· Article · Advances in Polar Science
doi: 10.13679/j.advps. 2019.0030 June 2020 Vol. 31 No. 2: ***-***

Structure and function of a novel cold regulated cold shock domain containing protein from an obligate psychrophilic yeast, *Glaciozyma antarctica*

Jennifer CHARLES, Makdi MASNODDIN, Farhan NAZAIE & Nur Athirah YUSOF*

Biotechnology Research Institute, Universiti Malaysia Sabah, Jalan UMS 88400 Kota Kinabalu, Sabah, Malaysia

Received on 9 September 2019; accepted on 31 January 2020; published online 1 June 2020

Abstract Cold shock domain (CSD)-containing proteins are one of the groups of the evolutionarily conserved nucleic acid-binding proteins in all three domains of life consisting of an ancient beta-barrel fold that serves to bind nucleic acids. The cDNA of a novel protein-coding gene containing CSD was cloned from *Glaciozyma antarctica* designated as Ga16676. The full length of *Ga16676* gene with the size of 1335 bp encodes for an N-terminal CSD with conserved nucleic acids binding motif RNP1 and RNP2. The *Ga16676* gene was cloned in pET30 Ek/LIC, sequenced, expressed and its resistance towards cold was characterized. Recombinant protein expression of Ga16676 showed overexpressed soluble expression in both supernatant and pellet forms at 20 ℃. The effects of recombinant CSD protein overexpression on colony formation shows that *E. coli* cells were able to grow at 37 ℃ and 20 ℃ but not at 4 ℃ while *E. coli*_Ga16676 cells were able to grow at all temperatures tested. In addition, *E. coli*_Ga16676 cells showed higher growth rate compared to empty *E. coli* cells at 10 ℃. Structural analysis of Ga16676 reveals some interesting findings such as more aromatic interactions for efficient binding in low energy environment, a longer loop that may contribute to structural flexibility and clustering of charged amino acids on the protein surface that is important for protein stability and flexibility.

Keywords cold shock protein, CSP, cold adaptation, cold stress, stress response, stress tolerance, *Glaciozyma antarctica*, psychrophilic yeast

Citation: Charles J, Masnoddin M, Nazaie F, et al. Structure and function of a novel cold regulated cold shock domain containing protein from an obligate psychrophilic yeast, *Glaciozyma Antarctica*. Adv Polar Sci, 2020, 31(2): **-**, doi: 10.13679/j.advps.2019.0030

1 Introduction

More than three-quarters or approximately 80% of the Earth's surface is covered by cold-oceans with a temperature below 5 ℃, including areas such deep oceans, glaciers, and polar regions (Russell, 1990; Margesin et al., 2007). The extremely low temperature in the Antarctic made it almost impossible for the survival of most organisms (Rogers et al., 2012). Indeed, temperature remains a critical environmental factor controlling the evolution and biodiversity of life on earth. The microorganisms that occupy these regions are identified as psychrophiles which are able to proliferate faster at 15 ℃ or below but unable to proliferate above 20 ℃ (Margesin et al., 2007). Due to the extreme environments, psychrophiles meet a variety of stressful and changing environmental conditions that challenge their optimal development. *Glaciozyma antartica* is an obligate psychrophilic yeast that strives in such harsh environments (Hashim et al., 2013). It

* Corresponding author, ORCID: 0000-0002-9298-7425, E-mail: nrathirah.yusof@ums.edu.my

was isolated from the sea ice near Casey Station, Antarctica (Firdaus-Raih et al., 2018). This yeast lives in the Antarctic marine waters where temperatures range from −2.2 ℃ to 4 ℃. According to Boo et al. (2013), *G. antarctica* is classified as an obligate psychrophilic yeast based on its optimum temperature, reported as 12 ℃, and tolerance of temperatures up to 20 ℃. One of the main challenges for organisms in cold environments is low temperatures that decrease the rates of biochemical reactions and thereby affect cell growth (D'Amico et al., 2006). Temperature downshift causes the cell membrane fluidity to decrease which affects protein translocation and secretion hence, cells produce cold shock proteins to help overcome these challenges (Phadtare and Severinov, 2010).

Experimental studies also have shown that cold-shock protein is responsible for the stabilization of membrane fluidity caused by cold stress. This involved the increasing production of unsaturated fatty acids, which involves fatty acid desaturase activity (Nakagawa et al., 2002). Cold shock proteins are small nucleic acid-binding proteins ranging from 65 to 75 amino acids in length (Graumann and Marahiel, 1996; Czapski and Trun, 2014). Cold shock proteins act upon harmful effects of temperature downshift and thus help the cells to adapt (Phadtare, 2004). Interestingly, studies have shown that Csps not only are important during temperature downshift but also play a wider role in stress tolerance of bacteria (Michaux et al., 2012; Schärer et al., 2013; Wang et al., 2014; Derman et al., 2015). The cold shock protein A (CspA) in *E. coli* function as RNA chaperone which aid in low affinity for RNA binding, low sequence specificity, and ability to destabilize the secondary structures in RNA resulted in them to be more susceptible to RNases, therefore preventing formation of secondary structure which is responsible to facilitating the efficient translation of mRNAs (Vasina and Baneyx, 1996). Cold shock proteins were initially identified in bacteria, a sudden drop in temperature (37 ℃℃ to 10 ℃) induced a 200-fold increase in CspA, which was independent of transcriptional activity (Jones and Inouye, 1994; Gottesman, 2018). The cold shock proteins are the most evolutionary conserved proteins (Jones and Inouye, 1994; Lindquist and Mertens, 2018) which can be found in psychrophilic, mesophilic, thermophilic and even hyperthermophilic bacteria (Phadtare, 2004; Jin et al., 2014).

