



Investigating the Functional Role of Hypothetical Proteins From an Antarctic Bacterium *Pseudomonas* sp. Lz4W: Emphasis on Identifying Proteins Involved in Cold Adaptation

Johny Ijaq¹, Deepika Chandra², Malay Kumar Ray² and M. V. Jagannadham^{1*}

¹Metabolomics Facility, School of Life Sciences, University of Hyderabad, Hyderabad, India, ²CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India

OPEN ACCESS

Edited by:

Alfredo Benso,
Politecnico di Torino, Italy

Reviewed by:

Padmani Sandhu,
Chandigarh University, India
Fiona Maree Mccarthy,
University of Arizona, United States

*Correspondence:

M. V. Jagannadham
jagannadham_visiting@
uohyd.ac.in

Specialty section:

This article was submitted to
Computational Genomics,
a section of the journal
Frontiers in Genetics

Received: 30 November 2021

Accepted: 07 February 2022

Published: 11 March 2022

Citation:

Ijaq J, Chandra D, Ray MK and
Jagannadham MV (2022) Investigating
the Functional Role of Hypothetical
Proteins From an Antarctic Bacterium
Pseudomonas sp. Lz4W: Emphasis on
Identifying Proteins Involved in
Cold Adaptation.
Front. Genet. 13:825269.
doi: 10.3389/fgene.2022.825269

Exploring the molecular mechanisms behind bacterial adaptation to extreme temperatures has potential biotechnological applications. In the present study, *Pseudomonas* sp. Lz4W, a Gram-negative psychrophilic bacterium adapted to survive in Antarctica, was selected to decipher the molecular mechanism underlying the cold adaptation. Proteome analysis of the isolates grown at 4°C was performed to identify the proteins and pathways that are responsible for the adaptation. However, many proteins from the expressed proteome were found to be hypothetical proteins (HPs), whose function is unknown. Investigating the functional roles of these proteins may provide additional information in the biological understanding of the bacterial cold adaptation. Thus, our study aimed to assign functions to these HPs and understand their role at the molecular level. We used a structured *in silico* workflow combining different bioinformatics tools and databases for functional annotation. *Pseudomonas* sp. Lz4W genome (CP017432, version 1) contains 4493 genes and 4412 coding sequences (CDS), of which 743 CDS were annotated as HPs. Of these, from the proteome analysis, 61 HPs were found to be expressed consistently at the protein level. The amino acid sequences of these 61 HPs were submitted to our workflow and we could successfully assign a function to 18 HPs. Most of these proteins were predicted to be involved in biological mechanisms of cold adaptations such as peptidoglycan metabolism, cell wall organization, ATP hydrolysis, outer membrane fluidity, catalysis, and others. This study provided a better understanding of the functional significance of HPs in cold adaptation of *Pseudomonas* sp. Lz4W. Our approach emphasizes the importance of addressing the “hypothetical protein problem” for a thorough understanding of mechanisms at the cellular level, as well as, provided the assessment of integrating proteomics methods with various annotation and curation approaches to characterize hypothetical or uncharacterized protein data. The MS proteomics data generated from this study has been deposited to the ProteomeXchange through PRIDE with the dataset identifier–PXD029741.

Keywords: *Pseudomonas* sp. Lz4W, LC-MS/MS, hypothetical proteins, cold adaptation, functional annotation

INTRODUCTION

Bacterial adaptation to cold environments is a fascinating area of research and reveals a great potential for biotechnological applications. The adaptive capability of bacteria to low temperatures depends on regulating many integrated signalling and molecular mechanisms which enable bacteria to sense the environmental changes, process the signals, and orchestrate the physiological changes accordingly (Moradali et al., 2017). Studies on cold adaptation of Antarctic bacteria revealed many of their survival strategies like the ability to control membrane fluidity, catalyze biochemical reactions at frozen temperatures, sense the temperature cues, regulate transcription and translation machinery, and modify protein conformation (Ray et al., 1994; Chintalapati, et al., 2004; Shivaji and Prakash, 2010; Prakash et al., 2010). Modifications in the cell envelop components such as peptidoglycan, lipopolysaccharides (LPS), and exopolysaccharides appear to be an important strategy for cold adaptation (Tribelli and Lopez, 2018). Additionally, cold adaptation in psychrophiles is attributed to membrane proteins, chaperones, RNA helicases, cold shock proteins, carotenoids, short- and branched-chain fatty acids, and antioxidative enzymes. Furthermore, an increase in the expression of proteins related to replication, transport, and folding was reported as a mechanism of cold adaptation (De Maayer et al., 2014; Bargiela et al., 2020). Likewise, changes in the catalytic activity of the enzymes, changes in central metabolic activities, post-translational modifications, synthesis of storage polymers, like polyhydroxyalkanoates (PHAs) were reported to be advantageous for bacterial adaptation to cold temperatures (Tribelli and Lopez, 2018; Bargiela et al., 2020).

In our laboratory, we have been using *Pseudomonas* sp. Lz4W as a model system to understand the cold adaptation of bacteria (Ray et al., 1998). *Pseudomonas* sp. Lz4W is a gram-negative bacterium belonging to the γ -proteobacteria class. This strain was isolated from the Schirmacher Oasis, Antarctica (70°45'12"S and 11°46'E) (Shivaji et al., 1989). It is capable of surviving zero and sub-zero temperatures. Under laboratory conditions, it was observed to be capable of growing up to 30°C, whereas 22°C being the optimum growth temperature (Kannan et al., 1998). Pandiyan and MK Ray reported the first draft genome sequence of this bacterium to decipher the genomic basis of its psychrophilicity (Pandiyan and Ray, 2013). Over the past decade, extensive research studies have focused at deciphering the underlying molecular mechanisms of *Pseudomonas* sp. Lz4W adaptability to cold temperatures (Chattopadhyay and Jagannadham 2001; Kumar et al., 2002; Jagannadham and Chowdhury, 2012).

The genetic architecture of an organism determines the ability of an organism to respond to environmental variations. To address this, the genome and the proteome of *Pseudomonas* sp. Lz4W has been investigated by previous studies. However, there are few coding sequences in the genome that are predicted to code for proteins but having no evidence of expression at the translation level. Proteins encoded by such sequences were assigned as "hypothetical" or "uncharacterized" (Desler et al., 2009; Dall'Agnol et al., 2014). These hypothetical proteins were found to constitute a major portion of microbial genomes. From recent studies, it was evident that functional characterization of

hypothetical proteins (HPs) could provide additional understanding of the physiological architecture of the cell (Rabbi et al., 2021; Araújo et al., 2020; Pranavathiyani et al., 2020; Sen and Verma, 2020; da Costa et al., 2018; Naqvi et al., 2015; Shahbaaz et al., 2013). Genomic analysis of *Pseudomonas* sp. Lz4W also predicted few ORFs as hypothetical proteins. Our previous proteomic studies on membrane proteins of *Pseudomonas* sp. Lz4W also revealed the expression of a few HPs (Jagannadham et al., 2011; Jagannadham and Chowdhury, 2012). But functional studies on these hypothetical proteins are lacking. Hence, studying these functionally unknown sequences could provide additional insight into potential mechanisms governing cold adaptation of *Pseudomonas* sp. Lz4W.

HPs were found to be important in connecting the missing links in genomic and proteomic information. From our previous study (Ijaq et al., 2020) using mass spectrometry (MS)-based proteomics, we provided evidence of expression for HPs at the protein level. Furthermore, these days a wide range of computational biology tools are available to unravel the functional roles of the identified hypothetical protein sequences (Ijaq et al., 2015; Ijaq et al., 2019). Keeping this in mind, for this study, we set an objective to identify and characterize the hypothetical sequences from the genome of *Pseudomonas* sp. Lz4W. Proteomics-based methods combined with annotation-based workflow were used to accomplish the functional characterization of HPs concerning their role in cold adaptation. This annotation helped in assigning various classes to the HPs such as membrane proteins, lipoproteins, transporters, enzymes, chaperones and virulence factors. We believe that this study will give insights regarding the functional roles of HPs present in *Pseudomonas* sp. Lz4W. Further experimental validation of these identified proteins may eventually help in developing new biotechnological applications.

MATERIALS AND METHODS

Bacterial Culture and Growth Conditions

Pseudomonas sp. Lz4W was isolated from the Schirmacher Oasis region of Antarctica (70°45'12" S and 11°46' E) (Shivaji et al., 1989). The strain was cultured using Antarctic bacterium media (ABM) comprising of peptone 5%, yeast extract 2%, and agar 1.5% (for solid media). Bacterial cells were grown in three different batches at 4°C and the cells were harvested in the early log phase. The bacterial cultures were then centrifuged (8000 rpm for 20 min) and the cell pellets were further processed to prepare the sample repeats for proteomic analysis.

Protein Isolation

Cell pellets were resuspended in lysis buffer (20 mM Tris-Cl pH 8, 150 mM NaCl, and 2 mM EDTA). PMSF (Phenylmethylsulfonyl fluoride) and protease inhibitor cocktail were added to the suspension. Resuspended cells were then lysed with 30 min lysozyme treatment at 4°C. Cell lysis was furthered by sonication (3-5 cycles of 30 s ON and 30 s OFF, at a medium frequency) using Diagenode's Bioruptor Plus sonication system. Cell debris and unbroken cells were removed as pellets by centrifugation at

10,000 rpm at 4°C for 10 min. The supernatant was collected and this constitutes the whole-cell extract. This whole-cell extract was treated with dithiothreitol (1 mM) and iodoacetamide (Gundry et al., 2009) to reduce the disulfide bonds of proteins. The supernatant was then further fractionated by ultracentrifugation (100,000Xg for 1 h) to identify the low abundant proteins. This step resulted in two fractions—supernatant as soluble protein fraction of whole-cell extract and pellet as an insoluble fraction (Ames, 1974).

