Malaysian Journal of Microbiology, Vol 19(6) Special Issue 2023, pp. 766-776
DOI: http://dx.doi.org/10.21161/mjm.230024

Malaysian Journal of Microbiology
Published by Malaysian Society for Microbiology
(In SCOPUS since 2011)

Functional analysis of conserved hypothetical proteins from the Antarctic yeast, *Glaciozyma antarctica* PI12

Makdi Masnoddin[1,2], Clemente Michael Wong Vui Ling[1] and Nur Athirah Yusof[1]*

[1]Biotechnology Research Institute, Universiti Malaysia Sabah, Jalan UMS 88400, Kota Kinabalu, Sabah, Malaysia.
[2]Preparatory Centre for Science and Technology, Universiti Malaysia Sabah, Jalan UMS 88400, Kota Kinabalu, Sabah, Malaysia.
Email: nrathirah.yusof@ums.edu.my

ABSTRACT

Aims: Recent discoveries have revealed that *Glaciozyma antarctica* PI12 has been discovered to encode numerous protein-coding genes that are crucial for thermal adaptation. However, more than 35% of the protein-coding genes for this species were identified as hypothetical proteins (HP). Nevertheless, over 35% of the protein-coding genes for this species were classified as hypothetical proteins (HP). Previous studies have documented the role of these uncharacterized proteins in the physiological regulation and cold adaptation of psychrophilic microorganisms. Thus, we aim to identify the structural features of the conserved HPs that were ideal for their function in response to temperature stress.
Methodology and results: Three conserved HPs of *G. antarctica*, designated GaHP2, GaHP3 and GaHP4, were cloned, expressed purified and their function and structure were evaluated. Functional analysis showed that these proteins maintained their activities at low temperatures below 25 °C, but at a lower reaction rate. Meanwhile, thermal unfolding assays revealed the stability of GaHP2 and GaHP4 at high temperatures (43 °C), suggesting their non-ATP-binding chaperone activity. The comparative structural analysis demonstrated that the HPs exhibited cold-adapted traits, most notably increased flexibility in their 3D structures. For GaHP2, the aromatic residues can be linked to its heat stability. GaHP4's cold shock domain implies it regulates gene transcription and translation during temperature fluctuations.
Conclusion, significance and impact of study: This study has established the structure-function relationships of the *G. antarctica* HPs and provided fundamental experimental data highlighting their importance in thermal stress response by maintaining a balance between molecular stability and structural flexibility.

Keywords: Antarctic yeast, cold adaptation, conserved hypothetical protein, quantitative PCR

INTRODUCTION

The Antarctic Sea ice's constantly cold and often hostile conditions have shaped its biological diversity, resulting in one of the most exclusive environments on Earth (Convey and Peck, 2019). *G. antarctica* PI12 is an obligate psychrophilic yeast that lives in Antarctic marine waters that was reported to remain metabolically active at sub-zero temperatures and exhibit tolerance to temperatures in a range of -12 °C to 20 °C (Boo *et al*., 2013; Soon *et al*., 2018). Despite this, the mechanism by which this psychrophilic yeast adjusts to temperature changes in its native environment, particularly the rapid cooling in Antarctica, remains unknown (Soon *et al*., 2018; Wong *et al*., 2019). To comprehend how *G. antarctica* adapts to cold, harsh environments, one must first grasp its physiology and biology in their entirety. However, the expressed sequence tag (EST) and RNAseq studies revealed 319 transcripts that are currently unique to *G. antarctica*, with 82 of these transcripts having unknown functions (Bharudin *et al*., 2018; Wong *et al*., 2019). A recent transcriptomic evaluation of *G. antarctica* PI12 indicated that 49% of differentially expressed genes (DEGs) were uncharacterized (Wong *et al*., 2019).

Therefore, more research is needed to determine the particular roles of these large proportions of poorly understood genes in enabling cells to adapt to extremely low temperatures. Numerous potential protein-coding genes remain poorly understood in terms of their presence and subsequent involvement in the survival mechanisms of *G. antarctica* (Firdaus-Raih *et al*., 2018). Previous reports indicated that 2920 (37%) of the protein-coding genes in *G. antarctica* PI12 were designated as HPs (Firdaus-Raih *et al*., 2018). The genes are either homologous to genes with unknown functions and are referred to as conserved hypothetical genes, or they have no known homologs and are termed hypothetical, non-characterized, or unknown since it is uncertain whether

*Corresponding author

ISSN (print): 1823-8262, ISSN (online): 2231-7538

they encode actual proteins. The term hypothetical proteins (HPs) refer to a group of proteins that are unrelated to any previously identified proteins (Eisenstein *et al*., 2000). Meanwhile, conserved HPs are those that have been conserved throughout several evolutionary lineages but have yet to be experimentally described (Galperin and Koonin, 2004).

The limited knowledge regarding conserved HPs presents challenges in identifying crucial signalling and stress response mechanisms within a biological system. Analysing these HPs can provide valuable insights into their functional mechanisms, thereby enhancing their utility and unlocking their full potential for biotechnological applications (Naveed *et al*., 2018). The upregulation of 62 unknown proteins at low temperatures in the Antarctic bacteria *Cryobacterium* sp. SO1 implies that these groups of unknown proteins are important for cold adaptation (Teoh *et al*., 2021). Previous research has indicated that HPs were crucial to the early response to cold and freeze stress in *G. antarctica*, but these findings have not been verified (Soon *et al*., 2018; Wong *et al*., 2019). Resolving the structures of proteins that are conserved among organisms will provide a basis for understanding their structural evolution, mechanistic features, and molecular functions (Hawkins and Kihara, 2007; Tomoike *et al*., 2009). The information can then be validated with further experiments, which will eventually help in understanding the cells' mechanisms of survival and adaptation in the harsh Antarctic environment (Turner *et al*., 2014; Convey and Peck, 2019).

The current work describes the elucidation of the function and structure of the HPs involved in thermal stress response in *G. antarctica* PI12, as well as determining the relationship between the protein molecular architecture and function. A series of bioinformatics and *in vitro* analyses of the HPs involved in thermal stress response were performed to demonstrate the proteins' cold-adapted activity at low temperatures (4 °C) and moderate temperatures (25 °C), as well as their stability and chaperone function at high temperatures (43 °C).

MATERIALS AND METHODS

Retrieval of sequence genome data

The protein-coding genes were retrieved from the transcriptome analysis of the *G. antarctica* PI12 genome (Firdaus-Raih *et al*., 2018). The transcriptome analysis of the *G. antarctica* PI12 genome provides data from RT-qPCR analysis of mRNA expression at optimal growth temperatures (15 °C) and those that simulate cold stress (0 °C and −12 °C).

Physicochemical characterization

The physicochemical properties of the HPs as raw sequence format were determined using the ProtParam tool (http://web.expasy.org/protparam/) of ExPASy (Gasteiger *et al*., 2005). The parameters include the composition of amino acids, their molecular masses and their theoretical isoelectric point (pI).