In contrast with the various studies conducted with heat shock in yeast, the mechanisms of the cold stress response are still poorly understood. Cold-shock proteins have been identified in yeast and recently, protein with thermotolerance and molecular chaperone properties has been studied in *G. antartica* (Yusof et al., 2016; Firdaus Raih et al., 2018). Cold-shock response has been studied in eukaryotes, but no well-defined conservation of cold-shock proteins has been found (Panadero et al., 2006; Yang et al., 2012). Similar to heat shock, cold-stress was reported to be regulated by classical stress pathways involving the high-osmolarity glycerol mitogen-activated protein kinase Hog1, and the transcriptional factors Msn2 and Msn4 (Panadero et al., 2006). However, there is still very limited information about the signaling events resulting from changes in temperature. Furthermore, the mechanism of how cells react with environmental stress varies between species. Thus, in order to understand how *G. antarctica* adapts to cold, harsh environments, one must understand every aspect of its physiology and biology. In this study, we present the first exploration of the role of cold shock domain in a novel protein from an obligate psychrophilic yeast. *Ga16676* is a novel protein-coding gene with uncharacterized function, also known as hypothetical protein, which has been identified in our study (unpublished data) as having significant expression patterns. The gene contains a CSD with the size of 207 bp which has sequence variants that confer thermal adaptation. We aimed to shed light on the cold adaptation strategies acquired by this cold shock domain containing protein. For this purpose, we have cloned the whole gene of *Ga16676* with the size of 1335 bp and proceeded to functional and structural analysis. We found out several structural features that allowed this protein to adapt to the low temperatures. The elucidation of the molecular adaptations that rendered this protein cold active activities expends our current understanding about protein cold adaptation and for biotechnological applications.

2 Materials and methods

2.1 *G. antarctica* culture

G. antarctica cells were cultured in yeast peptone dextrose broth (10 % (w/v) yeast extract, 20 % (w/v) peptone, and 20 % (w/v) dextrose) at 12 ℃ until the OD_{600} reached approximately 0.6–0.8. The cells were harvested and stored at −80 ℃ until further usage.

2.2 RNA extraction

The following materials were treated overnight with diethylpyrocarbonate (DEPC) 0.1 % (v/v) and subsequently autoclaved: pestle and mortar, microcentrifuge tubes, pipette tips, and spatula. The RNA extraction was performed using TRIzol® reagent (Invitrogen, USA). The concentration and purity of total RNA were measured using a Nanodrop spectrophotometer at 260 nm (ThermoScientific, USA). The RNA quality was determined after running total RNA on a 1 % (w/v) agarose gel at 120 V for 1 h. Total RNA was stored at −80 ℃.

2.3 Cloning and sequence analysis

Ga16676 complementary DNA (cDNA) was amplified using specific primers that promoted the ligation-independent cloning of target genes into the pET30 Ek/LIC (Merck-Millipore, Germany) vector (Table 1). The direct ligation of amplified *Ga16676* cDNA with the pET30

Ek/LIC vector was prepared according to the manufacturer's instructions. Then, the ligation mix was transformed into an *E. coli* BL21(DE3) expression host. The positive transformants were identified using sequencing analysis. The *Ga16676* sequence was deposited in GenBankTM under accession number MH719220.

Table 1 Primers used in this study

Primers	Sequence (5' – 3')
Ga16676F-FS	CCACCCATGGCTACATCAGCTT
Ga16676R-FS	GCTCCTCAAGAGAAGTCGCGAT
Ga16676F-LIC	GACGACGACAAGATGGCTACATCAGCT
Ga16676R-LIC	GAGGAGAAGCCCGGTCAAGAGAAGTCG

Notes: FS indicated Ga16676 full sequence primers while LIC indicated ligation-independent cloning system primers.

2.3.1 Expression of Ga16676 in E. coli BL21(DE3)

The expression of recombinant Ga16676 was assessed in LB media containing 50 μg mL^{-1} of kanamycin. Induction was performed using 1 mmol L^{-1} isopropyl β-D-thiogalactopyrosidase (IPTG) after reaching an OD_{600} of ~0.6 and overproduction of recombinant protein was induced for 16 h, at 16 ℃, 20 ℃ and 37 ℃ at 180 rpm respectively. Overexpression of the Ga16676 was analyzed using 15% SDS-PAGE. The protein bands were visualized by Coomassie brilliant blue R250 staining.

2.3.2 Colony formation of Ga16676 in E. coli BL21(DE3)

The Ga16676 culture was incubated at 37 ℃ until OD_{600} of ~0.6 and induced with 1 mmol L^{-1} IPTG. An amount of 10 μL of the culture was plated on LB plates supplemented with 50 μg mL^{-1} of kanamycin and streaked using sterile inoculation loop. The plates were incubated in 4 ℃, 20 ℃ and 37 ℃ until the formation of the streaked colony was observed. *E. coli* BL21(DE3) consisting plasmid without *Ga16676* insert was used as control.

2.4 Cell growth in different temperatures

The Ga16676 culture was incubated at 37 ℃ until OD_{600} of ~0.6 and induced with 1 mmol L^{-1} IPTG. The cells were transferred to 10 ℃, 20 ℃ and 37 ℃ and incubated at 180 r min^{-1} respectively. The optical density of the culture (OD_{600}) was observed and recorded until stationary phase in the respective temperatures. *E. coli* BL21(DE3) consisting plasmid without Ga16676 insert was used as control.

2.5 Modelling Ga16676 tertiary structures

The three-dimensional Ga16676 structure was modelled to the *Bacillus caldolyticus* (PDB entry 1HZA) using the SWISS-MODEL program. The structure quality was evaluated using PROCHECK (Laskowski et al., 1993), Verify3D (Eisenberg et al., 1997) and ANOLEA (Melo et al., 1997). The superimposed Ga16676 CSD model and template and comparative analysis were performed using CHIMERA USCF (Pettersen et al., 2004).