Proteins present in the soluble fraction were precipitated following the acetone precipitation method (Nejadi et al., 2014) with small modifications. The soluble fraction was mixed with chilled acetone in a 1:3 ratio and kept overnight at -20°C. Proteins precipitated were then separated as pellet using centrifugation (13,000 rpm at 4°C for 20 min). Pellet was air-dried and subsequently resuspended in 1X Laemmli buffer. Insoluble fraction pellet obtained from the previous step was also resuspended in 1X Laemmli buffer.

SDS PAGE and In-Gel Digestion

Soluble and insoluble fractions were collected in equal amounts (40 µg each) and subjected to SDS-PAGE in two different lanes. The gel was stained with Coomassie Blue. A small portion of gel with prominent bands was excised and used for in-gel digestion with trypsin as described earlier with slight modifications (Jagannadham 2008). Extracted peptides were desalted by using C18 resin-based Zip-Tips obtained from Millipore.

Mass Spectrometry Analysis and Data Acquisition

LC-MS/MS analysis of tryptic digested peptides was carried out using Q Exactive HF (Thermo Fisher Scientific). Digested peptides were reconstituted in 0.1% formic acid and separated by nano-LC using an Easy-nLC 1,200 (Thermo Fisher Scientific, United States). An equal amount of peptide samples (400 ng each) were loaded into NanoDrop 2000/2000c spectrometer (Thermo Fisher Scientific) with absorbance at 205 nm for quantification. PepMap™ RSLC C18 column (75 µm × 15 cm length, 3 µm particle size, and 100 Å pore size, Thermo Fisher Scientific, United States) was used to separate the tryptic digested peptides. A binary solvent system was used with the mobile phases comprising (A) 5% acetonitrile with 0.1% (vol/vol) formic acid and (B) 95% acetonitrile with 0.1% (vol/vol) formic acid. Peptides were eluted with a linear gradient of Buffer A to Buffer B with a flow rate of 300 nL/min over a 90 min gradient run. Separations are performed by changing the percentage of B. The binary gradient elution was as follows: 0 min, 0% B; 79 min, 90% B; 90 min, 0% B. For MS scan, mass range m/z 400–1750 was used. Dynamic exclusion interval of 15 s was selected with exclude isotope mode ON. For MS/MS, normalized collision energy of 28% was used for collision-induced dissociation (CID) activation. To retrieve MS/MS spectra, top 10 peaks was selected as scan event.

Protein Identification

Raw files generated by Q Exactive HF were submitted to proteome discoverer software (version 1.4, Thermo Fischer).

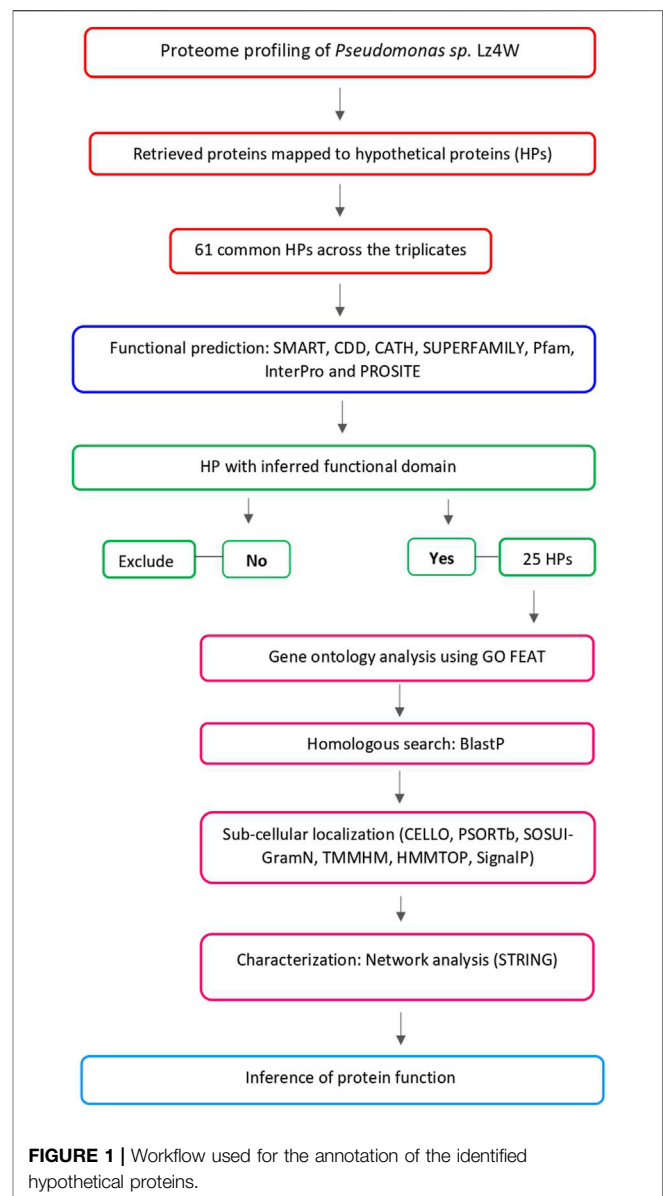


FIGURE 1 | Workflow used for the annotation of the identified hypothetical proteins.

SEQUEST search engine was used to search the mass spectra of peptides against the sequence database. For the identification of the proteins, all MS/MS spectra of peptides were searched against *Pseudomonas* sp. Lz4W database (GenBank Accession: CP017432.1, BioProject: PRJNA170013) following the criteria: precursor mass tolerance/peptide tolerance, ± 10 ppm; fragment mass tolerance/MS/MS tolerance, ± 0.6 Da; spectra selection range of precursor mass, 350–5,000 Da; dynamic modification, oxidation of methionine (15.99 Da); and static modification, carbamidomethylation of cysteine (57.02 Da). Two missed cleavages were set for the identification of proteins. False discovery rate (FDR) was set to 1% for the identification of peptides with high confidence.

Proteome profiling of triplicate cultures resulted in three sets of protein data - Set A, Set B, and Set C, identifying 949, 919, and 872 proteins respectively from each set. Initially, we searched for

the common proteins across the three sets. From the common proteins, proteins annotated as 'hypothetical proteins' were retrieved and further investigated for their functional roles.

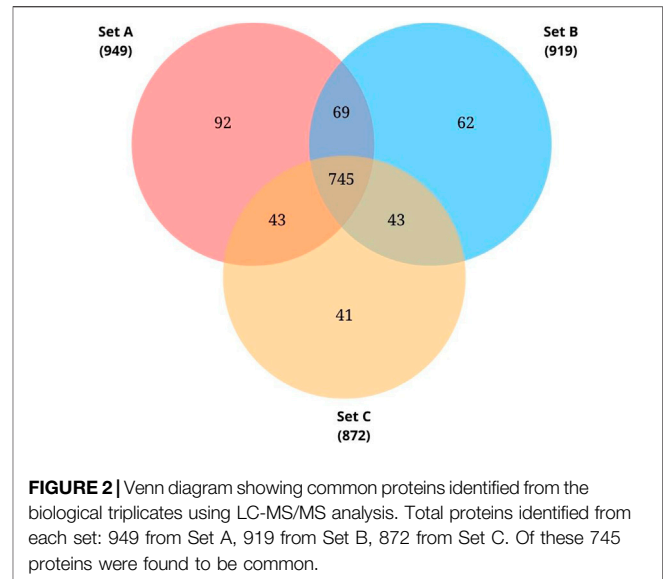
Sequence Retrieval

Pseudomonas sp. Lz4W genome was used for retrieving the sequences. *Pseudomonas sp.* Lz4W genome was sequenced and assembled in our laboratory (Pandiyani and Ray, 2013). Its genome consists of 5,018,820 bp corresponding to 4493 genes, of which 4,393 genes code for the proteins (coding sequences). From these, a total of 743 coding sequences (CDS) were annotated as hypothetical proteins. Genomic data was retrieved from the National Center for Biotechnology Information database (<https://www.ncbi.nlm.nih.gov/nucore/>) under accession number CP017432. The sequences of the expressed HPs were retrieved in FASTA format from NCBI (<https://www.ncbi.nlm.nih.gov/>) using their primary accession numbers. Corresponding UniProt entries were crosschecked in the Uniprot database (UniProt Proteome ID: UP000011925) to retrieve the UniProt accession numbers.

Functional Annotation

Hypothetical proteins that were identified at the translation level were subjected to functional annotation to uncover their functional roles. An overview of the annotation workflow is presented in **Figure 1**. First, they were provided with putative functional annotation based on the functional domains and motifs using various publicly available bioinformatics tools. Simple Modular Architecture Research Tool (SMART) (version 9.0) was used to predict the function of a protein based on the domain architecture. It searches proteome information from Swiss-Prot, SP-TrEMBL, and stable Ensembl for a similar domain (Letunic et al., 2012). CATH (version 4.3) (Sillitoe et al., 2015) was used to classify the domains within the structural hierarchy. Conserved Domain Database (CDD) (version 3.19) (Lu et al., 2020) was used to explore conserved domains that are associated with molecular function. Pfam (version 35.0) (Finn et al., 2014), SUPERFAMILY (version 2.0) (Gough et al., 2001), InterPro (version 87.0) (Blum, et al., 2021), and PROSITE (release 2021_04) (Sigrist et al., 2013) were used to classify the HPs into functional families to predict the function based on similarity. InterProScan (version 5.45–87.0) was used as the underlying software to search the hypothetical protein sequences against InterPro's protein family signatures. (Blum, et al., 2021). Functional domain was inferred to each hypothetical protein if the similar domain was found consistently across the databases used. For all the databases, we used default parameters.