Sub-cellular localization

To assign the location of the HPs in the cell, the proteins of interest were checked for transmembrane helices via TMHMM server version 2.0 (Krogh *et al*., 2001; Sonnhammer and Krogh, 2008) and signal peptides via SignalP-5.0 Server (Almagro Armenteros *et al*., 2019).

Sequence comparison

A similarity search was performed using the Basic Local Alignment Search Tool (BLAST) against the NCBI non-redundant (nr) database (Boeckmann *et al*., 2003) to find homologous proteins from related organisms that can be predicted to have the same function as the query protein (Johnson *et al*., 2008).

Function prediction

InterProScan (https://www.ebi.ac.uk/interpro/about/interproscan) was used to predict the presence of domains and important sites in any functional protein families (Quevillon *et al*., 2005; Mitchell *et al*., 2019). Depending on what the constituent signatures represent, each InterPro entry is labelled with a type. A new entry type, homologous superfamily, has been added to the current list of types as part of InterPro's release 65.0. Proteins in a homologous superfamily have a shared evolutionary history, as evidenced by structural similarities (Mitchell *et al*., 2019).

Determination and validation of three-dimensional structures

The amino acid sequence of the selected HPs was converted from the original DNA sequence via ExPASy Translate Tool (https://web.expasy.org/translate/) and used as targets for homology modelling using the Phyre2 server (www.sbg.bio.ic.ac.uk/phyre2) (Kelley *et al*., 2015). The generated 3D models were then subjected to structure refinement using ModRefiner webserver (https://zhanggroup.org/ModRefiner/). The energy minimized structures were assessed with PROCHECK Ramachandran plots, VERIFY3D and ANOLEA-web. All the homology modelled proteins were superimposed with the template using UCSF Chimera 1.10 (Pettersen *et al*., 2004; Yang *et al*., 2012).

Protein-protein interaction analysis

As the detection of structural similarities in proteins can give elucidation of the biochemical functions, the homology modelled hypothetical proteins were subjected to analysis of protein-protein interactions using the Protein Interactions Calculator (PIC) server (http://pic.mbu.iisc.ernet.in/) (Tina *et al*., 2007). The PIC server uses the coordinate set of a protein or assembly to

 ISSN (print): 1823-8262, ISSN (online): 2231-7538

Malays. J. Microbiol. Vol 19(6) Special Issue 2023, pp. 766-776
DOI: http://dx.doi.org/10.21161/mjm.230024

Table 1: Primer design for PCR amplification of selected *G. antarctica* genes.

Transcript		Primer	% GC content	Melting temperature(°C)
gahp2	F	5'-GGT GAT GAT GAT GAC AAG ATG AAC CGC C-3'	50.0	61.2
	R	5'-GGA GAT GGG AAG TCA TTA CTA GCA CCA CC-3'	51.7	61.6
gahp3	F	5'-GGT GAT GAT GAT GAC AAG ATG GCA CCT CAG-3'	50.0	62.1
	R	5'-GGA GAT GGG AAG TCA TTA TCA CAT CCC AGC-3'	50.0	61.7
gahp4	F	5'-GGT GAT GAT GAT GAC AAG ATG GCT ACA TCA G-3'	45.2	60.2
	R	5'-GGA GAT GGG AAG TCA TTA TCA AGA GAA GTC G-3'	45.2	59.6

compute disulphide bonds, interactions between hydrophobic residues, ionic interactions and aromatic-aromatic interactions within the protein.

Cloning of the genes coding for the selected HPs

To verify the accuracy of the predicted functions for the HPs from the *G. antarctica* PI12 genome, gene cloning was conducted to prepare a system for heterologous expression of the selected HPs in *E. coli* BL21 (DE3). The *G. antarctica* PI12 cells were obtained from the Biotechnology Research Institute, Universiti Malaysia Sabah, Sabah, Malaysia. Briefly, a single colony of *G. antarctica* PI12 was inoculated in a starter culture of 5 mL yeast peptone dextrose broth (Sigma-Aldrich, USA) on a rotary incubator shaker at 200 rpm and 12 °C temperature until its mid-log phase (OD600 nm=1.5). Subsequently, the cells were harvested for total RNA extraction using GENEzol™ Reagent (Geneaid, Taiwan) according to the protocols outlined by (Bharudin *et al*., 2014). The extracted RNA was directly used for the first strand complementary deoxyribonucleic acid (cDNA) synthesis using the SuperScript™ IV First-Strand Synthesis System (Invitrogen, USA). The cDNA was then used for gene amplification and subsequent recombinant DNA construction. Genes coding for the selected HPs of *G. antarctica* PI12 HPs were identified and amplified using MyTaq™ Red Mix DNA Polymerase (Bioline, Singapore) with a specific primer design for each of the selected genes (Table 1). The amplified genes were verified using LIC sequencing forward (5'-TAATACGACTCACTATAGGG-3') and reverse (5'-GAGCGGATAACAATTTCACAGG-3') primers.

Protein expression and purification

Following the successful cloning of all the selected recombinant proteins from *G. antarctica* PI12 into the *E. coli* BL21 (DE3) host, the starter culture of the positive transformant was grown in 500 mL LB broth supplemented with 100 µg/mL ampicillin. The culture was grown in 37 °C with shaking at 200 rpm to an OD600 of ~0.5-0.6. Subsequently, the cultures were induced by adding 0.1 mM isopropyl-β-D-thiogalactsidase (IPTG) (Sigma-Aldrich, USA) and were incubated at 16 °C for 18 h with shaking at 200 rpm. Cells were then harvested by centrifugation at 4,000× *g* for 20 min and resuspended in 20 mM Tris-HCl (pH 8), 150 mM NaCl and disrupted by sonication. After sonication, the cells were centrifuged at 10,000× *g* for 20 min and then purified from the supernatant (soluble fractions) via nickel-NTA affinity chromatography. The binding buffer comprised of 50 mM Tris-HCl and 150 mM sodium chloride at pH 8.0, whereas the elution buffer comprised of 50 mM Tris-HCl, 150 mM sodium chloride and 500 mM imidazole at pH 8.0. The purified proteins were further purified by gel filtration chromatography via Superdex® HiLoad 10/300 GL gel filtration column (GE Healthcare, USA) with 20 mM Tris-HCl pH 8.0, 300 mM NaCl, 2 mM EDTA as the equilibration buffer. The purified proteins were analyzed for purity using 15% SDS-PAGE. The protein bands were visualized by Coomassie brilliant blue R250 staining (Nacalai, Japan).