3 Results

3.1 Ga16676 sequence analysis

The Ga16676 sequence analysis revealed that this protein-coding gene was a conserved hypothetical protein containing CSD with the size of 49 kDa. The cold shock domain was 47.83% identical to *Bacillus caldolyticus* cold shock protein, 48.57% identical to *Listeria monocytogenes* cold shock CcpB, 49.28% identical to *Geobacillus stearothermophilus* cold shock CspB and 50% identical to *Bacillus cereus* cold shock protein CspD (Table 2). The full length of Ga16667 was 445 amino acids while the cold shock domain contained 69 amino acids. The cold shock domain was found to be conserved in all prokaryotic and eukaryotic single-strand nucleic-acid binding proteins. The amino acid sequence of Ga16676 CSD showed the presence of 5 β-strands which created a common β-barrel of a cold shock domain structure. The sequence also contains RNA binding motifs that were highly conserved among living organisms. Interestingly, residue substitutions of Ile to Val were observed in RNA binding motif-1 (RNP1) while Asp to Glu in RNA binding motif-2 (RNP2) (Figure 1). The RNA binding motifs play an important role in binding both single-stranded RNA and DNA (Manival et al., 2001). A total number of seven aromatic residues (Trp-11, Phe-12, Phe-18, Phe-20, His-33, Phe-34 and Tyr-42) that interacted with RNA or ssDNA were found in the sequence. Ga16676 CSD retained all of the aromatic residues except Phe-12 and His-33. In this study, the Phe-18 and Phe-31 residues that play a critical role in nucleic acid binding activity were conserved in Ga16676 CSD (Jung et al., 2010). A total of 8 unique residue substitutions was found in Ga16676 CSD. These were substitutions of Val to Cys, Trp to Phe, Asp to Glu, Lys to Arg, Gln to Glu, Phe to Tyr, Iso to Leu and Pro to Tyr. Phylogenetic tree analysis showed Ga16676 CSD

Table 2 Sequence identity of Ga16676 cold shock domain to reference sequences

Accession (species, gene name)	Sequence identity	*E*-value	Query coverage/%
Q816H3.1 (*Bacillus cereus* ATCC 14579, CspD)	50	1e−12	15
P42016.1 (*Geobacillus stearothermophilus*, CspB)	49.28	9e−13	15
P41016.1 (*Bacillus caldolyticus*, CspB)	47.83	1e−12	15
P62169.1 (*Bacillus anthracis*, CspE)	46.27	1e−11	15
Q803L0.1 (*Danio rerio*, Lin-28A)	43.37	2e−10	17
O54310.1 (*Thermotoga maritime*, Csp)	48.48	1e−10	14

Figure 1 Multiple sequence alignment of Ga16676 CSD (amino acid position 1-114) with other cold shock domain proteins from *E. coli*, *B. caldolyticus*, *B. subtilis* and *T. maritime*. The RNA binding motifs RNP1 and RNP2 were boxed. The *β*-strands were shown using top line. The aromatic residues that interacted with RNA or ssDNA were labelled with asterisks (*) while the hydrophobic residues forming hydrophobic core were labelled with black box (▪).

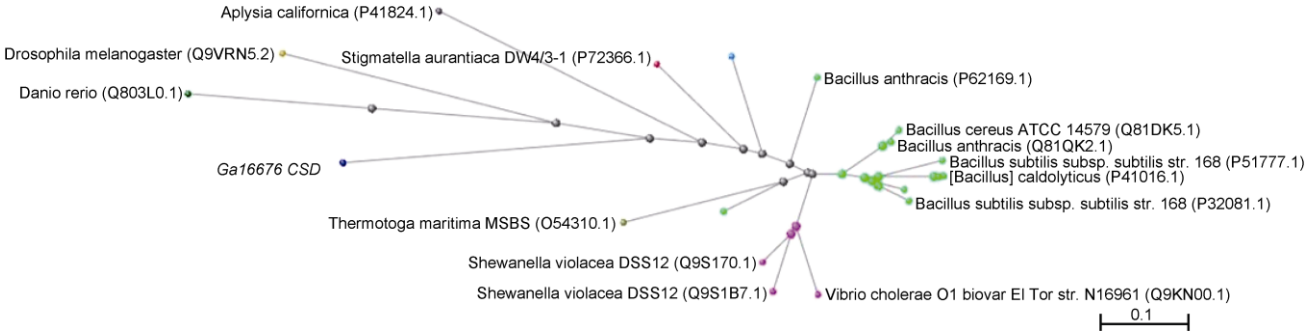


Figure 2 Phylogenetic analysis of Ga16676 CSD with other CSD from bacteria to eukarya at protein level.

had high similarity at the protein level with other CSD from bacteria to eukarya (Figure 2). Since Ga16676 CSD was unclustered with other organisms, it was possible that the sequence evolved to acquire cold adaptation strategy.

3.2 Recombinant protein expression in different temperatures and IPTG concentrations

Recombinant Ga16676 cells showed different protein expression profile when cells were grown at different temperatures of 10 ℃, 20 ℃ and 37 ℃ and IPTG concentrations of 0.5 mmol Ł$^{-1}$ and 1 mmol Ł$^{-1}$ respectively (Figure 3). At 37 ℃, recombinant Ga16676 expressed target protein in supernatants both induced with 1 or 0.5 mmol Ł$^{-1}$ IPTG. Interestingly, at 20 ℃, the recombinant Ga16676 was overexpressed in supernatants both in 1 and 0.5 mmol Ł$^{-1}$ IPTG and in pellet form in 1 mmol Ł$^{-1}$ IPTG. At 10 ℃, the recombinant Ga16676 was expressed in supernatant induced with 1 mmol Ł$^{-1}$ IPTG but none was found in pellet. Since the amount of protein loaded in each well was standardized, Ga16676 was found to be overexpressed at 20 ℃ both in the supernatant and pellet forms.