Functional annotations were further strengthened by matching the protein functional domains, motifs, and families along with GO terms. GO FEAT tool was used for GO based functional annotation of HP sequences (Araujo et al., 2018). E-value of $1e-03$ was selected to ensure significant matches. Subsequently, sequence similarity search was performed to identify if a hypothetical protein sequence is homologous to sequences that are already available in existing databases.



BlastP (Altschul et al., 1990; Altschul et al., 1997) was used against NCBI nonredundant (nr) database, and hits with an identity $>90\%$ were considered.

Sub-Cellular Localization

Knowledge about the subcellular localization of a protein can provide an estimate of a protein's function at the cellular level. PSORTb, CELLO and SOSUI-GramN were used to assign the sub-cellular location of the HPs. PSORTb (version 3.0.3) provides prediction of subcellular localization of bacterial and archaeal proteins. (Gardy et al., 2003). CELLO (version 2.5) is a two level support vector machine (SVM) based system used for the prediction of localization of proteins from both bacteria and eukaryotes. (Yu et al., 2006). SOSUI-GramN system is specific to predict the extracellular and membrane protein localization in gram negative bacteria. (Imai et al., 2008). SignalP (version 5.0) was used to predict the presence of signal peptide cleavage sites (Petersen et al., 2011). TMHMM (version 2.0) (Krogh et al., 2001) and HMMTOP (version 2.0) (Tusnády and Simon, 2001) was used to predict the transmembrane helices in proteins thus identifying the membrane proteins.

Protein-Protein Interaction Analysis

STRING (version 11.5) was used for protein-protein interaction analysis of the HPs. It provides functional associations between proteins based on the information from different sources like scientific literature, experimental evidence, predictions from co-expression, genome context, curated pathway databases and calculates the final 'confidence score' for each association (Szklarczyk et al., 2021). For this study, a confidence score above 0.7 was considered to ensure the reliability of the association. *Pseudomonas sp.* Lz4W reference genome was used to generate the interaction networks. Interactions were further visualized using Cytoscape (version 3.9.0) (Shannon

TABLE 1 | Hypothetical proteins identified from *Pseudomonas* sp. Lz4W along with their functional analysis.

S.No	GenBank accession	UniProt ID	Putative function	Category	
Proteins assigned with functional domains					
1	AUB73822.1	A0A2H4VVI5	Endolytic peptidoglycan transglycosylase (RlpA)	Enzymes	
2	AUB74731.1	A0A2H4VY66	Thiosulfate sulfurtransferase		
3	AUB74781.1	A0A2H4VYB4	ATPase activity		
4	AUB74810.1	A0A2H4VYB1	Pyrimidine/purine nucleoside phosphorylase		
5	AUB77590.1	A0A2H4W6A7	Glycosyl transferases group 1		
6	AUB75791.1	A0A2H4W154	Ligase, Condensation domain		
7	AUB75792.1	A0A2H4W178	Ligase, Condensation domain		
8	AUB75991.1	A0A2H4W1Q4	Sulfatase		
9	AUB76494.1	A0A2H4W376	Flavinator of succinate dehydrogenase		
10	AUB76830.1	A0A2H4W456	Aminoglycoside phosphotransferase (APH)		
Membrane proteins					
11	AUB74223.1	A0A2H4VWR7	Outer membrane protein	Membrane proteins	
12	AUB76580.1	A0A2H4W3E2	Outer membrane protein (OmpA)		
13	AUB76959.1	A0A2H4W4H7	Outer membrane protein (OmpA)		
14	AUB77315.1	A0A2H4W5I1	Outer membrane porin (OprD)		
Lipoproteins					
15	AUB73831.1	A0A2H4VVK8	Lipopolysaccharide-assembly lipoprotein (LptE)	Lipoproteins	
16	AUB74026.1	A0A2H4VW39	Lipoprotein		
17	AUB74624.1	A0A2H4VXV2	Lipoprotein with Alanine-zipper		
Transport proteins					
18	AUB75187.1	A0A2H4VZE1	Intermembrane phospholipid transport system lipoprotein (MlaA)	Transport proteins	
19	AUB75954.1	A0A2H4W1N5	Copper resistance protein C (CopC), Copper transport protein		
20	AUB77446.1	A0A2H4W5Z0	Tim 44 like domain		
Chaperones					
21	AUB76897.1	A0A2H4W4C7	HSP70 family of molecular chaperone	Chaperones	
22	AUB76544.1	A0A2H4W3C6	Outer membrane chaperone (skp)		
Cellular process					
23	AUB73565.1	A0A2H4VUU7	Ycel-like domain	Cellular process	
Virulence factors					
24	AUB73322.1	A0A2H4VU41	Virulence factor	Virulence factors	
25	AUB74265.1	A0A2H4VWT9	Putative quorum-sensing-regulated virulence factor		
Proteins assigned with domain of unknown function					
26	AUB73312.1	A0A2H4VU37	Domain of unknown function	Domain of Unknown Function (DUF)	
27	AUB73583.1	A0A2H4VUV2	Domain of unknown function		
28	AUB73639.1	A0A2H4VW22	Domain of unknown function		
29	AUB73986.1	A0A2H4VW02	Domain of unknown function		
30	AUB74271.1	A0A2H4VWW5	Domain of unknown function		
31	AUB74729.1	A0A2H4VY42	Domain of unknown function		
32	AUB74842.1	A0A2H4VYJ4	Domain of unknown function		
33	AUB77556.1	A0A2H4W6A6	Domain of unknown function		
34	AUB75101.1	A0A2H4VZ73	Domain of unknown function		
35	AUB75536.1	A0A2H4W0E2	Domain of unknown function		
36	AUB75846.1	A0A2H4W1A8	Domain of unknown function		
37	AUB76324.1	A0A2H4W2Q7	Domain of unknown function		
38	AUB76750.1	A0A2H4W3Y2	Domain of unknown function		
39	AUB77015.1	A0A2H4W4N4	Domain of unknown function		
40	AUB77395.1	A0A2H4W5R3	Domain of unknown function		
41	AUB75938.1	A0A2H4W1Q3	Domain of unknown function		
Proteins identified with no functional domain					
42	AUB73328.1	A0A2H4VU50	None		Uncharacterized
43	AUB73856.1	A0A2H4VVM1	None		
44	AUB73929.1	A0A2H4VWW4	None		
45	AUB73991.1	A0A2H4VW15	None		
46	AUB73994.1	A0A2H4VW21	None		
47	AUB74225.1	A0A2H4VWQ3	None		
48	AUB74336.1	A0A2H4VWZ7	None		
49	AUB74383.1	A0A2H4VX83	None		
50	AUB74509.1	A0A2H4VXH3	None		
51	AUB74552.1	A0A2H4VXN7	None		
52	AUB74606.1	A0A2H4VXS5	None		

(Continued on following page)

TABLE 1 | (Continued) Hypothetical proteins identified from *Pseudomonas* sp. Lz4W along with their functional analysis.

S.No	GenBank accession	UniProt ID	Putative function	Category
53	AUB74636.1	A0A2H4VXV4	None	
54	AUB75651.1	A0A2H4W0S2	None	
55	AUB75652.1	A0A2H4W0U3	None	
56	AUB75757.1	A0A2H4W111	None	
57	AUB75993.1	A0A2H4W1U9	None	
58	AUB76598.1	A0A2H4W3K1	None	
59	AUB76896.1	A0A2H4W4C0	None	
60	AUB77182.1	A0A2H4W552	None	
61	AUB77347.1	A0A2H4W5N7	None	

et al., 2003). All the bioinformatics tools and databases used in this study are listed in **Supplementary Table S1**.

RESULTS AND DISCUSSION

The complete genome sequence for the species *Pseudomonas* sp. Lz4W was deposited in the NCBI database in 2013 by Ray and colleagues (Pandiyan and Ray, 2013). As on November 2021, a total of 4393 protein-coding genes were predicted on this genome. Among these, 743 CDS (16.9%) were labelled as hypothetical proteins whose function is unknown. In the present study, we tried investigating the protein expression of *Pseudomonas* sp. Lz4W during cold adaptation to check whether any of these CDS labelled as HPs are expressed at the protein level. Proteome profiling of the triplicate cultures resulted in 949, 919, and 872 proteins respectively (**Supplementary Table S2**). 745 common proteins were identified across the three sets. From these 745 proteins, 61 proteins were found to be annotated as “hypothetical proteins” (**Figure 2**). A complete list of these 61 HPs and their MS data is presented in **Supplementary Table S3**. Uniprot annotation of the corresponding entries is shown in **Supplementary Table S4**. From the UniProt information, it was found that all these 61 HPs are either predicted proteins or proteins inferred from homology. From this proteomic analysis, identifying these proteins at the translational level provides evidence for their expression in the tested growth conditions. The sequences of these proteins were extracted from NCBI and used for further functional characterization using available bioinformatics tools and resources.