Functional determination

Enzymatic ATP hydrolysis

The ATPase activity was evaluated by using malachite green reagent (Sigma-Aldrich, USA) according to the manufacturer's instruction. The assay reaction mixture was composed of 10 µg recombinant protein incubated in a 30 µL reaction volume containing 20 µL 40 mM Tris, 80 mM NaCl, 8 mM $MgAc_2$, 1 mM EDTA at pH 7.5 and 10 µL 4 mM ATP (Sigma-Aldrich, USA). The ATPase reaction mixture was incubated at room temperature or 4 °C for 30 min. After incubation, the reaction was stopped by adding 200 µL of malachite green reagent and incubated for an additional 30 min at room temperature to generate the colorimetric product. The product mixtures were loaded onto a 96-well plate and the absorbance values of colorimetric products were read using a SpectraMax spectrophotometer (Molecular Devices, San Jose, CA, USA) at 620 nm. All the samples were run in triplicate. Phosphate standard values for colorimetric detection were prepared according to the manufacturer's instructions.

Inhibition of citrate synthase thermal aggregation

Citrate synthase, a model substrate was used to test for chaperone function of proteins *in vitro*. Citrate synthase (Sigma, USA) was diluted with 40 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) KOH buffer at pH 7.5 to a final concentration of 150 nM in 1.0 mL cuvettes and heated at 43 °C with 300 nM of the purified recombinant proteins. Aggregation of citrate synthase was monitored by measuring turbidity at 320 nm

 ISSN (print): 1823-8262, ISSN (online): 2231-7538

Malays. J. Microbiol. Vol 19(6) Special Issue 2023, pp. 766-776
DOI: http://dx.doi.org/10.21161/mjm.230024

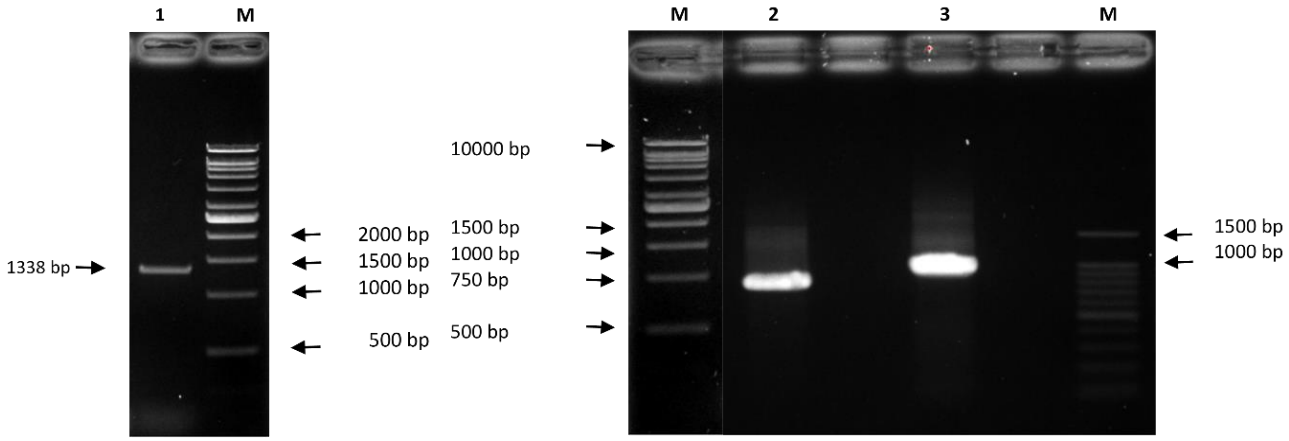

Figure 1: Agarose gel electrophoresis analysis of *G. antarctica* genes amplification at 60 °C annealing temperature. The PCR products of *gahp2* (Lane 1) genes were indicated by a single band at an approximate size of 765 bp. The PCR product of *gahp3* (Lane 2) and *gahp4* (Lane 3) were indicated by a single band at an approximate size of 942 bp and 1338 bp respectively.

Table 2: Transcriptomic analysis for the conserved HPs related to thermal stress response in *Glaciozyma antarctica* PI12.

Transcript	LAN ID	Length (bp)	Fold change (log2) 12 °C vs 0 °C	Description
gahp2	LAN_16_492	765	1.63	Predicted uncharacterized protein, conserved
gahp3	LAN_01_092	942	2.05	Predicted uncharacterized protein, conserved
gahp4	LAN_16_676	1338	1.45	Predicted uncharacterized protein, conserved

with a SPECTRA Max PLUS spectrophotometer every 5 and 10 min for 1 h. Bovine serum albumin (BSA) was used at 150 nM to evaluate the non-specific protection of citrate synthase (Chen *et al*., 2007).

RESULTS AND DISCUSSION

Analysis of the conserved hypothetical proteins from the *G. antarctica* PI12

Genome data mining of the transcriptomic analysis of *G. antarctica* PI12 revealed three coding sequences annotated to conserved HPs associated with temperature stress response (Table 2). All coding sequences reported a level of expression at low temperature (0 °C) of more than 1.5-fold (significant *p*-value $<10^{-5}$). The selection of these proteins was mostly based on the lack of information currently available, thereby presenting an opportunity for novel discoveries. The *gahp2*, *gahp3* and *gahp4* genes were amplified at the target sites of 1149 bp, 765 bp, 942 bp and 1338 bp respectively (Figure 1), followed by cloning and verification through DNA sequencing. The range of sequences considered for selection was limited to those with a length of 500 to 1500 bp. This limitation was imposed due to the tendency of shorter sequences to produce peptides of small size and longer sequences to give rise to proteins with multiple domains, which complicates the characterization procedures (Bertone, 2001; Gupta *et al*., 2016).

Prediction of physicochemical properties and subcellular localization

The physicochemical analysis in Table 3 shows that the GaH2, GaHP3 and GaHP4 proteins had molecular weight values of about 29 kDa, 35 kDa and 45 kDa, respectively. The isoelectric point (pI) values calculated vary from 4.89 to 8.84. This parameter is important in protein purification because it indicates the point at which the protein's mobility is zero in an electro-focusing device, and hence the point at which the protein will be eluted (Gasteiger *et al*., 2005; Novák and Havlíček, 2016).

The TMHMM 2.0 and SignalP 5.0 servers predicted no transmembrane helices or signal peptides for the GaH2, GaHP3 and GaHP4. As a result, it was expected that these conserved HPs would be located in the cytoplasm. Proteins in this cellular localization are involved in functional processes such as biosynthesis and transport, which contribute to the secretion of substrates or even other proteins (Nakashima and Nishikawa, 1994). The physicochemical characterization and subcellular localization analysis play an important role in the

 ISSN (print): 1823-8262, ISSN (online): 2231-7538

Malays. J. Microbiol. Vol 19(6) Special Issue 2023, pp. 766-776
DOI: http://dx.doi.org/10.21161/mjm.230024

Table 3: The physicochemical properties of *P. cryoconitis* and *G. antarctica* HPs retrieved by the Protparam tool.

Protein ID	Physicochemical properties		
	No. of amino acid	Molecular weight (Da)	Isoelectric point (pI)
GaHP2	254	28609.73	5.46
GaHP3	313	35444.18	6.08
GaHP4	445	44889.36	8.84

Table 4: The BLAST search results for the selected HPs from *G. antarctica* against the NCBI non-redundant UniProtKB/SwissProt sequences database.