3.3 Colony formation of Ga16676 in E. coli BL21(DE3)

The ability of Ga16676 in *E. coli* BL21(DE3) to grow at different temperatures was observed by its colony formation (Figure 4). In this study, *E. coli* BL21(DE3) with empty vector was used as control. At 37 ℃ after overnight incubation, both control and Ga16676 cells grew. The control grew better compared to Ga16676 cells. Incubation at 20 ℃ for 3 d yielded the same result where both cells grew on the tested LB plate. Contrast to incubation at 4 ℃ for 8 d, only Ga16676 cells grew but not the control.

3.4 Cell growth of Ga16676 in E. coli BL21(DE3) at different temperatures

The cell growth of Ga16676 and control were observed at different temperatures of 37 ℃, 20 ℃ and 10 ℃ (Figure 5). In this study, *E. coli* BL21(DE3) with empty vector was used as control. At 37 ℃, control cells grew better compared to Ga16676. Incubation at 20 ℃ showed the same result as 37 ℃ where the control cells grew better compared to Ga16676. However, at 10 ℃ incubation, Ga16676 cells grew better compared to control. After 2 h exposure to 10 ℃, Ga16676 cells started to increase in its growth density until 11 h of incubation. However, as for the control, exposure to 10 ℃ for 2 h had slowed down the cell growth where the cells increment started to become stationary after 10 h of exposure.

3.5 3D structure analysis of GaSGT1

Ga16676 CSD domain had the highest similarity with CSD structure from *B. caldolyticus* (PDB: 1HZA) with 48.57% identity. The structure of 1HZA was used as a template where the tertiary structure of Ga16676 CSD was modelled using SWISS-MODEL. The model was evaluated using PROCHECK, Verify3D and ANOLEA. PROCHECK analysis showed that the build model had 100% amino acids in favoured and allowed regions. Furthermore, model verification using Verify3D showed that the constructed model obtained a positive score of more than 80%. Analysis using ANOLEA showed acceptable energy calculations at the atomic level in the protein model structure.

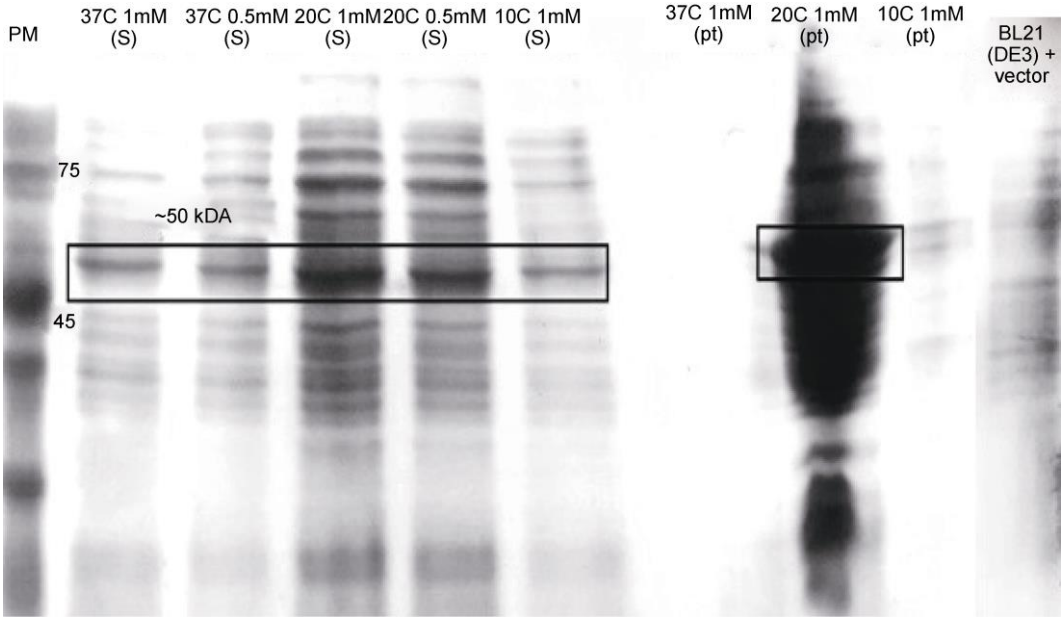
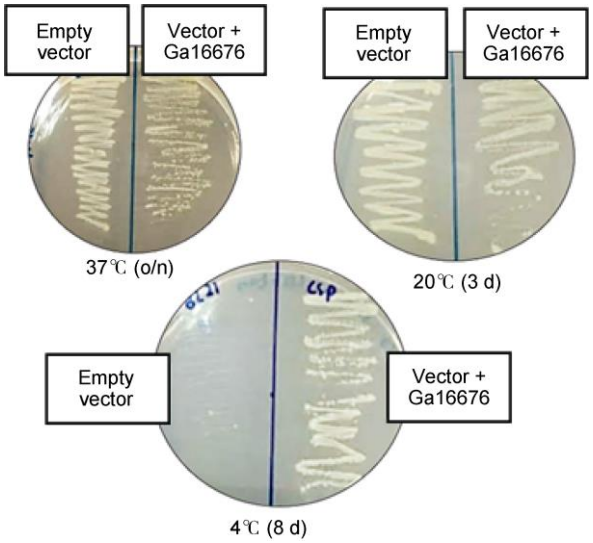


Figure 3 15% SDS-PAGE analysis of recombinant Ga16676 protein expression at different temperatures and IPTG concentration. From far left, lane-1 was loaded with protein marker (PM), lane-2 contained Ga16676 supernatant after incubation at 37 ℃ induced with 1 mmol Ł$^{-1}$ IPTG, lane-2 contained Ga16676 supernatant after incubation at 37 ℃ induced with 0.5 mmol Ł$^{-1}$ IPTG, lane-3 contained Ga16676 supernatant after incubation at 20 ℃ induced with 1 mmol Ł$^{-1}$ IPTG, lane-4 contained Ga16676 supernatant after incubation at 20 ℃ induced with 0.5 mmol Ł$^{-1}$ IPTG, lane-5 contained Ga16676 supernatant after incubation at 10 ℃ induced with 1 mmol Ł$^{-1}$ IPTG lane-6 contained Ga16676 pellet after incubation at 37 ℃ induced with 1 mmol Ł$^{-1}$ IPTG, lane-7 contained Ga16676 pellet after incubation at 20 ℃ induced with 1 mmol Ł$^{-1}$ IPTG, lane-8 contained Ga16676 pellet after incubation at 10 ℃ induced with 1 mmol Ł$^{-1}$ IPTG and lane-9 contained expression of BL21(DE3) harboring expression vector as control.