Functional Annotation Using Domain Analysis

Initially, we looked for the functional domains in the sequences. Domains are functional and structural units of proteins that are evolutionarily conserved across each protein family or superfamily (Ponting and Russell, 2002). As they are highly conserved within a functional protein family, looking for a presence of a functional domain in the hypothetical protein sequences could provide insights into their function (Eisenstein et al., 2000; Sen and Verma, 2020). Sequences of all the 61 HPs were extensively analyzed for the presence of functional domains using SMART, CDD (CD-Search tool),

CATH, Pfam, SUPERFAMILY, InterPro, and PROSITE. The predicted results were shown in **Supplementary Table S5**. Out of 61 HPs, 25 proteins were assigned with functional domains. Putative function was assigned to these proteins only if a similar functional domain is identified across the different tools and databases used. For the remaining 36 proteins, no specific conserved domains were found for 20 proteins and the remaining 16 proteins were found to have a DUF (domain of unknown function) domain, whose function is not yet characterized (**Table 1**).

Based on the identified putative function, the 25 HPs with known functional domains were categorized into different functional groups namely membrane proteins, lipoproteins, cellular process, transport proteins, chaperones and virulence factors (**Figure 3**).

Membrane Proteins

In our analysis, a few HPs, AUB74223.1, AUB76580.1, AUB76959.1, and AUB77315.1 were predicted to be associated with outer membrane (OM). Membrane composition, structure, and proteins play a significant role in cold adaptation. From our previous studies, it was demonstrated that the membranes and membrane components play a crucial role in the cold adaptation of bacteria (Chattopadhyay and Jagannadham 2001; Kumar et al., 2002; Jagannadham and Chowdhury, 2012). Functional genomics studies of psychrophilic bacteria have shown the increased abundance of outer membrane proteins during cold adaptation (Ting et al., 2010; Frank et al., 2011). HPs AUB74223.1, AUB76580.1 and AUB76959.1 were found to contain OmpA (outer membrane protein A) like domain. Proteins with OmpA like domain play a vital role in maintaining the outer membrane structure and stability (Wang, 2002), and may serve as a porin facilitating the transport of small solutes (Smith et al., 2007; Confer and Ayalew, 2013). Hypothetical protein (HP) AUB77315.1 was found to contain the OprD domain belonging to the outer membrane porin family. OprD is one of the major porin proteins present in the outer membrane of the genus '*Pseudomonas*'. Members of the OprD family are involved in the specific uptake of various small molecules ranging from amino acids to peptides, usually below 200 Da (Li et al., 2012; Chevalier et al., 2017).

Lipoproteins

Hypothetical proteins AUB73831.1, AUB74026.1, and AUB74624.1 were annotated as lipoproteins. Bacterial lipoproteins are membrane-associated proteins with

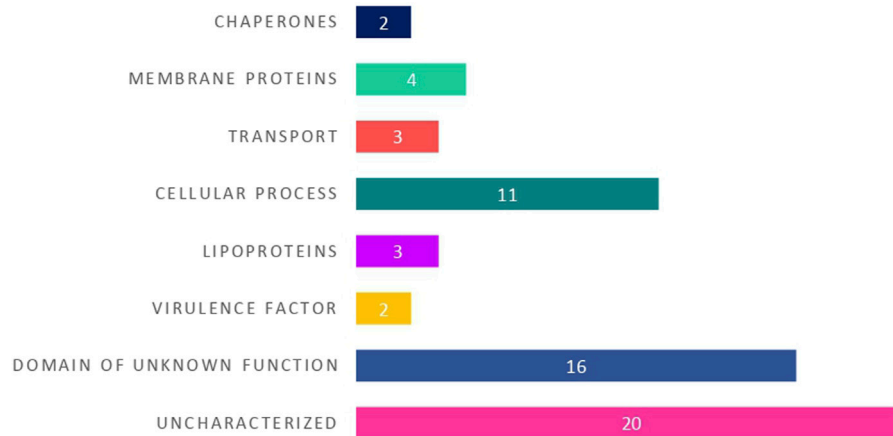


FIGURE 3 | The distribution of 25 hypothetical proteins into different functional groups based on their predicted protein domains and families.

hydrophilic protein moiety attached to hydrophobic fatty acid chains. These proteins are anchored to the membrane with their N-terminally linked fatty acids (Wilson and Bernstein, 2016). They are involved in diverse functions like outer membrane biogenesis, transport of various molecules, signal transduction, cell motility, and maintenance of cell shape (Narita and Tokuda, 2017). Previous studies on bacterial cold adaptation (Ting et al., 2010) have described the role of lipoproteins in optimizing membrane structure and function at low temperatures, indicating their role in cold adaptation. HP AUB73831.1 was annotated as LptE (lipopolysaccharide-assembly lipoprotein). LptE, together with LptD, assists in the proper assembly of lipopolysaccharide (Chimalakonda et al., 2011). Being a part of the cell envelope, proper assembly of LPS confers an advantage to bacterial cold adaptation (Tribelli and López, 2018). HP AUB74624.1 was predicted to be a lipoprotein with alanine-zipper.

Cellular Process

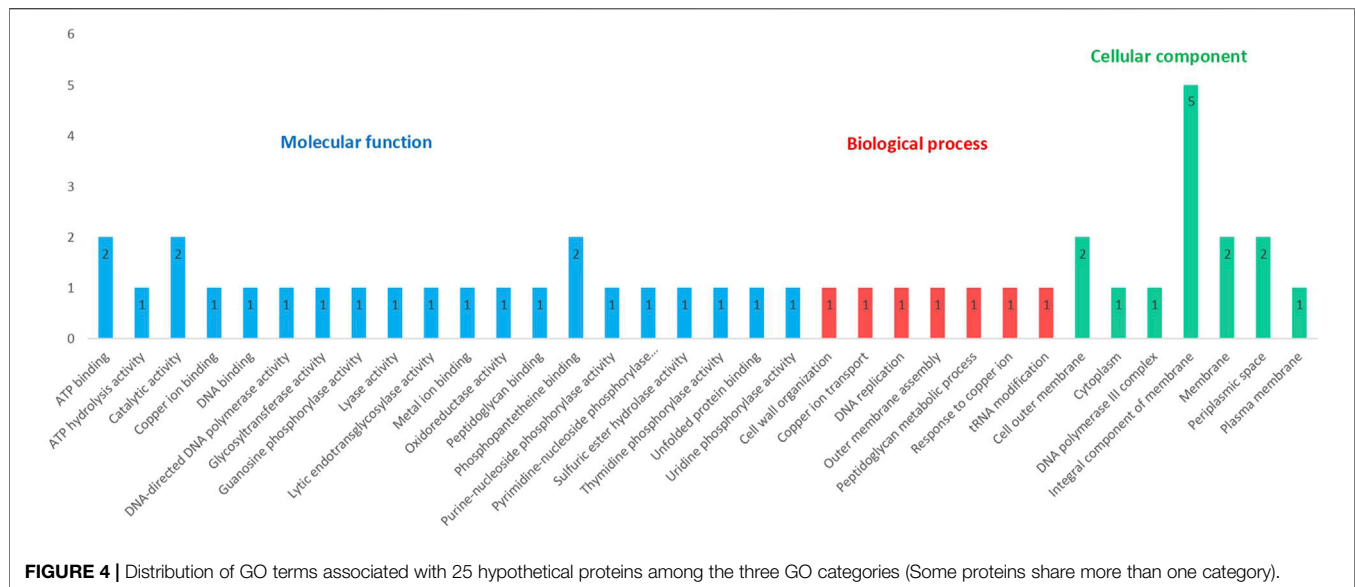
Microbial adaptation to cold temperatures necessitates changes in metabolism which is in turn regulated by enzymes. From our analysis, 10 HPs were characterized as enzymes. HP AUB73822.1 was predicted to be Rare lipoprotein A (RlpA). RlpA is a lytic peptidoglycan transglycosylase, which catalyzes specific cleavages in peptidoglycan chains, thus regulating the biosynthesis of peptidoglycan (Jorgenson et al., 2014). Two putative ligases (AUB75791.1, AUB75792.1) were also detected in this study. They contain a condensation domain that is known to catalyze a condensation reaction to form peptide bonds through a non-ribosomal peptide (NRP) biosynthesis mechanism (Izoré et al., 2021). Bacterial NRPs can serve as antibiotics, toxins, siderophores, pigments, cytostatics, etc. (Süssmuth and Mainz, 2017). HP AUB77590.1 was predicted to be a glycosyltransferase and is assumed to be involved in the biosynthesis of lipopolysaccharide and exopolysaccharides, important components in microbial cold adaptation (Soumya and Nampoothiri, 2021; Sen and Verma 2020; Deming and Young, 2017). Additional transferases identified were thiosulphate

sulphurtransferase (AUB74731.1), and aminoglycoside phosphotransferase (APH) (AUB76830.1). Aminoglycosides inhibit bacterial cell growth by impairing protein synthesis (Cavallo and Martinetto, 1981). APH inactivates aminoglycoside antibiotics via phosphorylation thereby defending the protein synthesis mechanism (Kotra et al., 2000). HP AUB76494.1 was identified to be a flavinator of succinate dehydrogenase (SdhE) (also known as Sdh5 in humans). SdhE is necessary for succinate dehydrogenase (SDH) activity which regulates cellular respiration via electron transport system. (McNeil et al., 2012). The ability to regulate cellular respiration at low temperatures could provide a distinct advantage for bacteria to cope with frozen conditions (Amato and Christner, 2009). Few other important enzymes predicted in this group of HPs were DNA polymerase III subunit with ATPase activity (AUB74781.1), sulphatase (AUB75991.1), and Pyrimidine/purine nucleoside phosphorylase (AUB74810.1).

HP AUB73565.1 was found to contain a YceI-like domain. The family of proteins with this domain was predicted to aid in isoprenoid quinone biosynthesis, thus playing a role in the electron transport system. In *E. coli*, YceI is induced by high pH (Stancik et al., 2002; Handa et al., 2005).

