the catalytic residues between *CpsGmhA* and *EcDiaA* is much more similar than that of other pairs of compared proteins because the two residues Glu65 and Gln172 are differently oriented. This result indicates that *CpsGmhA* and *EcDiaA* are closely related.

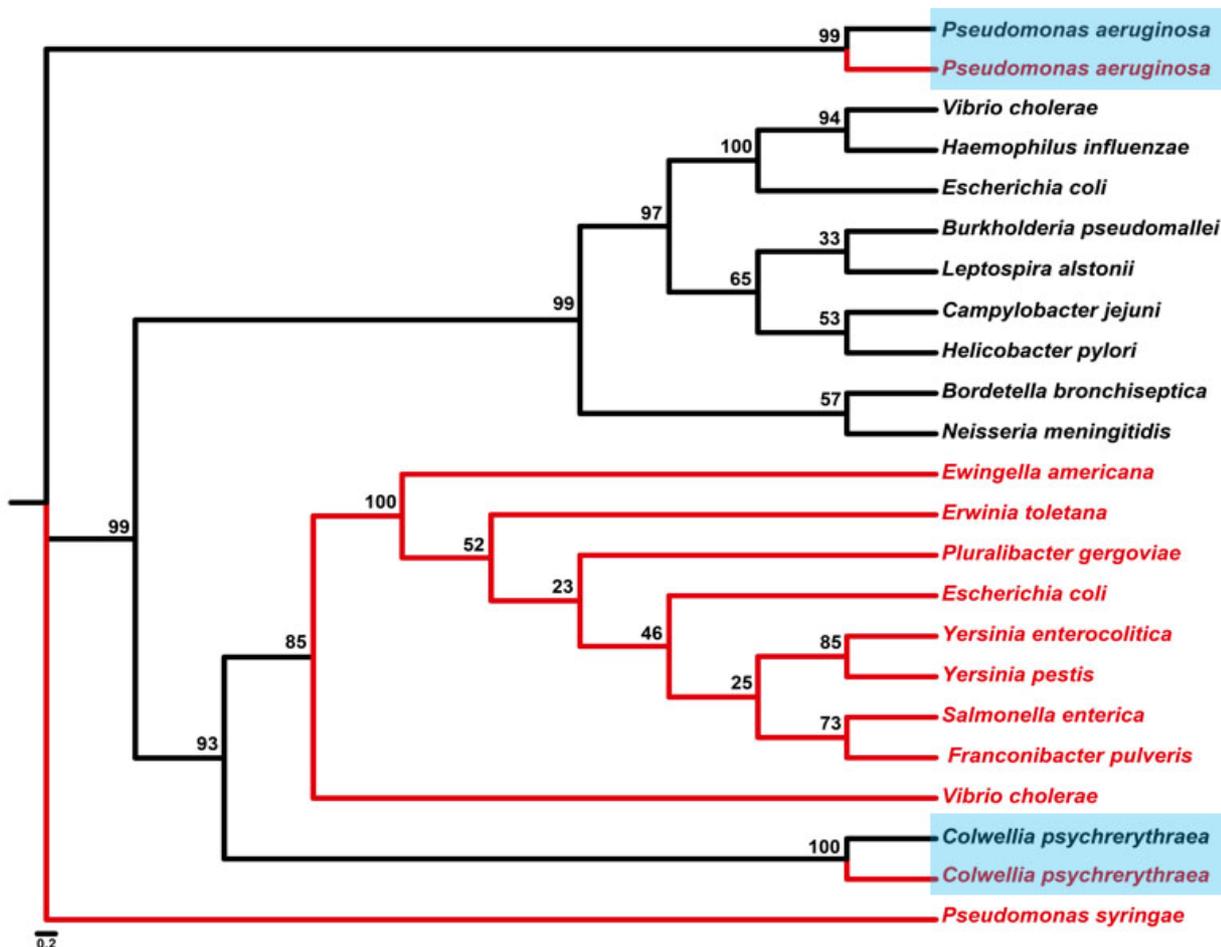


Fig. 5. Phylogenetic analysis of GmhA and DiaA. Phylogenetic relationships of DiaA and GmhA based on their amino acids. The consensus cladogram inferred from the amino acid sequence data of DiaA and GmhA using maximum likelihood (ML) analysis under GTRCAT and protein GAMMA models. The group of DiaA is highlighted in red color. The paired GmhA and DiaA from *C. psychrerythraea* and *P. aeruginosa* are highlighted in sky blue. The numbers above the branches indicate bootstrap percentages.