Figire 4 Colony formation of Ga16676 cells and control on LB plates supplimented with ampicillin at different incubation temperatures.

Superimposing of Ga16676 CSD model and its template gave an RMSD value of 0.168Å. Comparative analysis of Ga16676 CSD model and 1HZA showed that the constructed model contained a longer loop which connected the RNP1 and RNP2 (Figure 6a). Besides that, the residue substitutions as mentioned in sequence analysis showed an increment of aromatic interactions at the RNA binding motifs. Some selective residues substitutions like Val to Cys, Asp to Glu, Gln to Glu, Phe to Tyr, and Pro to Tyr substituted uncharged/less charged amino acids to more charged amino acids, non-polar to polar and hydrophobic to less hydrophobic amino acids. Interestingly, these substitutions increased the distance between residues interactions at the loop regions which connected the RNP1 and RNP2. Analysis of the protein surface revealed that Ga16676 CSD model had a higher solvent accessible surface area (4902.63Å) compared to its counterpart, 1HZA (4442.95Å). Moreover, there was no distinct clustering of charged amino acids on Ga16676 CSD model as compared to 1HZA. The protein surface of Ga16676 CSD model had considerable clustering of charged amino acids (Figure 6b).

4 Discussion

Most studies on Csps were done from *E. coli* and a number of homologues from psychrophiles have been reported but mostly through genomic DNA studies. However, to our knowledge, there is no functional study that has been performed on Csps from an obligate psychrophilic yeast. In order to have a better insight into the molecular mechanism

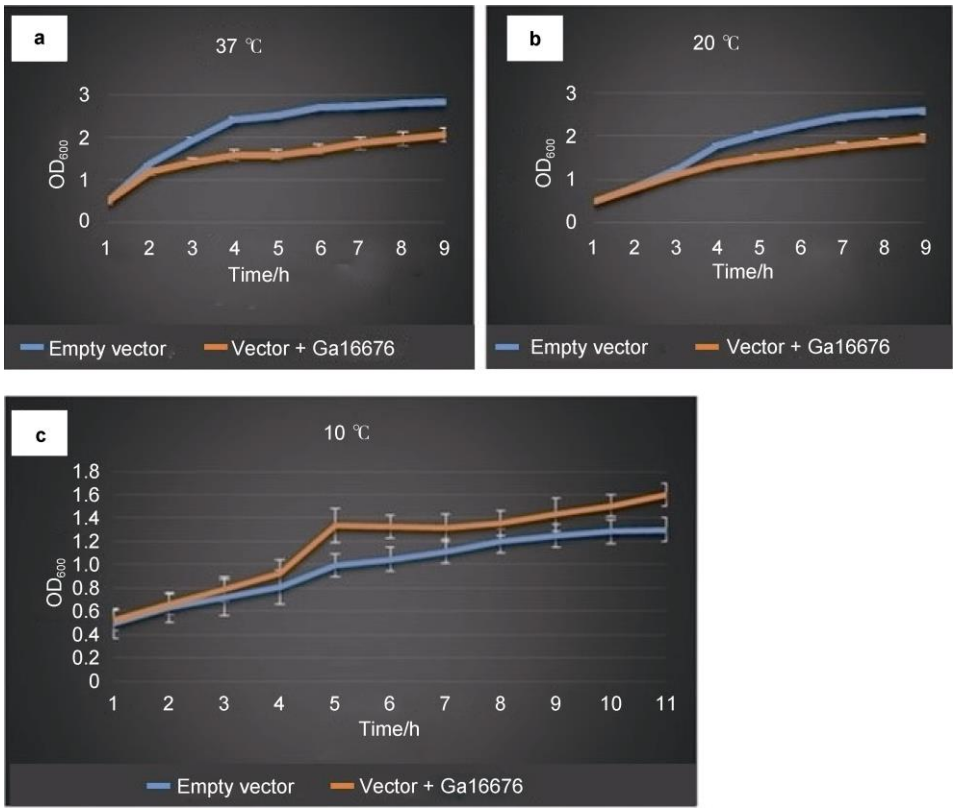


Figure 5 Cell growth comparison between Ga16676 cells and control at different temperatues of 37 ℃ (**a**), 20 ℃ (**b**) and 10 ℃ (**c**).