Transport proteins

Regulating the proteins of the transport systems is one of the mechanisms by which the bacteria survive extreme temperatures (Baraúna et al., 2017). HP AUB75187.1 was predicted to be MlaA lipoprotein. This protein is a part of an ABC transport system that facilitates the transport of phospholipids from the outer membrane to the inner membrane which is responsible for maintaining the lipid asymmetry in the outer membrane. Such asymmetric distribution of lipids in the OM maintains the membrane structure and confers resistance to gram-negative bacteria (Malinverni and Silhavy, 2009). HP AUB75954.1 was predicted to be a copper resistance protein C (CopC), which mediates copper homeostasis (Cha and Cooksey, 1991). HP AUB77446.1 was predicted to contain a Tim44-like domain. In eukaryotes, Tim44 is a part of translocation machinery that



imports nuclear-encoded proteins into the mitochondria. The bacterial homolog of Timm44, TimA, is likely to serve as a protein transporter (Clements et al., 2009).

Chaperones

Cold injury can cause protein denaturation or misfolding. Molecular chaperones could assist in the correct folding and stability of the proteins during various stress conditions. They also help in the refolding of misfolded proteins and solubilization of the aggregated proteins. (Chattopadhyay, 2006). In this study, two proteins (AUB76544.1, AUB76897.1) were categorized as chaperones. HP AUB76544.1 was annotated as Skp. This protein (also known as OmpH and HlpA) is a molecular chaperone that is important for the proper folding and insertion of several outer membrane proteins (Walton and Sousa, 2004). HP AUB76897 was found to comprise Hsp70 like domain implying its role as a chaperone in maintaining the structural integrity and function of proteins.

Virulence Factors

The proteins AUB73322.1, AUB74265.1 were found to be virulence factors. Virulence factors are expressed by pathogens to promote pathogenesis. They play an important role in bacterial growth and survival (Moradali et al., 2017). HP AUB74265.1 was found to be regulated by quorum sensing. (Chevalier et al., 2017). Quorum sensing (QS) is a mechanism by which bacterial cells communicate with each other using chemical signals. This helps bacterial communities to control various biological processes essential for bacterial adaptation and survival under stress (Moradali et al., 2017). The role of virulence factors in adaptive mechanisms to cold stress is not well understood.

Gene Ontology Analysis

These 25 proteins with known functional domains and provided putative function were further analyzed to uncover the GO terms

associated with them. Out of 25 targets, 20 of them have GO term predictions. Proteins were distributed within three GO categories - molecular function, biological process, and cellular component. GO terms from three different categories are illustrated in **Figure 4**. Complete results from the GO FEAT platform are presented in **Supplementary Table S6**. 20 different GO terms were identified under molecular function. Best of these referred to ATP binding (AUB74781.1, AUB76897.1), catalytic activity (AUB75791.1, AUB75792.1), and phosphopantetheine binding (AUB75791.1; AUB75792.1). The biological process category contained 7 different GO terms, including cell wall organization (AUB73822.1), outer membrane assembly (AUB73831.1), and peptidoglycan metabolic process (AUB73822.1), processes that are central to cold adaptation. The cellular component category revealed 7 GO terms with major representation to cell membrane confirming the central role of the bacterial membrane in cold adaptation.

Sequence Similarity Search

From BlastP analysis, 21 HPs were found to have homologous sequences in the NCBI database with described function. The remaining 4 (AUB73831.1, AUB74265.1, AUB74624.1, AUB75938.1) were matched to hypothetical proteins. The results were found to be consistent with functional domain analysis, except for HP AUB73831.1. From domain analysis, AUB73831.1 was predicted to be a Lipopolysaccharide-assembly lipoprotein (LptE). But its homologous sequences in other species of *pseudomonas* were found to be hypothetical. BlastP results are presented in **Supplementary Table S7**.

Sub-Cellular Localization

Most often, protein function is correlated to its location. Therefore, the information about protein's localization could provide supporting evidence to the functions predicted from functional domain analysis. However, emphasis was on investigating the localization of HPs that revealed functional

domains. Proteins that were categorized as membrane proteins (AUB74223.1, AUB76580.1, AUB76959.1, and AUB77315.1) were predicted to be located in the outer membrane. Additionally, proteins AUB73831.1 (Lipopolysaccharide-assembly lipoprotein), AUB74624.1 (lipoprotein), and AUB77590.1 (glycosyltransferase) were also predicted to be localized in the outer membrane. Proteins that were annotated as enzymes (AUB74731.1, AUB74781.1, AUB75791.1, AUB75792.1, AUB76494.1, AUB76830.1, and AUB76897.1) were mostly found to be located in the cytoplasm. Further, proteins like AUB73565.1 (YceI family protein), AUB73822.1 (Peptidoglycan transglycosylase), AUB75954.1 (Copper resistance protein C), and AUB76544.1 (skp) were predicted to be periplasmic. Nine proteins were predicted to have transmembrane helices based on TMHMM and HMMTOP. Using SignalP- 5.0, a total of 14 proteins with cleavage sites were detected, of which 7 were “standard” secretory signal peptides (Sec/SPI) and 7 were lipoprotein signal peptides (Sec/SPII). Examination of the results reveals that, for most proteins, there was an agreement between the inferred function and the location predicted. Sub-cellular localization analysis of each HP is summarized in **Supplementary Table S8**.

Analysis of Protein-Protein Interactions

Further, we looked for the proteins interacting with 25 HPs. As proteins do not work alone, the function of a HP may be inferred by looking at its interaction neighbourhood. Among 25 HPs, 21 HPs were shown to have interactions with other functionally known proteins of *Pseudomonas* sp. Lz4W with a confidence score above 0.7. Out of these, annotation is not available for 3 proteins. The remaining 18 proteins were analyzed for their functions based on interactions with other proteins (**Supplementary Table S9**).

Three proteins AUB74223.1, AUB76580.1, and AUB76959.1, annotated as outer membrane proteins belonging to the OmpA family, have demonstrated high confidence interaction with NADH-quinone oxidoreductase subunit I (NDH-1) encoded by *nuoI*. NDH-1 has been described to be involved in shuttling electrons from NADH to quinones, as part of the bacterial respiratory chain (Leif et al., 1995). Therefore, this interaction of OmpA proteins with NDH-1 suggests their role in the electron transport system, thereby regulation of energy metabolism, which is essential for survival in frozen temperatures.

The protein AUB73831.1, assigned as LptE, revealed strong interaction with LPS-assembly protein LptD, which was shown to be involved in the assembly of lipopolysaccharide (LPS) at the surface of the outer membrane, thus confirming the function attributed to AUB73831.1 by this study (Chimalakonda et al., 2011). We confidently found protein AUB74026.1 (lipoprotein) to be interacting with penicillin-binding proteins, 1A and 1B. These penicillin-binding proteins along with other hydrolases and synthetases take part in peptidoglycan synthesis (Charpentier et al., 2002). Hence, this lipoprotein is presumed to be involved in cell wall formation by participating in the synthesis of cross-linked peptidoglycan from the lipid intermediates (Braun and Hantke, 2019).

The protein AUB73822.1, annotated as RlpA (lytic transglycosylase), exhibited high degree interaction with mrdB and mrdA proteins that are described to be associated with peptidoglycan cell wall biogenesis (Cho et al., 2016), one of the mechanisms regulated during bacterial cold adaptation. This confirms the role of the protein AUB73822.1 in peptidoglycan metabolism, as inferred by this study. Proteins AUB75791.1 and AUB75792.1 annotated as peptide synthetases presented strong interactions with 4'-phosphopantetheinyl transferase and beta-ketoacyl synthase-like protein. Studies have shown that both phosphopantetheinyl transferases (PPTs) and beta-ketoacyl synthase are components of enzymatic systems like fatty acid synthases (FAS), non-ribosomal peptide synthetases (NRPS), and polyketide synthases (PKS). They play an essential enzymatic role in fatty acid biosynthesis and synthesis of various bioactive secondary metabolites like polyketides, antibiotics, and non-ribosomal peptide metabolites (Smith et al., 2003; Copp and Neilan, 2006; Beld et al., 2014). However, the role of non-ribosomal peptides in cold adaptation is not clear. But the significance of membrane fatty acid composition for cell viability at frozen temperatures was well studied (Chintalapati et al., 2004). Protein AUB77590.1, explained as glycosyltransferase from domain analysis, exhibited interaction with a group of proteins involved in the synthesis of outer membrane LPS along with few glycosyl hydrolases that hydrolyze the glycosidic bonds between polysaccharides. This confirms the functional role of AUB77590.1 in LPS and exopolysaccharides biosynthesis, specialized adaptations to cope with cold (Soumya and Nampoothiri, 2021). The coding sequence encoding the protein AUB74731.1 with thiosulphate sulphurtransferase (TST) domain revealed a neighbourhood with genes encoding thioredoxin and ABC transporter ATP-binding protein. Sulfurtransferases play a major role in maintenance of Fe-S containing electron transport proteins. Thus, TST together with the thioredoxin system may play a role in energy transduction. In general, the thioredoxin system also functions to protect bacteria against oxidative stress. At low temperatures, oxygen solubility increases rising the levels of reactive oxygen species (ROS), leading to oxidative stress. Thioredoxin system act as scavengers of ROS, thus reducing oxidative stress (Zeller and Klug, 2006; Tribelli and Lopez, 2018). The protein AUB76830.1, with aminoglycoside phosphotransferase domain, showed interaction with a protein involved in guanine biosynthesis. But its role in cold adaptation is not known. The protein AUB76494.1, with flavinator like domain (SdhE), exhibited high confidence interaction with succinate dehydrogenase (SDH) flavoprotein subunit. This interaction confirms the role of AUB76494.1 in the flavinylation of the flavoprotein subunit of SDH, responsible for the assembly of the respiratory enzyme complex. The protein AUB74781.1 was inferred to be a DNA polymerase III subunit with ATPase activity. It revealed high confidence interactions with other subunits of DNA polymerase III (dnaE, dnaQ, holoA) and proteins that have been described to be involved in replicative synthesis in bacteria. These interactions confirm the role of AUB74781.1 in DNA replication (Bargiela et al., 2020). The protein AUB74810.1 explained as pyrimidine/

TABLE 2 | List of HPs and their inferred functions in relation to cold adaptation.