Protein ID	Sequence identity	e-value	Description
GaHP2	57%	7.00E-86	ORY79343.1 Protoglobin-domain-containing protein [*Leucosporidium creatinivorum*]
GaHP3	37%	1.00E-07	EGU12204.1 putative Microtubule-associated protein [*Rhodotorula toruloides* ATCC 204091]
GaHP4	42%	1e-64	XP_016275913.1 cold-shock DNA-binding domain containing protein [*Rhodotorula toruloides* NP11]

elucidation of proteins predicted with unknown function (Hawkins and Kihara, 2007). Identifying a protein's localization in the cellular space contributes to the protein's functional characterization, primarily because protein function is normally linked to its location (Dönnes and Höglund, 2004; Naqvi *et al*., 2015; Shen *et al*., 2019).

Predicted proteins with adaptational functions in response to temperature stress

The selected HPs were annotated for homologous proteins from related organisms using the BLASTp tool against the NCBI nonredundant UniProtKB/SwissProt sequences database. Table 4 showed low similarity to the protein in the database, where GaHP3 has the lowest percentage of sequence identity, reaching as low as 37%. As a result, this suggested that the selected HPs were still incompletely characterized (Altschul *et al*., 1997; Gazi *et al*., 2020).

For the prediction of a homologous protein family and functional domains via InterProScan, the three *G. antarctica* PI12 HPs were assigned to a distinct homologous superfamily, but only GaHP3 and GaHP4 consisted of domains and putative functions that could be attributed. GaHP2 was described as part of the Globin or protoglobin homologous superfamily, as well as two GO terms for biological processes, heme binding and oxygen binding. This suggests that this protein, like haemoglobin in humans, is involved in oxygen transport. Recent research discovered that when oxygen levels were low, *G. antarctica* PI12 used nitrite as an alternative electron terminal acceptor and coupled with several other common mechanisms, these processes enabled *G. antarctica* PI12 to successfully adapt to the cold (Wong *et al*., 2019). GaHP3 has a Zn(2)-C6 fungal-type DNA-binding domain and three different GO terms, including DNA-binding transcription factor activity for biological processes and zinc ion binding and transcription control for molecular functions. This implies that this protein has a role in gene transcription regulation in the cell. GaHP4 has two domains, the cold shock domain and the cold shock protein DNA-binding domain. GaHP4's cold shock domain hinted at the protein's role in the cell's response to temperature fluctuations through regulating gene transcription and translation (Raymond-Bouchard *et al*., 2018). Similar to GaHP3, a nucleic acid binding GO term for the molecular function was also discovered. This points to the protein's functions in cell response to low temperatures as well as gene transcription and translation regulation. The proteins in this class are known to operate as transcription factors or RNA chaperones and often feature a conserved cold shock domain containing RNA recognition motifs (Hébraud and Potier, 1999; Anderson *et al*., 2006). Transcription regulation is critical for all living species, as it enables the cell to respond to intra- and extracellular signals, such as environmental stressors or food scarcity (Suvorova *et al*., 2015).

Three-dimensional structures analysis

Figure 2 and Figure 3 show the predicted structure of GaHP2 and GaHP4 proteins modelled by the Phyre2 server. The computational method for protein structure determination has advanced to the point that it can predict protein structures with atomic precision regularly, even in cases where no similar structure exists (Jumper *et al*., 2021). This provides a reliable and cost-effective alternative to the months to years of laborious effort required to determine the structure of a single protein. In the current study, the generated GaHP2 and GaHP4 3D model has a high confidence score despite a sequence identity of less than 44% similar to PDB structures, indicating that the folds were possibly correct and accurate in the core (2-4 Å). PROCHECK analysis showed that each model had 100% amino acids in favoured and allowed regions. Furthermore, model

 ISSN (print): 1823-8262, ISSN (online): 2231-7538

Malays. J. Microbiol. Vol 19(6) Special Issue 2023, pp. 766-776
DOI: http://dx.doi.org/10.21161/mjm.230024

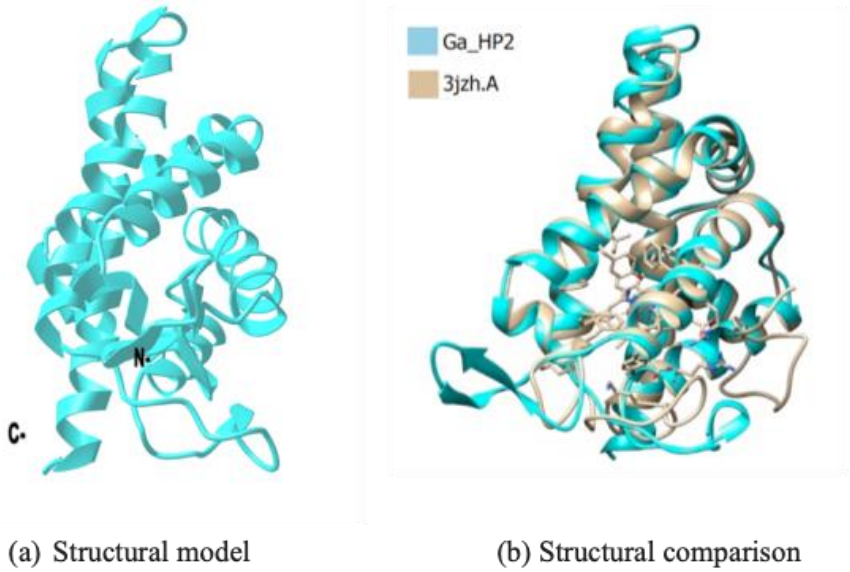

Figure 2: The three-dimensional structure analysis of GaHP2. (a) The structural model of GaHP2 was constructed using the Phyre2 server with 14% sequence identity to the structures of 3jzh as templates (Pesce *et al*., 2013). (b) Comparative structural analysis between GaHP2 model and *Methanosarcina acetivorans* Protoglobin (PDB ID: 3jzh).