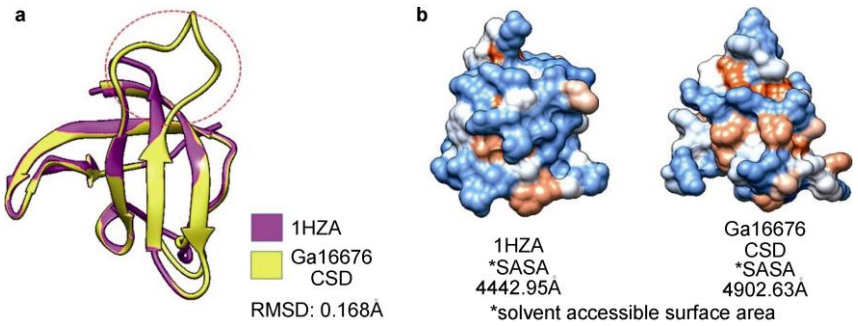


Figure 6 a, Comparative structural analysis between Ga16676 CSD model and *B. caldolyticus* CSD (1HZA). **b**, The charged surface of 1HZA and Ga16676 CSD model. The solvent accessible surface area (SASA) of Ga16676 CSD model showed that it has no distinct patches of charged amino acids compared to 1HZA protein surface.

that a psychrophilic yeast uses to overcome the temperature downshift, we studied the role of its cold shock domain containing protein. In the present study, our sequence analysis showed that Ga16676 protein from *G. antarctica* contained cold shock domain where the sequence of nucleic acid-binding proteins, RNP1 and RNP2 were highly conserved in Ga16676 (Figure 1). Residue substitutions of Ile to Val was observed in RNA binding motif-1 (RNP1) while Asp to Glu in RNA binding motif-2 (RNP2) showed interesting residue substitutions of hydrophobic to less hydrophobic and less polar to more polar residues. In addition, a total of 8 unique residue substitutions in Ga16676 C SD sequence which were substitution of Val to Cys, Trp to Phe, Asp to Glu, Lys to Arg, Gln to Glu, Phe to Tyr, Iso to Leu and Pro to Tyr showed favorable substitutions from non-polar to polar residues, less polar to more polar residues, and less charged to more charged residues. The residue substitution P67Y (Proline to Tyrosine) on the β-strand may contribute to the aromatic-aromatic interactions in strand-strand stabilization of β-sheets. The interactions between aromatics are believed to occupy a crucial role in protein folding, stability, and protein-protein or protein-ligand (Waters, 2002; Chakrabarti and Bhattacharyya, 2007). Therefore, the substitution of amino acid residues with aromatic residue in Ga16676 cold shock domain enables the stabilization of structure of a barrel

without H-bonding which allows the larger strand-strand distance to exist stably (Budyak et al., 2013). In addition, the high thermostability of *Thermotoga maritima* cold shock protein is reported due to the electrostatic interactions among the residues (Motono et al., 2008). Another report on an experimental mutational study on *Bacillus caldolyticus* cold shock protein shows that when favourable charged residues in *Thermotoga maritima* cold shock protein were grafted into *Bacillus caldolyticus* cold shock protein, the mutational effects significantly improved the thermostability of *Bacillus caldolyticus* cold shock protein without changing its two-state folding mechanism (Su et al., 2016). It was possible that the residue substitutions in Ga16676 CSD sequence which favoured hydrophilic, charged and polar residues helped to promote thermostability in the protein structure as a cold adaptation mechanism. Moreover, the phylogenetic analysis of the CSD obtained by evolutionary distance showed that the Ga16676 CSD was closely related to eukaryotes compared to the bacteria counterparts. Further, the analysis of sequence distribution on evolutionary scale also suggested sequence variation in Ga16676 CSD that possibly made the protein functionally more diverse compared to bacteria. Besides that, we also tested whether the recombinant Ga16676 protein could be expressed at 37 ℃, 20 ℃ and 10 ℃. As described previously, the recombinant Ga16676 protein was expressed at 10 ℃ in supernatant form and overexpressed in both supernatant and pellet forms at 20 ℃ after 16 h incubation. It has been reported that cold-shock inducible system produced high cold-active protein expression levels as well as reducing host toxicity (Bjerga and Williamson, 2015). Our result showed that the cold shock domain in Ga16676 may promote protein production at temperatures lower than 37 ℃. Furthermore, the colony formation of *E. coli* BL21(DE3) containing Ga16676 showed that the cells could grow at 4 ℃ after 8 d of incubation but not for *E. coli* BL21(DE3) cells with no insertion of Ga16676. Therefore, it is possible that the Ga16676 CSD involved in the acquisition of cold tolerance allowed the cells to grow at a low temperature of 4 ℃. Production of cold shock proteins is part of the mechanisms that permit low-temperature growth of microorganisms (Wouters et al., 2000). Besides that, further analysis on the cell growth showed that *E. coli* BL21(DE3) cells with Ga16676 were able to grow better at 10 ℃ compared to *E. coli* BL21(DE3) cells with no insertion of Ga16676. After 2 h of exposure to 10 ℃, the cell growth started to increase with higher doubling time until hour-11 compared to *E. coli* BL21 (DE3) cells with no insertion of Ga16676. Hence, the colony formation and cell growth experiment suggested that recombinant Ga16676 protein might play a protective role in *E. coli* BL21(DE3) against cold shock stress which allowed the cells to grow at low temperatures. A similar result was reported by Yusof et al. (2016), where a recombinant GaSGT1 protein was able to prevent protein aggregation during temperature assaults as low as 0 ℃. The importance of cold shock protein in cold adaptation was also documented in the gene expression of early cold response in *Saccharomyces cerevisiae*. When the budding yeast was exposed to a rapid temperature shift from 30 to 10 ℃, early cold response genes were induced during the first 2 h of cold treatment while late cold response genes were induced after 12 and 60 h (Schade, 2004). The 3D structure analysis was important for the elucidation of a protein structure and function. Ga16676 CSD model contained a longer loop which connected the RNP1 and RNP. The loop which connected the two most important nucleic acid-binding domains was possibly linked to higher flexibility in cold-adapted proteins. This finding has been reported in several studies of cold-adapted enzymes (Michetti et al., 2017). Selective residues substitutions which involved uncharged/less charged to a more charged amino acid in Ga16676 was also observed in *B. caldolytis* cold shock protein and its mesophilic homolog *B. subtilis* cold shock protein where this is important for protein stability (Delbrück et al., 2001). Based on the charged surface of 1HZA and Ga16676 CSD model, the value of the solvent-accessible surface area was higher than the 1HZA. These finding contrasted with the proposed theory of cold-adapted proteins contain less charged exposed residues, more exposed hydrophobic regions and less solvent accessibility to be more flexible for efficient activity at cold environments (Gianese et al., 2001; Zanphorlin et al., 2016). Hence, we proposed that accessibility to the solvents may be important to cold shock protein in *G. antarctica* in order to stabilize its structure during extreme temperature decrease. Another interesting finding was the proportional clustering of charged amino acids on Ga16676 CSD. In thermophiles, the high number of charged residues on a protein surface is important for thermostabilisation (Sælensminde et al., 2009). In *B. caldolyticus* cold shock protein, the clustering of charged amino acids is important for thermostability of the thermophilic protein (Delbrück et al., 2001). In our study, Ga16676 CSD protein surface has a considerable distribution of charged residues instead of being clustered. This finding suggested that the distribution of charged amino acids was equally important potentially for balanced stability and flexibility for protein to function in cold. Our analyses suggested that stability is important in cold shock protein structure where the rigidity may be 'softened' with flexible loops or linkers and considerable distribution of charged amino acids which allow conformational change for nucleic acid-binding.