S.No	Protein ID	Inferred function	Role in cold adaptation
1	AUB73822.1	Peptidoglycan transglycosylase involved in cell wall biosynthesis	Cell wall biosynthesis
2	AUB74731.1	Thiosulfate sulfurtransferase	Not known
3	AUB74781.1	ATPase involved in DNA replication	DNA replication machinery
4	AUB74810.1	Pyrimidine/purine nucleoside phosphorylase involved in AMP biosynthesis	Not known
5	AUB77590.1	Glycosyltransferase	Not known
6	AUB75791.1	Phosphopantetheine transferase involved in post translation modifications of fatty acid synthases (FAS)	Fatty acid metabolism required for membrane fluidity
7	AUB75792.1	Phosphopantetheinyl transferase involved in post translation modifications of fatty acid synthases (FAS)	Fatty acid metabolism required for membrane fluidity
8	AUB76494.1	Flavinator of succinate dehydrogenase (SdhE)	Cellular respiration
9	AUB76830.1	Aminoglycoside phosphotransferase (APH)	Not known
10	AUB74223.1	OmpA family protein involved in electron transport	Energy metabolism
11	AUB76580.1	Outer membrane protein (OmpA) involved in respiratory chain	Energy metabolism
12	AUB76959.1	Outer membrane protein (OmpA) involved in respiratory chain	Energy metabolism
13	AUB73831.1	Lipoplysaccharide-assembly lipoprotein (LptE) involved in the assembly of lipopolysaccharide	Membrane structure and stability
14	AUB74026.1	Lipoprotein	Peptidoglycan biosynthesis
15	AUB75954.1	Copper resistance protein C involved in copper homeostasis	Not known
16	AUB76897.1	Molecular chaperone	Preventing the aggregation of stress-denatured proteins
17	AUB76544.1	Chaperone (skp)	Assembly and insertion of beta-barrel proteins in outer membrane
18	AUB73565.1	YceI family protein participating in ETC	Not known

purine nucleoside phosphorylase showed strong interactions with proteins associated with the purine salvage pathway. However, the exact role of this protein in cold adaptation is unclear.

The protein AUB73565.1 which was found to contain a YceI-like domain showed high confidence interaction with Cytochrome b561. The gene that encodes cytochrome b561 is present upstream to yceI. Cytochrome b561 is an integral membrane protein involved in electron transfer activity. In some species, polyisoprenoid binding protein belonging to the YceI-like family was also found to co-exist with cytochrome b561. Thus, probably proteins with YceI-like domain are involved in electron transport system (Stancik et al., 2002; Handa et al., 2005). However, further analysis is needed to confirm its role in cold adaptation.

The gene encoding the protein AUB76544.1 demonstrated a proximal relationship with genes lpxA and lpxD. These genes were shown to be involved in the biosynthesis of lipid A, a component of LPS required for anchoring to the outer membrane of the cell (Kroeck et al., 2019). These functional interactions confirm the role of AUB76544.1 as a chaperone assisting in the proper folding and insertion of several outer membrane proteins. The protein AUB76897.1, containing Hsp70 like domain, exhibited a great degree of association with other chaperones like grpE, dnaJ, groEL. These proteins were closely related to promoting proper folding and assembly of unfolded polypeptides generated under stress conditions and were shown to have enhanced expression under cold stress (Chen et al., 2012). Hence it is evident that AUB76897.1 functions as a chaperone in maintaining the stability and function of stress-denatured proteins, crucial in cold conditions.

The protein AUB77315.1 was found to encompass the outer membrane porin (OprD) domain. From STRING analysis, the gene encoding OprD was found to be co-occurring with genes

coding for porins, signifying a shared function (Szklarczyk et al., 2021). Though the confidence score of this interaction is less than our selected cut-off (>0.7), the results are consistent across the annotation pipeline. Hence, this protein can be inferred as an outer membrane porin involved in the transport across the membrane.

Proteins AUB73322.1, AUB74624.1, AUB75954.1, and AUB77446.1 were found to interact with proteins for which annotation is not available. For proteins AUB75187.1, and AUB75991.1, the confidence score of interactions is less than our selected cut-off (>0.7). Hence the interactions of these proteins were not considered for functional characterization. PPI networks are presented in **Supplementary Figure S1**.

Finally, for each of the 25 HPs, we combined all the results from different annotation methods including domain analysis, gene ontology, sequence homology, sub-cellular localization, and PPIs to infer the final function. From our study, we could assign a function to 18 HPs with high confidence, of which 12 were found to be playing an important role in cold adaptation of *Pseudomonas* sp. Lz4W. The summary of this analysis is shown in **Table 2**. Complete results are provided in **Supplementary Table S10**.

CONCLUSION

In general, HPs constitute about 30–50% of the microbial genomes, and the HP data is rapidly accumulating due to the deluge of sequencing data from sequencing technologies like NGS and mass spectrometry. Characterizing this HP data is one of the challenges in modern biomedical research. The genomic analysis of the Antarctic bacterium *Pseudomonas* sp. Lz4W revealed 743 CDS annotated as hypothetical proteins. Our proteomic analysis confirmed the

expression of 61 hypothetical proteins at the translation level. This is the first study to characterize the HPs related to bacterial cold adaptation, revealing their role in membrane structure and stability, transport, protein folding, energy metabolism, and enzymatic regulation. Overall, our study once again highlighted the functional significance of hypothetical or uncharacterized sequences from the genomes. Furthermore, we have developed an integrated model based on mass spectrometry-based proteomics and systematic bioinformatics analysis to identify and characterize the uncharacterized proteins.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ Supplementary Material- MS Proteomics data is available on PRIDE and can be accessed under the ID: PXD029741.

AUTHOR CONTRIBUTIONS

MJ and MR designed the proteomics experiments; JI designed the bioinformatics study; DC performed the experiments; JI

interpreted the data and performed bioinformatics analysis; JI and MJ drafted and revised the manuscript. All authors read and approved the final manuscript.

FUNDING

This research received funding from CSIR- Centre for Cellular and Molecular Biology. DC received a research fellowship from CSIR, India. This work is in-house project of CCMB and not supported by any funding/grants.

ACKNOWLEDGMENTS

We thank Dr. B. Raman and Ms. D. B. Kameswari from the Proteomics facility, CSIR-CCMB for their help with mass spectral analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.825269/full#supplementary-material>