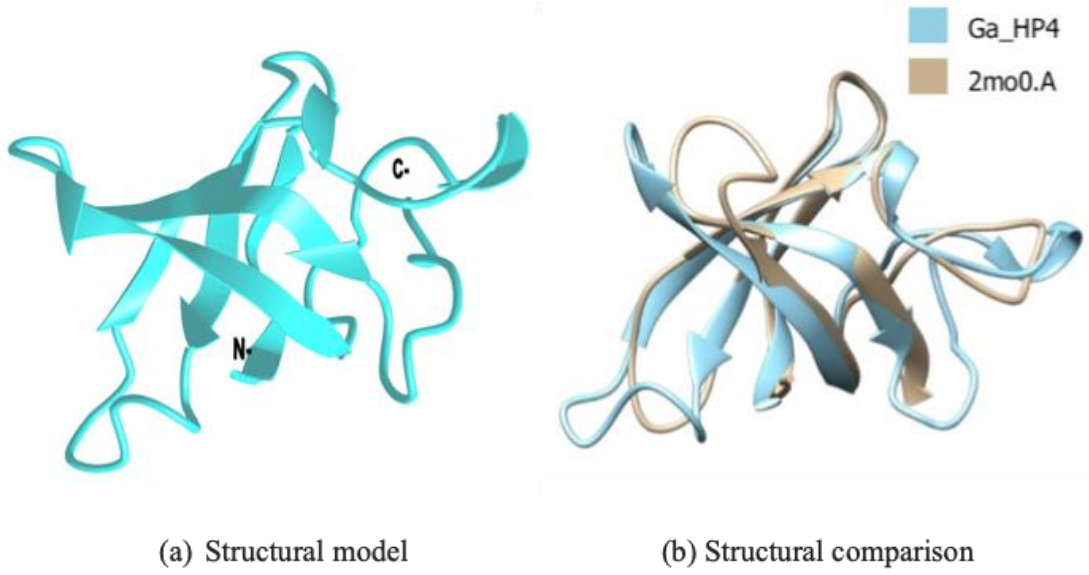

Figure 3: The three-dimensional structure analysis of GaHP4. (a) The structural model of GaHP4 was constructed using the Phyre2 server with 50% sequence identity to the structures of 2mo0 as templates (Madej *et al*., 2014). (b) Comparative structural analysis between GaHP4 model and *Thermus aquaticus* Y51MC23 cold shock DNA binding domain (PDB ID: 2mo0).

verification using Verify3D showed that the constructed models obtained a positive score of more than 92%. Analysis using ANOLEA showed acceptable energy calculations at the atomic level in the protein model structure.

GaHP2 model revealed a prominent α-helices pattern, as determined by structural comparison. By contrast, a 3D structural comparison of GaHP4 to the cold shock DNA-binding domain homolog from Thermus aquaticus indicated a prominent β-sheet. Despite this, each model was shown to have a large number of random loops secondary structures in the area between the α-helices and β-sheet. The predominance of looped regions shows that the protein structure is highly conserved (Neelamathi *et al.*, 2009). In the GaHP2 and GaHP4 models, each was shown to have multiple additional loops, therefore

 ISSN (print): 1823-8262, ISSN (online): 2231-7538

Malays. J. Microbiol. Vol 19(6) Special Issue 2023, pp. 766-776
DOI: http://dx.doi.org/10.21161/mjm.230024

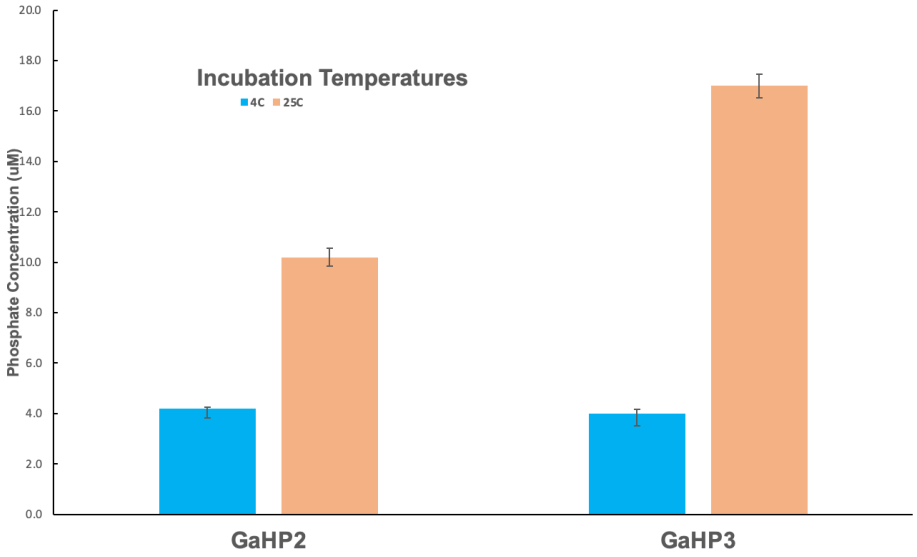

Figure 4: ATPase hydrolysis activity of GaHP2 and GaHP3 recombinant proteins at 4 °C and 25 °C. Data were representative of three biological replicates and were expressed as the mean ± SD.

indicating the cold-adapted features of the proteins (Kulakova *et al*., 2004; DasSarma *et al*., 2013; Michetti *et al*., 2017).

The superimposed GaHP2 model and its template gave a significant RMSD value of 0.5 Å (Figure 2b). Similarly, the superimposed GaHP4 model resulted in a significant RMSD value of 0.9 Å (Figure 3b). The 3D model of GaHP3 did not receive any structural homologs for comparative structure analysis. The intraprotein interactions analysis revealed that all models had lower hydrophobic interactions than their respective homologs. This suggested their structural flexibility, which enabled them to function at low temperatures (Siddiqui *et al*., 2013). Increased structural flexibility of specific residues, mainly in the active region, or the entire protein structure enables adaptation of psychrophilic enzymes by lowering the activation energy required for metabolism (Arnold *et al*., 2001). The findings are corroborated by a study of a cold-adapted protein, chitobiase from *Arthrobacter* sp. TAD20, in which increased structural flexibility resulted in increased activity at low temperatures, particularly those that significantly impede molecular motions (Lonhienne *et al*., 2001). A previous study on cold-adapted β-d-galactosidases from *Arthrobacter* sp. 32cB demonstrated that the environment's lack of free energy, caused by low temperature and high viscosity of water, can be compensated for by a higher efficiency of energy gain due to their high structural flexibility (Rutkiewicz *et al*., 2019).

Interestingly, GaHP2 showed higher aromatic-aromatic interaction than their mesophilic homologs, indicating unusual rigidity for a psychrophilic protein. Further inspection showed that the aromatic-aromatic interaction in GaHP2 was located at the heme-binding site. This suggested that the aromatic-aromatic interaction contributes to the stability of binding sites. This finding, in conjunction with a previous study comparing the cold-adapted citrate synthase from an *Arthrobacter* strain to the homologous enzyme from the hyperthermophile *Pyrococcus furiosus*, indicates that part of the cold-adapted enzyme's adaptation to low temperature may be due to increased accessibility of the active site (Gerike *et al*., 2001). Due to the high degree of conservation and relative rigidity of the catalytic site across homologous enzymes, the GaHP2 binding efficiency is consistent with the capacity of the *M. acetivorans* protoglobin (3zjh) homolog to reshape their haem distal site structure. By changing the conformation of another aromatic residue, Trp60, the aromatic residues Phe93 operate as ligand sensors and control access to the haem through the tunnel system (Klinger *et al*., 2003).

Similarly, GaHP4 showed more aromatic-aromatic interaction than the thermophilic cold shock protein (2mo0). However, the aromatic-aromatic interaction in GaHP4 involved Phe and Tyr residues as compared to Phe and Trp in the thermophiles 2mo0. Tyr has been previously described to bind strongly to both DNA and RNA, indicating a higher level of activity in GaHP4 compared to the other homologs (Anjana *et al*., 2012). It has been demonstrated that cold-adapted enzymes' high specific activity at low temperatures is a critical adaptation for compensating for the exponential reduction in chemical reaction rates when the temperature is low (Feller, 2013; Siddiqui *et al*., 2013).