Acknowledgments We are grateful to the Ministry of Higher Education Malaysia for funding our project (Grant no. FRG0463-2017). We thank Prof. Dr. Clemente Michael Wong Vui Ling from Biotechnology Research Institute who has assisted us with his advice, material support

and helpful discussions. We also thank those who involved in the whole genome sequencing and protein structural project of *Glaciozyma antarctica* in Malaysia Genome Institute and Universiti Kebangsaan Malaysia for their relentless efforts. We appreciate very much two anonymous reviewers, and guest editor for their helpful and constructive comments on the manuscript of this paper.

References

Bjerga G E K, Williamson A K. 2015. Cold shock induction of recombinant Arctic environmental genes. BMC Biotechnology, 15: 78. doi: 10.1186/s12896- 015-0185-1.

Budyak I L, Zhuravleva A, Gierasch L M. 2013. The role of aromatic–aromatic interactions in strand–strand stabilization of β-sheets. J Mol Biol, 425(18): 3522-3535, doi: 10.1016/j.jmb.2013.06.030.

Boo S Y, Wong C M V L, Rodrigues K F, et al. 2013. Thermal stress responses in Antarctic yeast, *Glaciozyma antarctica* PI12, characterized by real-time quantitative PCR. Polar Biol, 36(3): 381-389, doi: 10.1007/s00300-0121268-2.

Chakrabarti P, Bhattacharyya R. 2007. Geometry of non-bonded interactions involving planar groups in proteins. Prog Biophys Mol Biol, 95(1-3): 83-137, doi: 10.1016/j.pbiomolbio.2007.03.016.

Czapski T R, Trun N. 2014. Expression of csp genes in *E. coli* K-12 in defined rich and defined minimal media during normal growth, and after cold-shock. Gene, 547(1): 91-97, doi: 10.1016/j.gene.2014. 06.033.

D'Amico S, Collins T, Marx J C, et al. 2006. Psychrophilic microorganisms: challenges for life. EMBO Rep, 7(4): 385-389.

Delbrück H, Mueller U, Perl D, et al. 2001. Crystal structures of mutant forms of the *Bacillus caldolyticus* cold shock protein differing in thermal stability. J Mol Biol, 313(2): 359-369, doi: 10.1006/jmbi. 2001.5051.

Eisenberg D, Lüthy R, Bowie J U. 1997. Verify3D: Assessment of protein models with three-Dimensional profiles. Methods Enzymol, 277: 396-404, doi: 10.1016/S0076-6879(97)77022-8.

Firdaus-Raih M, Hashim N H F, Bharudin I, et al. 2018. The *Glaciozyma antarctica* genome reveals an array of systems that provide sustained responses towards temperature variations in a persistently cold habitat. PloS One, 13(1): e0189947.

Gianese G, Argos P, Pascarella S. 2001. Structural adaptation of enzymes to low temperatures. Protein Eng Del Sel, 14(3):141-148, doi:10.1093/protein/14.3.141.

Gottesman S. 2018. Chilled in translation: adapting to bacterial climate change. Mol Cell, 70(2): 193-194, doi: 10.1016/j.molcel.2018.04. 003.

Graumann P, Marahiel M A. 1996. Some like it cold: response of microorganisms to cold shock. Arch Microbiol, 166(5): 293-300.

Hashim N H F, Bharudin I, Nguong D L S, et al. 2013. Characterization of Afp1, an antifreeze protein from the psychrophilic yeast *Glaciozyma antarctica* PI12. Extremophiles, 17(1): 63-73.

Jin B, Jeong K W, Kim Y. 2014. Structure and flexibility of the thermophilic cold-shock protein of *Thermus aquaticus*. Biochem Biophys Res Commun, 451(3): 402-407.

Jones P G, Inouye M. 1994. The cold-shock response—a hot topic. Mol Microbiol, 11(5): 811-818.

Jung Y H, Yi J Y, Jung H J, et al. 2010. Overexpression of cold shock protein A of Psychromonas arctica KOPRI 22215 confers cold-resistance. Protein J, 29(2): 136-142, doi: 10.1007/s10930-010-9233-9.

Laskowski R A, MacArthur M W, Moss D S, et al. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst, 26(2): 283-291, doi: 10.1107/ S0021889892009944.

Lindquist J A, Mertens P R. 2018. Cold shock proteins: from cellular mechanisms to pathophysiology and disease. Cell Commun Signal, 16(1): 63, doi: 10.1186/s12964-018-0274-6

Manival X, Ghisolfi-Nieto L, Joseph G, et al. 2001. RNA-binding strategies common to cold-shock domain- and RNA recognition motif-containing proteins. Nucleic Acids Res, 29(11): 2223-2233, doi: 10.1093/nar/29.11.2223.