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol.* 215 (3), 403–410. doi:10.1016/S0022-2836(05)80360-2
- 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 Res.* 25 (17), 3389–3402. doi:10.1093/nar/25.17.3389
- Amato, P., and Christner, B. C. (2009). Energy Metabolism Response to Low-Temperature and Frozen Conditions in *Psychrobacter Cryohalolentis*. *Appl. Environ. Microbiol.* 75 (3), 711–718. doi:10.1128/AEM.02193-08
- Ames, G. F. (1974). Resolution of Bacterial Proteins by Polyacrylamide Gel Electrophoresis on Slabs. Membrane, Soluble, and Periplasmic Fractions. *J. Biol. Chem.* 249 (2), 634–644. doi:10.1016/s0021-9258(19)43074-3
- Araújo, C. L., Blanco, I., Souza, L., Tiwari, S., Pereira, L. C., Ghosh, P., et al. (2020). In Silico functional Prediction of Hypothetical Proteins from the Core Genome of *Corynebacterium Pseudotuberculosis* Biovar ovis. *PeerJ* 8, e9643. doi:10.7717/peerj.9643
- Araujo, F. A., Barh, D., Silva, A., Guimarães, L., and Ramos, R. (2018). GO FEAT: a Rapid Web-Based Functional Annotation Tool for Genomic and Transcriptomic Data. *Scientific Rep.* 8 (1), 1794. doi:10.1038/s41598-018-20211-9
- Baraúna, R. A., Freitas, D. Y., Pinheiro, J. C., Foador, A. R., and Silva, A. (2017). A Proteomic Perspective on the Bacterial Adaptation to Cold: Integrating OMICs Data of the Psychrotrophic Bacterium *Exiguobacterium Antarcticum* B7. *Proteomes* 5 (1), 9. doi:10.3390/proteomes5010009
- Bargiela, R., Lanthaler, K., Potter, C. M., Ferrer, M., Yakunin, A. F., Paizs, B., et al. (2020). Proteome Cold-Shock Response in the Extremely Acidophilic Archaeon, *Cuniculiplasma Divulgatum*. *Microorganisms* 8 (5), 759. doi:10.3390/microorganisms8050759
- Beld, J., Sonnenschein, E. C., Vickery, C. R., Noel, J. P., and Burkart, M. D. (2014). The Phosphopantetheinyl Transferases: Catalysis of a post-translational Modification Crucial for Life. *Nat. Product. Rep.* 31 (1), 61–108. doi:10.1039/c3np70054b
- Blum, M., Chang, H. Y., Chuguransky, S., Grego, T., Kandasamy, S., Mitchell, A., et al. (2021). The InterPro Protein Families and Domains Database: 20 Years on. *Nucleic Acids Res.* 49 (D1), D344–D354. doi:10.1093/nar/gkaa977
- Braun, V., and Hantke, K. (2019). “Lipoproteins: Structure, Function, Biosynthesis,” in *Bacterial Cell Walls and Membranes*. Editor A Kuhn (Cham: Springer). doi:10.1007/978-3-030-18768-2_3
- Cavallo, G., and Martinetto, P. (1981). The Mechanism of Action of Aminoglycosides”. *G Bacteriol Virol. Immunol.* 74 (7–12), 335–346.
- Cha, J. S., and Cooksey, D. A. (1991). Copper Resistance in *Pseudomonas syringae* Mediated by Periplasmic and Outer Membrane Proteins. *Proc. Natl. Acad. Sci. United States America* 88 (20), 8915–8919. doi:10.1073/pnas.88.20.8915
- Charpentier, X., Chalut, C., Rémy, M. H., and Masson, J. M. (2002). Penicillin-binding Proteins 1a and 1b Form Independent Dimers in *Escherichia coli*. *J. Bacteriol.* 184 (13), 3749–3752. doi:10.1128/JB.184.13.3749-3752.2002
- Chattopadhyay, M. K. (2006). Mechanism of Bacterial Adaptation to Low Temperature. *J. Biosci.* 31 (1), 157–165. doi:10.1007/bf02705244
- Chattopadhyay, M. K., and Jagannadham, M. V. (2001). Maintenance of Membrane Fluidity in Antarctic Bacteria. *Polar Biol.* 24, 386–388. doi:10.1007/s003000100232
- Chen, Z., Yu, H., Li, L., Hu, S., and Dong, X. (2012). The Genome and Transcriptome of a Newly Described Psychrophilic Archaeon, *Methanohalobium Psychrophilum* R15, Reveal its Cold Adaptive Characteristics. *Environ. Microbiol. Rep.* 4 (6), 633–641. doi:10.1111/j.1758-2229.2012.00389.x
- Chevalier, S., Bouffartigues, E., Bodilis, J., Maillot, O., Lesouhaitier, O., Feuilloley, M. G. J., et al. (2017). Structure, Function and Regulation of *Pseudomonas aeruginosa* Porins. *FEMS Microbiol. Rev. Sep.* 1 (415), 698–722. doi:10.1093/femsre/fux020
- Chimalakonda, G., Ruiz, N., Chng, S. S., Garner, R. A., Kahne, D., and Silhavy, T. J. (2011). Lipoprotein LptE Is Required for the Assembly of LptD by the Beta-Barrel Assembly Machine in the Outer Membrane of *Escherichia coli*. *Proc. Natl. Acad. Sci. U S A.* 108 (6), 2492–2497. doi:10.1073/pnas.1019089108
- Chintalapati, S., Kiran, M. D., and Shivaji, S. (2004). Role of Membrane Lipid Fatty Acids in Cold Adaptation. *Cel Mol Biol (Noisy-le-grand)* 50, 631–642.

- Cho, H., Wivagg, C., Kapoor, M., Zachary, B., Rochs, P. D. A., Suh, H., et al. (2016). Bacterial Cell wall Biogenesis Is Mediated by SEDS and PBP Polymerase Families Functioning Semi-autonomously. *Nat. Microbiol.* 1, 16172. doi:10.1038/nmicrobiol.2016.172
- Clements, A., Bursac, D., Gatsos, X., Perry, A. J., Civciristov, S., Celik, N., et al. (2009). The Reducible Complexity of a Mitochondrial Molecular Machine. *Proc. Natl. Acad. Sci. United States America* 106 (37), 15791–15795. doi:10.1073/pnas.0908264106
- Confer, A. W., and Ayalew, S. (2013). The OmpA Family of Proteins: Roles in Bacterial Pathogenesis and Immunity. *Vet. Microbiol.* 163 (3-4), 207–222. doi:10.1016/j.vetmic.2012.08.019
- Copp, J. N., and Neilan, B. A. (2006). The Phosphopantetheinyl Transferase Superfamily: Phylogenetic Analysis and Functional Implications in Cyanobacteria. *Appl. Environ. Microbiol.* 72 (4), 2298–2305. doi:10.1128/AEM.72.4.2298-2305.2006
- da Costa, W. L. O., Araújo, C. L. A., Dias, L. M., Pereira, L. C. S., Alves, J. T. C., Araújo, F. A., et al. (2018). Functional Annotation of Hypothetical Proteins from the Exiguobacterium Antarcticum Strain B7 Reveals Proteins Involved in Adaptation to Extreme Environments, Including High Arsenic Resistance. *PLoS One* 25, 13. doi:10.1371/journal.pone.0198965
- Dall'Agnol, H. P., Baraúna, R. A., de Sá, P. H., Ramos, R. T., Nóbrega, F., Nunes, C. I., et al. (2014). Omics Profiles Used to Evaluate the Gene Expression of Exiguobacterium Antarcticum B7 during Cold Adaptation. *BMC genomics* 15 (1), 986. doi:10.1186/1471-2164-15-986
- De Maayer, P., Anderson, D., Cary, C., and Cowan, D. A. (2014). Some like it Cold: Understanding the Survival Strategies of Psychrophiles. *EMBO Rep.* 15 (5), 508–517. doi:10.1002/embr.201338170
- Deming, J. W., and Young, J. N. (2017). “The Role of Exopolysaccharides in Microbial Adaptation to Cold Habitats,” in *Psychrophiles: From Biodiversity to Biotechnology*. Editor R Margesin (Cham: Springer). doi:10.1007/978-3-319-57057-0_12
- Desler, C., Suravajhala, P., Sanderhoff, M., Rasmussen, M., and Rasmussen, L. J. (2009). Silico Screening for Functional Candidates Amongst Hypothetical Proteins. *BMC Bioinformatics* 16 (10), 289. doi:10.1186/1471-2105-10-289
- Eisenstein, E., Gilliland, G. L., Herzberg, O., Moullet, J., Orban, J., Poljak, R. J., et al. (2000). Biological Function Made crystal clear—Annotation of Hypothetical Proteins via Structural Genomics. *Curr. Opin. Biotechnol.* 11, 25–30. doi:10.1016/s0958-1669(99)00063-4
- Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., et al. (2014). Pfam: the Protein Families' Database. *Nucleic Acids Res.* 42 (Database issue), D222–D230. doi:10.1093/nar/gkt1223
- Frank, S., Schmidt, F., Klockgether, J., Davenport, C. F., Gesell Salazar, M., Völker, U., et al. (2011). Functional Genomics of the Initial Phase of Cold Adaptation of Pseudomonas Putida KT2440. *FEMS Microbiol. Lett.* 318 (1), 47–54. doi:10.1111/j.1574-6968.2011.02237.x
- Gardy, J. L., Spencer, C., Wang, K., Ester, M., Tusnády, G. E., Simon, I., et al. (2003). S. PSORT-B: Improving Protein Subcellular Localization Prediction for Gram-Negative Bacteria. *Nucleic Acids Res.* 31 (13), 3613–3617. doi:10.1093/nar/gkg602
- Gough, J., Karplus, K., Hughey, R., and Chothia, C. (2001). Assignment of Homology to Genome Sequences Using a Library of Hidden Markov Models that Represent All Proteins of Known Structure. *J. Mol. Biol.* 313 (4), 903–919. doi:10.1006/jmbi.2001.5080
- Gundry, R. L., White, M. Y., Murray, C. I., Kane, L. A., Fu, Q., Stanley, B. A., et al. (2009). Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow. *Curr. Protoc. Mol. Biol.* 10, 10.25.1–10.25.23. doi:10.1002/0471142727.mb1025s88
- Handa, N., Terada, T., Doi-Katayama, Y., Hirota, H., Tame, J. R., and Park, S. Y. (2005). Crystal Structure of a Novel Polyisoprenoid-Binding Protein from Thermus Thermophilus HB8. *Protein Sci. a Publ. Protein Soc.* 14 (4), 1004–1010. doi:10.1110/ps.041183305
- Ijaq, J., Bethi, N., and Jagannadham, M. V. (2020). Mass Spectrometry-Based Identification and Characterization of Human Hypothetical Proteins Highlighting the Inconsistency across the Protein Databases. *J. Proteins Proteom* 11, 17–25. doi:10.1007/s42485-020-00030-1
- Ijaq, J., Chandrasekharan, M., Poddar, R., Bethi, N., and Sundararajan, V. S. (2015). Annotation and Curation of Uncharacterized Proteins- Challenges. *Front. Genet.* 6, 119. doi:10.3389/fgene.2015.00119
- Ijaq, J., Malik, G., Kumar, A., Kumar, A., Das, P. S., Meena, N., et al. (2019). A Model to Predict the Function of Hypothetical Proteins through a Nine-point Classification Scoring Schema. *BMC Bioinformatics* 20, 14. doi:10.1186/s12859-018-2554-y
- Imai, K., Asakawa, N., Tsuji, T., Akazawa, F., Ino, A., Sonoyama, M., et al. (2008). SOSUI-GramN: High Performance Prediction for Sub-cellular Localization of Proteins in Gram-Negative Bacteria. *Bioinformatics* 2 (9), 417–421. doi:10.6026/97320630002417
- Izoré, T., Candace Ho, Y. T., Kaczmarek, J. A., Gavrilidou, A., Chow, K. H., Steer, D. L., et al. (2021). Structures of a Non-ribosomal Peptide Synthetase Condensation Domain Suggest the Basis of Substrate Selectivity. *Nat. Commun.* 12, 2511. doi:10.1038/s41467-021-22623-0
- Jagannadham, M. V., Abou-Eladab, E. F., and Kulkarni, H. M. (2011). Identification of Outer Membrane Proteins from an Antarctic Bacterium Pseudomonas syringae Lz4W. *Mol. Cel Proteomics* 10 (6), M110. doi:10.1074/mcp.M110.004549
- Jagannadham, M. V. (2008). Identification of Proteins from Membrane Preparations by a Combination of MALDI TOF-TOF and LC-Coupled Linear Ion Trap MS Analysis of an Antarctic Bacterium Pseudomonas syringae Lz4W, a Strain with Unsequenced Genome. *Electrophoresis* 29 (21), 4341–4350. doi:10.1002/elps.200700750
- Jagannadham, M. V., and Chowdhury, C. (2012). Differential Expression of Membrane Proteins Helps Antarctic Pseudomonas syringae to Acclimatize upon Temperature Variations. *J. Proteomics* 75 (8), 2488–2499. doi:10.1016/j.jprot.2012.02.033
- Jorgenson, M. A., Chen, Y., Yahashiri, A., Popham, D. L., and Weiss, D. S. (2014). The Bacterial Septal Ring Protein RlpA Is a Lytic Transglycosylase that Contributes to Rod Shape and Daughter Cell Separation in Pseudomonas aeruginosa. *Mol. Microbiol.* 93 (1), 113–128. doi:10.1111/mmi.12643
- Kannan, K., Janiyani, K. L., Shivaji, S., and Ray, M. K. (1998). Histidine Utilisation Operon (Hut) Is Upregulated at Low Temperature in the Antarctic Psychrotrophic Bacterium Pseudomonas syringae. *FEMS Microbiol. Lett.* 161 (1), 7–14. doi:10.1111/j.1574-6968.1998.tb12922.x
- Kotra, L. P., Haddad, J., and Mobashery, S. (2000). Aminoglycosides: Perspectives on Mechanisms of Action and Resistance and Strategies to Counter Resistance. *Antimicrob. Agents Chemother.* 44 (12), 3249–3256. doi:10.1128/AAC.44.12.3249-3256.2000
- Kroec, K. G., Sacco, M. D., Smith, E. W., Zhang, X., Shoun, D., Akthar, A., et al. (2019). Discovery of Dual-Activity Small-Molecule Ligands of Pseudomonas aeruginosa LpxA and LpxD Using SPR and X-ray Crystallography. *Sci. Rep.* 9, 15450. doi:10.1038/s41598-019-51844-z
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* 305 (3), 567–580. doi:10.1006/jmbi.2000.4315
- Kumar, G. S., Jagannadham, M. V., and Ray, M. K. (2002). Low-temperature-induced Changes in Composition and Fluidity of Lipopolysaccharides in the Antarctic Psychrotrophic Bacterium Pseudomonas syringae. *J. Bacteriol.* 184 (23), 6746–6749. doi:10.1128/JB.184.23.6746-6749.2002
- Leif, H., Sled, V. D., Ohnishi, T., Weiss, H., and Friedrich, T. (1995). Isolation and Characterization of the Proton-Translocating NADH: Ubiquinone Oxidoreductase from Escherichia coli. *Eur. J. Biochem.* 230 (2), 538–548. doi:10.1111/j.1432-1033.1995.0538h.x
- Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: Recent Updates to the Protein Domain Annotation Resource. *Nucleic Acids Res.* 40 (Database issue), D302–D305. doi:10.1093/nar/gkr931
- Li, H., Luo, Y. F., Williams, B. J., Blackwell, T. S., and Xie, C. M. (2012). Structure and Function of OprD Protein in Pseudomonas aeruginosa: from Antibiotic Resistance to Novel Therapies. *Int. J. Med. Microbiol.* 302 (2), 63–68. doi:10.1016/j.ijmm.2011.10.001
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, NR, et al. (2020). CDD/SPARCLE: the Conserved Domain Database in 2020. *Nucleic Acids Res.* 48 (D1), D265–D268. doi:10.1093/nar/gkz991
- Malinverni, J. C., and Silhavy, T. J. (2009). An ABC Transport System that Maintains Lipid Asymmetry in the Gram-Negative Outer Membrane. *Proc. Natl. Acad. Sci. U S A.* 106 (19), 8009–8014. doi:10.1073/pnas.0903229106