ATPase assay analysis

The rate of ATP hydrolysis was determined by the amount of free phosphate produced by the proteins following a 30 min incubation at 4 °C and 25 °C (room temperature) in the presence of ATP and malachite green reagent (Figure 4). The results indicated that ATP activity values of G. antarctica PI12 recombinant proteins;

 ISSN (print): 1823-8262, ISSN (online): 2231-7538

Malays. J. Microbiol. Vol 19(6) Special Issue 2023, pp. 766-776
DOI: http://dx.doi.org/10.21161/mjm.230024

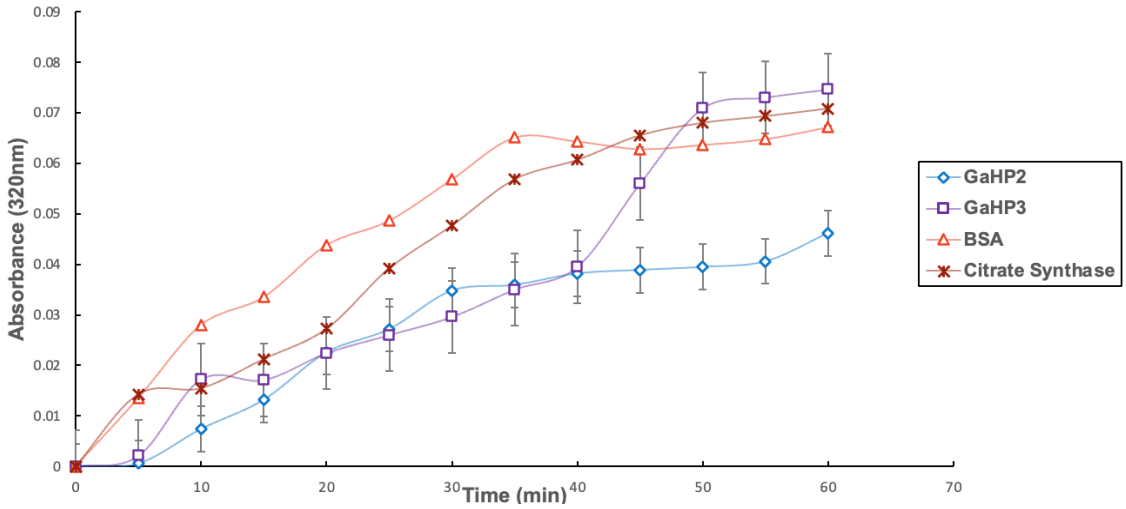

Figure 5: Heat-induced denaturation of citrate synthase. A total of 300 nM recombinant proteins of *G. antarctica* PI12 were purified and incubated at 43 °C with 150 nM citrate synthase and turbidity was measured at 320 nm. Data were representative of three biological replicates and were expressed as the mean ± SD. BSA at 300 nM was a control sample for citrate synthase non-specific protection.

GaHP2 and GaHP3 are considerably greater at room temperature than at lower temperatures (4 °C). This suggests that both proteins have an ideal operating temperature greater than 12 °C, the temperature at which they grow optimally. Similar results were reported with another cold-adapted protein from *G. antarctica* PI12, the novel recombinant chitinase, which has an optimal activity temperature of 15 °C but retains activity at temperatures ranging from 5 °C to 25 °C (Ramli *et al*., 2011). This finding is consistent with the functional annotation of GaHP3 and GaHP4, which suggests that these proteins are involved in regulating gene transcription in response to thermal stress (Raymond-Bouchard *et al*., 2018).

Thermal unfolding assay analysis

Purified GaHP2 and GaHP3 recombinant proteins were further evaluated for chaperone-like activity by measuring heat-induced aggregation of a non-native protein, citrate synthase, at 43 °C for 60 min (Figure 5). The results indicated that the addition of GaHP2 was found to reduce citrate synthase thermal unfolding. GaHP3 exhibited a significant reduction in heat aggregation during the first 40 min, but thereafter decreased significantly.

This indicated that GaHP2 and GaHP3 (for up to 40 min) have the ability to protect citrate synthase from heat-induced aggregation. This corresponded with the presence of aromatic-aromatic interactions, as demonstrated by the analysis of the 3D models for GaHP2 and GaHP4. The reversibility of some cold-adapted enzymes' thermal unfolding has been attributed to the fact that when they are unfolded, a smaller number of hydrophobic groups are exposed to the aqueous solvent, thereby preventing or limiting the irreversible aggregation process typical of more stable proteins (Marx *et al*., 2007). This indicates the chaperone activity of the proteins and their ability to recognise and bind unfolded proteins *in vitro*, hence avoiding aggregation (Maikova *et al*., 2016). A comparison of psychrophilic, mesophilic and thermophilic NAD^+ dependent DNA ligases shows that psychrophilic enzymes have higher catalytic efficiency due to enhanced flexibility of appropriate parts of the molecular structure (Georlette *et al*., 2003). Thermal denaturation appears to be reversible in certain cases, such as the α-amylase from *Pseudoalteromonas haloplanktis* (Feller, 2003; Siddiqui *et al*., 2005). Interestingly, it has been discovered that heat inactivation of cold-adapted enzymes frequently occurs prior to the structure unfolding (Collins *et al*., 2002).

CONCLUSION

The functional analysis of the three *G. antarctica* PI12 HPs (GaHP2, GaHP3 and GaHP4) indicated cold-adapted traits, most notably increased flexibility in comparison to their mesophilic or thermophilic counterparts. This is mostly owing to their lower hydrophobic interactions and extra loop attributes, which have been related to increased flexibility for cold-adapted activities. The structural dynamics of the HPs were corroborated by the functional experiments, which revealed that all HPs exhibited ATPase activity at low and moderate temperatures, demonstrating their cold-adapted property. However, GaHP2 demonstrated increased stability and protection against the thermal unfolding of citrate synthase at a high temperature of 43 °C. The existence of aromatic clusters in GaHP2 has been linked to this remarkable stability. Thus, it is perfectly feasible to elucidate that the HPs examined in this work adopt strategies to maintain a balance between molecular stability and structural flexibility.

 ISSN (print): 1823-8262, ISSN (online): 2231-7538

ACKNOWLEDGEMENTS

This research was funded by the Yayasan Penyelidikan Antartika Sultan Mizan (YPASM) grant (Ref UMS Code: GLS0032-2019) which provided the primary source of funding for the project. Additionally, we appreciate the support from the Malaysian Ministry of Higher Education (MOHE), grant number FRGS/1/2017/STG05/UMS/02/2 and registered in UMS with the code number FRGS0463-2017.