Margesin R, Neuner G, Storey K B. 2007. Cold-loving microbes, plants, and animals fundamental and applied aspects. Naturwissenschaften, 94(2): 77-99.

Melo F, Devos D, Depiereux E, et al. 1997. ANOLEA: a www server to assess protein structures. Proc Int Conf Intell Syst Mol Biol, 5: 187-190, doi: 10.1093/nar/gkh440.

Michaux C, Martini C, Shioya K, et al. 2012. CspR, a cold shock RNA-binding protein involved in the long-term survival and the virulence of Enterococcus faecalis. *J Bacteriol*, 194: 6900-6908, doi: 10.1128/JB.01673-12.

Michetti D, Brandsdal B O, Bon D, et al. 2017. A comparative study of cold- and warm-adapted Endonucleases A using sequence analyses and molecular dynamics simulations. PLoS One, 12(2): e0169586, doi: 10.1371/journal.pone.0169586.

Motono C, Gromiha M M, Kumar S. 2008. Thermodynamic and kinetic determinants of Thermotoga maritima cold shock protein stability: a structural and dynamic analysis. Proteins, 71(2): 655-669, doi: 10.1002/prot.21729

Nakagawa Y, Sakumoto N, Kaneko Y, et al. 2002. Mga2p is a putative sensor for low temperature and oxygen to induce OLE1 transcription in Saccharomyces cerevisiae. Biochem Biophys Res Commun, 291(3): 707-713.

Panadero J, Pallotti C, Rodríguez-Vargas S, et al. 2006. A downshift in temperature activates the high osmolarity glycerol (HOG) pathway, which determines freeze tolerance in Saccharomyces cerevisiae. J Biol Chem, 281(8): 4638-4645.

Perl D, Mueller U, Heinemann U, et al. 2000. Two exposed amino acid residues confer thermostability on a cold shock protein. Nat Struct Mol Biol, 7(5): 380-383, doi: 10.1038/75151.

Pettersen E F, Goddard T D, Huang C C, et al. 2004. UCSF Chimera—A visualization system for exploratory research and analysis. J Comput Chem, 25(13): 1605-1612, doi: 10.1002/jcc.20084.

Phadtare S. 2004. Recent developments in bacterial cold-shock response. Curr Issues Mol Biol, 6(2): 125-136.

Phadtare S, Alsina J, Inouye M. 1999. Cold-shock response and cold-shock proteins. Curr Opin Microbiol, 2(2): 175-180.

Phadtare S, Severinov K. 2010. RNA remodeling and gene regulation by cold shock proteins. *RNA Biol*, 7: 788-795, doi: 10.4161/rna.7.6. 13482.

Rogers A D, Johnston N M, Murphy E J, et al. 2012. Antarctic ecosystems: an extreme environment in a changing world. The Royal Society. John Wiley & Sons.

Russell N J. 1990. Cold adaptation of microorganisms. Phil Trans R Soc

Lond B Biol Sci, 326(1237): 595-611.

Sælensminde G, Halskau Ø Jr, Jonassen I. 2009. Amino acid contacts in proteins adapted to different temperatures: hydrophobic interactions and surface charges play a key role. Extremophiles, 13: 11, doi: 10.1007/s00792-008-0192-4

Schade B, Jansen G, Whiteway M, et al. 2004. Cold adaptation in budding yeast. Mol Biol Cell, 15(12): 5492-5502.

Schärer K, Stephan R, Tasara T. 2013. Cold shock proteins contribute to the regulation of listeriolysin O production in Listeria monocytogenes. *Foodborne Pathog Dis*, 10(12): 1023-1029, doi: 10.1089/fpd.2013.1562.

Su J G, Han X M, Zhao S X, et al. 2016. Impacts of the charged residues mutation S48E/N62H on the thermostability and unfolding behavior of cold shock protein: insights from molecular dynamics simulation with Gō model. J Mol Model, 22(4): 91, doi: 10.1007/s00894-016- 2958-4.

Vasina J A, Baneyx F. 1996. Recombinant protein expression at low temperatures under the transcriptional control of the major Escherichia coli cold shock promoter cspA. Appl Environ Microbiol, 62(4): 1444-1447, doi: 10.1128/aem.62.4.1444-1447.1996.

Verghese J, Abrams J, Wang Y, et al. 2012. Biology of the heat shock response and protein chaperones: budding yeast (*Saccharomyces cerevisiae*) as a model system. Microbiol Mol Biol Rev, 76(2): 115-158.

Vigh L, Maresca B, Harwood J L. 1998. Does the membrane's physical state control the expression of heat shock and other genes? Trends Biochem Sci, 23(10): 369-374.

Waters M L. 2002. Aromatic interactions in model systems. Curr Opin Chem Biol, 6(6): 736-741.

Wouters J A, Rombouts F M, Kuipers O P, et al. 2000. The role of cold-shock proteins in low-temperature adaptation of food-related bacteria. Syst Appl Microbiol, 23(2): 165-173, doi: 10.1016/S0723-2020(00)80001-6.

Yang C, Wang L, Siva V S, et al. 2012. A novel cold-regulated cold shock domain containing protein from scallop *Chlamys farreri* with nucleic acid-binding activity. PLoS One, 7(2): e32012, doi: 10.1371/journal. pone.0032012.

Yusof N A, Hashim N H F, Beddoe T, et al. 2016. Thermotolerance and molecular chaperone function of an SGT1-like protein from the psychrophilic yeast, *Glaciozyma antarctica*. Cell Stress Chaperones, 21(4): 707-715.

Zanphorlin L M, de Giuseppe P O, Honorato R V, et al. 2016. Oligomerization as a strategy for cold adaptation: Structure and dynamics of the GH1 β-glucosidase from *Exiguobacterium antarcticum* B7. Sci Rep, 6, 23776, doi: 10.1038/srep23776.