Next, we compared the structural similarity among the four structures. Structural comparisons show that *CpsGmhA* has higher similarity to *EcDiaA* with a r.m.s.d of 0.7 Å, but relatively low similarity to both *EcGmhA* and *HpHobA* with an r.m.s.d of 1.7 Å and 3.7 Å, respectively. Thus, the overall tetrameric structures of *CpsGmhA* and *EcDiaA* are completely overlaid (Fig. 4C). To investigate the sequence-based relationship between *CpsGmhA* and *CpsDiaA*, we performed a phylogenetic analysis of GmhA and DiaA (Fig. 5). GmhA and DiaA could be divided into two distinct groups. However, we found that GmhA and DiaA from *C. psychrerythraea* and *Pseudomonas aeruginosa* are not clearly separated. Indeed, GmhA and DiaA from the two species showed over 98% similarity because they differ in only two or three residues. This result indicates that *PaDiaA* could be classified as an isotype of *PaGmhA* and *CpsGmhA* as an isotype of *CpsDiaA*. Taken together, the structural and sequence similarities of *CpsGmhA*, *CpsDiaA* and *EcDiaA* may provide insight into the functional role of GmhA as an isotype of DiaA in *C. psychrerythraea*.

DISCUSSION

To date, eight crystal structures of GmhA have been reported from various mesophilic species (Harmer, 2010; Kim and Shin, 2009; Seetharaman et al., 2006; Taylor et al., 2008). The overall structures are divided into two forms: open and closed. The conformations of the apo-structures are also variable in that the GmhAs of *E. coli*, *Vibrio cholerae*, and *P. aeruginosa* are open, whereas those of *B. pseudomallei*, *C. jejuni*, and *C. psychrerythraea* are closed. The overall structure of psychrophilic *CpsGmhA* is very similar to the homologous structures of the closed forms. Previous studies of GmhA crystal structures demonstrate that either two or four molecules are contained in the asymmetric unit. However, only one molecule of *CpsGmhA* was found in the asymmetric unit. The biologically reliable tetrameric structure was revealed by the C222 symmetry operation. Additional analysis of oligomerization using the program PDBEPIA (v1.47) further indicated that biological assembly of the protein results in a tetramer (Krissinel and Henrick, 2007).

The reaction mechanisms of GmhA have been proposed by complex structures with either product or substrate (Harmer, 2010; Taylor et al., 2008). To date, there have been three GmhA complex structures reported in either substrate form or product-bound form. The substrate S7P is only present in the *EcGmhA*

structure as an open conformation, and the product M7P is found bound in closed structures from *P. aeruginosa* and *B. pseudomallei* (Harmer, 2010; Taylor et al., 2008). Moreover, only one subunit of the tetrameric *EcGmhA* contains the substrate S7P, whereas *BpGmhA* and *PaGmhA* both have M7P products in each of the four subunits (Fig. 4C). Thus, the two subunits are not symmetrical about the vertical axis, as is the case with other GmhA structures. This may indicate that a local conformational change occurs upon substrate binding. However, there was no difference in symmetry between the protein and its apo-structure. On the other hand, it is also likely that a change in only one molecule of the substrate would not induce an overall conformational change.

As shown in the previous section, only *CpsGmhA* contains a catalytic glutamine, rather than a glutamate residue at the equivalent position (Gln62) in the active site. Therefore, *CpsGmhA* may not have isomerase activity, based on mutagenesis data reported previously (Harmer, 2010; Taylor et al., 2008). Furthermore, based on sequence identity, *CpsGmhA* is more closely related to *EcDiaA* (62.2% identical) than to *EcGmhA* (37% identical) or *HpHobA* (13.9% identical). Interestingly, *CpsGmhA* has the two signature residues, Leu190 and Phe191, in common with those of *EcDiaA*, which are critical for DnaA binding, whereas *EcGmhA* does not contain these residues (Fig. 5A) (Keyamura et al., 2007). In addition, there are only three residue discrepancies between *CpsGmhA* and *CpsDiaA*, and therefore no significant functional differences would be expected. Since a functional study of *CpsGmhA* has not yet been done, we propose that the *CpsGmhA* protein could be a putative isotype of *CpsDiaA*. Taken together, further investigation of *CpsGmhA* is required to confirm the functional similarity with *CpsDiaA*.