REFERENCES

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S. *et al.* (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37(4), 420-423.**

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25(17), 3389-3402.**

Anderson, K. L., Roberts, C., Disz, T., Vonstein, V., Hwang, K., Overbeek, R. *et al.* (2006). Characterization of the Staphylococcus aureus heat shock, cold shock, stringent, and SOS responses and their effects on log-phase mRNA turnover. *Journal of Bacteriology* **188(19), 6739-6756.**

Anjana, R., Kirti Vaishnavi, M., Sherlin, D., Pavan Kumar, S., Naveen, K., Sandeep Kanth, P. *et al.* (2012). Aromatic-aromatic interactions in structures of proteins and protein-DNA complexes: A study based on orientation and distance. *Bioinformation* **8(24), 1220.**

Arnold, F. H., Wintrode, P. L., Miyazaki, K. and Gershenson, A. (2001). How enzymes adapt: Lessons from directed evolution. *Trends in Biochemical Sciences* **26(2), 100-106.**

Bertone, P. (2001). SPINE: An integrated tracking database and data mining approach for identifying feasible targets in high-throughput structural proteomics. *Nucleic Acids Research* **29(13), 2884-2898.**

Bharudin, I., Abu Bakar, M. F., Hashim, N. H. F., Mat Isa, M. N., Alias, H., Firdaus-Raih, M. *et al.* (2018). Unravelling the adaptation strategies employed by *Glaciozyma antarctica* PI12 on Antarctic sea ice. *Marine Environmental Research* **137, 169-176.**

Bharudin, I., Zaki, N. Z., Bakar, F. D. A., Mahadi, N. M., Najmudin, N., Illias, R. M. *et al.* (2014). Comparison of RNA extraction methods for transcript analysis from the psychrophilic yeast, *Glaciozyma antarctica*. *Malaysian Applied Biology* **43(2), 71-79.**

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E. *et al.* (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Research* **31(1), 365-370.**

Boo, S. Y., Wong, C. M. V. L., Rodrigues, K. F., Najimudin, N., Murad, A. M. A., Mahadi, N. M. *et al.* (2013). Thermal stress responses in Antarctic yeast, *Glaciozyma antarctica* PI12, characterized by real-time quantitative PCR. *Polar Biology* **36(3), 381-389.**

Chen, T., Villeneuve, T. S., Garant, K. A., Amons, R., and MacRae, T. H. (2007). Functional characterization of artemin, a ferritin homolog synthesized in Artemia embryos during encystment and diapause. T*he FEBS Journal* **274(4), 1093-1101.**

Collins, T., Meuwis, M. A., Stals, I., Claeyssens, M., Feller, G., and Gerday, C. (2002). A novel family 8 xylanase, functional and physicochemical characterization. *Journal of Biological Chemistry* **277(38), 35133-35139.**

Convey, P. and Peck, L. S. (2019). Antarctic environmental change and biological responses. *Science Advances* **5(11), eaaz0888.**

DasSarma, S., Capes, M. D., Karan, R. and DasSarma, P. (2013). Amino acid substitutions in cold-adapted proteins from *Halorubrum lacusprofundi*, an extremely halophilic microbe from Antarctica. *PLoS ONE* **8(3), e58587.**

Dönnes, P. and Höglund, A. (2004). Predicting protein subcellular localization: Past, present, and future. *Genomics, Proteomics and Bioinformatics* **2(4), 209-215.**

Eisenstein, E., Gilliland, G. L., Herzberg, O., Moult, J., Orban, J., Poljak, R. J. *et al.* (2000). Biological function made crystal clear – Annotation of hypothetical proteins via structural genomics. *Current Opinion in Biotechnology* **11(1), 25-30.**

Feller, G. (2003). Molecular adaptations to cold in psychrophilic enzymes. *Cellular and Molecular Life Sciences* **60(4), 648-662.**

Feller, G. (2013). Psychrophilic enzymes: From folding to function and biotechnology. *Scientifica* **2013, Article ID 512840.**

Firdaus-Raih, M., Hashim, N. H. F., Bharudin, I., Bakar, M. F. A., Huang, K. K., Alias, H. *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.**

Galperin, M. Y. and Koonin, E. V. (2004). 'Conserved hypothetical' proteins: Prioritization of targets for experimental study. *Nucleic Acids Research* **32(18), 5452-5463.**

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D. *et al.* (2005). Protein identification and analysis tools on the ExPASy server. *In*: The Proteomics Protocols Handbook. Walker, J. M. (ed.). Humana Press, New York. **pp. 571-607.**

Gazi, M. A., Mahmud, S., Fahim, S. M., Md. Rezaul, I., Das, S., Mahfuz, M. *et al.* (2020). Questing functions and structures of hypothetical proteins from *Campylobacter jejuni*: A computer-aided approach. *Bioscience Reports* **40(6), BSR20193939.**

Georlette, D., Damien, B., Blaise, V., Depiereux, E., Uversky, V. N., Gerday, C. *et al.* (2003). Structural

and functional adaptations to extreme temperatures in psychrophilic, mesophilic, and thermophilic DNA ligases. *Journal of Biological Chemistry* **278(39), 37015-37023.**

Gerike, U., Danson, M. J. and Hough, D. W. (2001). Cold-active citrate synthase: Mutagenesis of active-site residues. *Protein Engineering* **14(9), 655-661.**

Gupta, D., Banerjee, S., Pailan, S. and Saha, P. (2016). *In silico* identification and characterization of a hypothetical protein of *Mycobacterium tuberculosis* EAI5 as a potential virulent factor. *Bioinformation* **12(3), 182-191.**

Hawkins, T. and Kihara, D. (2007). Function prediction of uncharacterized proteins. *Journal of Bioinformatics and Computational Biology* **5(1), 1-30.**

Hébraud, M. and Potier, P. (1999). Cold shock response and low temperature adaptation in psychrotrophic bacteria. *Journal of Molecular Microbiology and Biotechnology* **1(2), 211-219.**

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T. L. (2008). NCBI BLAST: A better web interface. *Nucleic Acids Research* **36, W5-9.**

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al*. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* **596(7873), 583-589.**

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10(6), 845-858.**

Klinger, C., Roßbach, M., Howe, R. and Kaufmann, M. (2003). Thermophile-specific proteins: The gene product of aq_1292 from *Aquifex aeolicus* is an NTPase. *BMC Biochemistry* **4, 12.**

Krogh, A., Larsson, B., Von Heijne, G. and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology* **305(3), 567-580.**

Kulakova, L., Galkin, A., Nakayama, T., Nishino, T. and Esaki, N. (2004). Cold-active esterase from *Psychrobacter* sp. Ant300: Gene cloning, characterization, and the effects of Gly→ Pro substitution near the active site on its catalytic activity and stability. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1696(1), 59-65.**

Lonhienne, T., Zoidakis, J., Vorgias, C. E., Feller, G., Gerday, C. and Bouriotis, V. (2001). Modular structure, local flexibility and cold-activity of a novel chitobiase from a psychrophilic Antarctic bacterium. *Journal of Molecular Biology* **310(2), 291-297.**