ACKNOWLEDGMENTS

We would like to thank the staff at the X-ray core facility of the Korea Basic Science Institute (KBSI; Ochang, Korea) and Yeon-Gil Kim, the beamline staff at PLS-5C of the Pohang Light Source (Pohang, Korea) for the data collection. This work was supported by the Korea Polar Research Institute (KOPRI; grant No. PE14330 to JHL) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology to JHC (2013R1A1A1061391).

REFERENCES

Battye, T.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 67, 271-281.

Bazaka, K., Crawford, R.J., Nazarenko, E.L., and Ivanova, E.P. (2011). Bacterial extracellular polysaccharides. In *Bacterial adhesion* (Springer), pp. 213-226.

Brooke, J.S., and Valvano, M.A. (1996). Biosynthesis of inner core lipopolysaccharide in enteric bacteria: identification and characterization of a conserved phosphoheptose isomerase. *J. Biol. Chem.* 271, 3608-3614.

Carillo, S., Casillo, A., Pieretti, G., Parrilli, E., Sannino, F., Bayer-Giraldi, M., Cosconati, S., Novellino, E., Ewert, M., Deming, J.W., et al. (2015). A unique capsular polysaccharide structure from the psychrophilic marine bacterium *Colwellia psychrerythraea* 34H that mimics antifreeze (glyco)proteins. *J. Am. Chem. Soc.* 137, 179-189.

Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540-552.

Chattopadhyay, M. (2006). Mechanism of bacterial adaptation to low temperature. *J. Biosci.* 31, 157-165.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2009). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 66, 12-21.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. (2003). Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31, 3497-3500.

DeLano, W.L. (2002). The PyMOL molecular graphics system.

Feller, G. (2013). Psychrophilic enzymes: from folding to function and biotechnology. *Scientifica (Cairo)* 2013, 512840.

Feller, G., and Gerday, C. (2003). Psychrophilic enzymes: hot topics in cold adaptation. *Nat. Rev. Microbiol.* 1, 200-208.

Gouet, P., Courcelle, E., and Stuart, D.I. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305-308.

Harmer, N.J. (2010). The structure of sedoheptulose-7-phosphate isomerase from *Burkholderia pseudomallei* reveals a zinc binding site at the heart of the active site. *J. Mol. Biol.* 400, 379-392.

Holm, L., and Rosenstrom, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38, W545-549.

Huston, A.L., Krieger-Brockett, B.B., and Deming, J.W. (2000). Remarkably low temperature optima for extracellular enzyme activity from Arctic bacteria and sea ice. *Environ. Microbiol.* 2, 383-388.

Huston, A.L., Methe, B., and Deming, J.W. (2004). Purification, characterization, and sequencing of an extracellular cold-active aminopeptidase produced by marine psychrophile *Colwellia psychrerythraea* strain 34H. *Appl. Environ. Microbiol.* 70, 3321-3328.

Ishida, T., Akimitsu, N., Kashioka, T., Hatano, M., Kubota, T., Ogata, Y., Sekimizu, K., and Katayama, T. (2004). DiaA, a novel DnaA-binding protein, ensures the timely initiation of *Escherichia coli* chromosome replication. *J. Biol. Chem.* 279, 45546-45555.

Junge, K., Eicken, H., Swanson, B.D., and Deming, J.W. (2006). Bacterial incorporation of leucine into protein down to -20 degrees C with evidence for potential activity in sub-eutectic saline ice formations. *Cryobiology* 52, 417-429.

Keyamura, K., Fujikawa, N., Ishida, T., Ozaki, S., Suetsugu, M., Fujimitsu, K., Kagawa, W., Yokoyama, S., Kurumizaka, H., and Katayama, T. (2007). The interaction of DiaA and DnaA regulates the replication cycle in *E. coli* by directly promoting ATP-DnaA-specific initiation complexes. *Genes Dev.* 21, 2083-2099.

Kim, M.S., and Shin, D.H. (2009). A preliminary X-ray study of sedoheptulose-7-phosphate isomerase from *Burkholderia pseudomallei*. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 65, 1110-1112.

Kneidinger, B., Graninger, M., Puchberger, M., Kosma, P., and Messner, P. (2001). Biosynthesis of nucleotide-activated D-glycero-D-manno-heptose. *J. Biol. Chem.* 276, 20935-20944.

Krembs, C.e., Eicken, H., Junge, K., and Deming, J. (2002). High concentrations of exopolymeric substances in Arctic winter sea ice: implications for the polar ocean carbon cycle and cryoprotection of diatoms. *Deep Sea Res. Part I* 49, 2163-2181.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774-797.

Marx, J.G., Carpenter, S.D., and Deming, J.W. (2009). Production of cryoprotectant extracellular polysaccharide substances (EPS) by the marine psychrophilic bacterium *Colwellia psychrerythraea* strain 34H under extreme conditions. *Can J. Microbiol.* 55, 63-72.

Methe, B.A., Nelson, K.E., Deming, J.W., Momen, B., Melamud, E., Zhang, X.J., Moul, J., Madupu, R., Nelson, W.C., Dodson, R.J., et al. (2005). The psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea* 34H through genomic and proteomic analyses. *Proc. Natl. Acad. Sci. USA* 102, 10913-10918.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 67, 355-367.

Natrajan, G., Hall, D.R., Thompson, A.C., Gutsche, I., and Terradot, L. (2007). Structural similarity between the DnaA-binding proteins HobA (HP1230) from *Helicobacter pylori* and DiaA from *Escherichia coli*. *Mol. Microbiol.* 65, 995-1005.

Nikaido, H. (2003). Molecular basis of bacterial outer membrane

- permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593-656.
- Pattengale, N.D., Alipour, M., Bininda-Emonds, O.R., Moret, B.M., and Stamatakis, A. (2010). How many bootstrap replicates are necessary? *J. Comput. Biol.* **17**, 337-354.
- Pikuta, E.V., Hoover, R.B., and Tang, J. (2007). Microbial extremophiles at the limits of life. *Crit. Rev. Microbiol.* **33**, 183-209.
- Raetz, C.R., and Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635-700.
- Russell, N.J. (1990). Cold adaptation of microorganisms. *Philos. Trans. R Soc. Lond. B. Biol. Sci.* **326**, 595-608, discussion 608-511.
- Russell, N.J. (1998). Molecular adaptations in psychrophilic bacteria: potential for biotechnological applications. *Adv. Biochem. Eng. Biotechnol.* **61**, 1-21.
- Seetharaman, J., Rajashankar, K.R., Solorzano, V., Kniewel, R., Lima, C.D., Bonanno, J.B., Burley, S.K., and Swaminathan, S. (2006). Crystal structures of two putative phosphoheptose isomerases. *Proteins* **63**, 1092-1096.
- Shen, B., Hohmann, S., Jensen, R.G., and Bohnert, H.J. (1999). Roles of sugar alcohols in osmotic stress adaptation. Replacement of glycerol by mannitol and sorbitol in yeast. *Plant Physiol.* **121**, 45-52.
- Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.
- Taylor, P.L., Blakely, K.M., de Leon, G.P., Walker, J.R., McArthur, F., Evdokimova, E., Zhang, K., Valvano, M.A., Wright, G.D., and Junop, M.S. (2008). Structure and function of sedoheptulose-7-phosphate isomerase, a critical enzyme for lipopolysaccharide biosynthesis and a target for antibiotic adjuvants. *J. Biol. Chem.* **283**, 2835-2845.
- Vagin, A., and Teplyakov, A. (2009). Molecular replacement with MOLREP. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 22-25.
- Vaguine, A.A., Richelle, J., and Wodak, S. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **55**, 191-205.
- Wells, L.E., and Deming, J.W. (2006). Characterization of a cold-active bacteriophage on two psychrophilic marine hosts. *Aquat. Microb. Ecol.* **45**, 15-29.