Madej, T., Lanczycki, C. J., Zhang, D., Thiessen, P. A., Geer, R. C., Marchler-Bauer, A. *et al.* (2014). MMDB and VAST+: Tracking structural similarities between macromolecular complexes. *Nucleic Acids Research* **42(D1), D297-D303.**

Maikova, A., Zalutskaya, Z., Lapina, T. and Ermilova, E. (2016). The HSP70 chaperone machines of Chlamydomonas are induced by cold stress. *Journal of Plant Physiology* **204, 85-91.**

Marx, J. C., Collins, T., D'Amico, S., Feller, G. and Gerday, C. (2007). Cold-adapted enzymes from marine Antarctic microorganisms. *Marine Biotechnology* **9(3), 293-304.**

Michetti, D., Brandsdal, B. O., Bon, D., Isaksen, G. V., Tiberti, M. and Papaleo, E. (2017). A comparative study of cold-and warm-adapted Endonucleases A using sequence analyses and molecular dynamics simulations. *PLoS ONE* **12(2), e0169586.**

Mitchell, A. L., Attwood, T. K., Babbitt, P. C., Blum, M., Bork, P., Bridge, A. *et al.* (2019). InterPro in 2019: Improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Research* **47(D1), D351-D360.**

Nakashima, H., and Nishikawa, K. (1994). Discrimination of intracellular and extracellular proteins using amino acid composition and residue-pair frequencies. *Journal of Molecular Biology* **238(1), 54-61.**

Naqvi, A. A. T., Shahbaaz, M., Ahmad, F. and Hassan, M. I. (2015). Identification of functional candidates amongst hypothetical proteins of *Treponema pallidum* ssp. Pallidum. *PLoS ONE* **10(4), e0124177.**

Naveed, M., Chaudhry, Z., Ali, Z., Amjad, M., Zulfiqar, F. and Numan, A. (2018). Annotation and curation of hypothetical proteins: Prioritizing targets for experimental study. *Advancements in Life Sciences* **5(3), 73-87.**

Neelamathi, E., Vasumathi, E., Bagyalakshmi, S. and Kannan, R. (2009). *In silico* prediction of structure and functional aspects of a hypothetical protein of *Neurospora crassa*. *Journal of Cell and Tissue Research* **9(3), 1989.**

Novák, P., and Havlíček, V. (2016). Protein extraction and precipitation. *In*: Proteomic Profiling and Analytical Chemistry. Ciborowski, P. and Silberring, J. (eds.). Elsevier, Netherlands. **pp. 52-62.**

Pesce, A., Tilleman, L., Donné, J., Aste, E., Ascenzi, P., Ciaccio, C. *et al.* (2013). Structure and haem-distal site plasticity in *Methanosarcina acetivorans* protoglobin. *PLoS ONE* **8(6), e66144.**

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. *et al.* (2004). UCSF Chimera – A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25(13), 1605-1612.**

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R. *et al.* (2005). InterProScan: Protein domains identifier. *Nucleic Acids Research* **33, 116-120.**

Ramli, A. N. M., Mahadi, N. M., Rabu, A., Murad, A. M. A., Bakar, F. D. A. and Illias, R. M. (2011). Molecular cloning, expression and biochemical characterisation of a cold-adapted novel recombinant chitinase from *Glaciozyma antarctica* PI12. *Microbial Cell Factories* **10, 94.**

Raymond-Bouchard, I., Goordial, J., Zolotarov, Y., Ronholm, J., Stromvik, M., Bakermans, C. *et al.*

(2018). Conserved genomic and amino acid traits of cold adaptation in subzero-growing Arctic permafrost bacteria. *FEMS Microbiology Ecology* **94(4), fiy023.**

Rutkiewicz, M., Bujacz, A. and Bujacz, G. (2019). Structural features of cold-adapted dimeric GH2 β-D-galactosidase from *Arthrobacter* sp. 32cB. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **1867(9), 776-786.**

Shen, Y., Tang, J. and Guo, F. (2019). Identification of protein subcellular localization via integrating evolutionary and physicochemical information into Chou's general PseAAC. *Journal of Theoretical Biology* **462, 230-239.**

Siddiqui, K. S., Poljak, A., Guilhaus, M., Feller, G., D'Amico, S., Gerday, C. *et al.* (2005). Role of disulfide bridges in the activity and stability of a cold-active α-amylase. *Journal of Bacteriology* **187(17), 6206-6212.**

Siddiqui, K. S., Williams, T. J., Wilkins, D., Yau, S., Allen, M. A., Brown, M. V. *et al.* (2013). Psychrophiles. *Annual Review of Earth and Planetary Sciences* **41(1), 87-115.**

Sonnhammer, E. L. L. and Krogh, A. (2008). A hidden Markov model for predicting transmembrane helices in protein sequence. *Sixth International Conference on Intelligent Systems for Molecular Biology* **6, 175-182.**

Soon, J., Koh, P., Vui, C. M., Wong, C. M. V. L., Najimudin, N. and Mahadi, M. (2018). Gene expression patterns of *Glaciozyma antarctica* PI12 in response to cold and freeze stress. *Polar Science* **20, 45-54.**

Suvorova, I. A., Korostelev, Y. D. and Gelfand, M. S. (2015). GntR family of bacterial transcription factors and their DNA binding motifs: Structure, positioning and co-evolution. *PLoS ONE* **10(7), e0132618.**

Teoh, C. P., Lavin, P., Lee, D. J. H., González-Aravena, M., Najimudin, N., Lee, P. C., *et al*. (2021). Genomics and transcriptomics analyses provide insights into the cold adaptation strategies of an Antarctic bacterium, Cryobacterium sp. SO1. *Polar Biology* **44(7), 1305-1319.**

Tina, K. G., Bhadra, R. and Srinivasan, N. (2007). PIC: Protein interactions calculator. *Nucleic Acids Research* **35(Suppl 2), 473-476.**

Tomoike, F., Wakamatsu, T., Nakagawa, N., Kuramitsu, S. and Masui, R. (2009). Crystal structure of the conserved hypothetical protein TTHA1606 from *Thermus thermophilus* HB8. *Proteins: Structure, Function and Bioinformatics* 76(1), 244-248.

Turner, J., Barrand, N. E., Bracegirdle, T. J., Convey, P., Hodgson, D. A., Jarvis, M. *et al.* (2014). Antarctic climate change and the environment: An update. *Polar Record* **50(3), 237-259.**

Wong, C. M. V. L., Boo, S. Y., Voo, C. L. Y., Zainuddin, N. and Najimudin, N. (2019). A comparative transcriptomic analysis provides insights into the cold-adaptation mechanisms of a psychrophilic yeast, *Glaciozyma antarctica* PI12. *Polar Biology* **42(3), 541-553.**

Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C. C., Pettersen, E. F. *et al.* (2012). UCSF Chimera, MODELLER, and IMP: An Integrated Modeling System. *Journal of Structural Biology* **179(3), 269-278.**

 ISSN (print): 1823-8262, ISSN (online): 2231-7538