- McNeil, M. B., Clulow, J. S., Wilf, N. M., Salmond, G. P., and Fineran, P. C. (2012). SdhE Is a Conserved Protein Required for Flavinylation of Succinate Dehydrogenase in Bacteria. *J. Biol. Chem.* 287 (22), 18418–18428. doi:10.1074/jbc.M111.293803
- Moradali, M. F., Ghods, S., and Rehm, B. H. A. (2017). *Pseudomonas aeruginosa* Lifestyle: A Paradigm for Adaptation, Survival, and Persistence. *Front. Cel. Infect. Microbiol.* 7 (7), 39. doi:10.3389/fcimb.2017.00039
- Naqvi, A. A., 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. doi:10.1371/journal.pone.0124177
- Narita, S. I., and Tokuda, H. (2017). Bacterial Lipoproteins; Biogenesis, Sorting and Quality Control. *Biochim. Biophys. Acta Mol. Cel Biol Lipids* 1862 (11), 1414–1423. doi:10.1016/j.bbalip.2016.11.009
- Nejadi, N., Masti, S. M., Tavirani, M. R., and Golmohammadi, T. (2014). Comparison of Three Routine Protein Precipitation Methods: Acetone, TCA/acetone Wash and TCA/acetone. *J. paramedical Sci.* 5. doi:10.22037/JPS.V5I4.7699
- Pandiyani, A., and Ray, M. K. (2013). Draft Genome Sequence of the Antarctic Psychrophilic Bacterium *Pseudomonas syringae* Strain Lz4W. *Genome Announc* 1 (3), e00377–13. doi:10.1128/genomeA.00377-13
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: Discriminating Signal Peptides from Transmembrane Regions. *Nat. Methods* 8 (10), 785–786. doi:10.1038/nmeth.1701
- Ponting, C. P., and Russell, R. R. (2002). The Natural History of Protein Domains. *Annu. Rev. Biophys. Biomol. Struct.* 31, 45–71. doi:10.1146/annurev.biophys.31.082901.134314
- Prakash, J. S. S., Krishna, P. S., Sirisha, K., Kanesaki, Y., and Suzuki, I. (2010). An RNA Helicase, CrhR, Regulates the Low-Temperature-Inducible Expression of Three Genes for Molecular Chaperones in *Synechocystis* Sp. PCC 6803. *Microbiology* 156, 442–451. doi:10.1099/mic.0.031823-0
- Pranavathyani, G., Prava, J., Rajeev, A. C., and Pan, A. (2020). Novel Target Exploration from Hypothetical Proteins of *Klebsiella pneumoniae* MGH 78578 Reveals a Protein Involved in Host-Pathogen Interaction. *Front Cel Infect Microbiol* 3 (10), 109. doi:10.3389/fcimb.2020.00109
- Rabbi, M. F., Akter, S. A., Hasan, M. J., and Amin, A. (2021). Silico Characterization of a Hypothetical Protein from *Shigella dysenteriae* ATCC 12039 Reveals a Pathogenesis-Related Protein of the Type-VI Secretion System. *Bioinform Biol. Insights* 22 (15), 11779322211011140. doi:10.1177/11779322211011140
- Ray, M. K., Kumar, G. S., Janiyani, K., Kannan, K., Jagtap, P., Basu, M. K., et al. (1998). Adaptation to Low Temperature and Regulation of Gene Expression in Antarctic Psychrotrophic Bacteria. *J. Biosci.* 23, 423–435. doi:10.1007/BF02936136
- Ray, M. K., Kumar, G. S., and Shivaji, S. (1994). Phosphorylation of Lipopolysaccharides in the Antarctic Psychrotroph *Pseudomonas syringae*: a Possible Role in Temperature Adaptation. *J. Bacteriol.* 176 (14), 4243–4249. doi:10.1128/jb.176.14.4243-4249.1994
- Sen, T., and Verma, N. K. (2020). Functional Annotation and Curation of Hypothetical Proteins Present in A Newly Emerged Serotype 1c of *Shigella Flexneri*: Emphasis on Selecting Targets for Virulence and Vaccine Design Studies. *Genes* 11 (3), 340. doi:10.3390/genes11030340
- Shahbaaz, M., Hassan, M. I., and Ahmad, F. (2013). Functional Annotation of Conserved Hypothetical Proteins from *Haemophilus Influenzae* Rd KW20. *PLoS One* 8 (12), e84263. doi:10.1371/journal.pone.0084263
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303
- Shivaji, S., and Prakash, J. S. S. (2010). How Do Bacteria Sense and Respond to Low Temperature? *Arch. Microbiol.* 192 (2), 85–95. doi:10.1007/s00203-009-0539-y
- Shivaji, S., Rao, N. S., Saisree, L., Sheth, V., Reddy, G. S., and Bhargava, P. M. (1989). Isolation and Identification of *Pseudomonas* Spp. From Schirmacher Oasis, Antarctica. *Appl. Environ. Microbiol.* 55 (3), 767–770. doi:10.1128/aem.55.3.767-770.1989
- Sigrist, C. J., de Castro, E., Cerutti, L., Cuche, B. A., Hulo, N., Bridge, A., et al. (2013). New and Continuing Developments at PROSITE. *Nucleic Acids Res.* 41 (Database issue), D344–D347. doi:10.1093/nar/gks1067
- Sillitoe, I., Lewis, T. E., Cuff, A., Das, S., Ashford, P., Dawson, N. L., et al. (2015). CATH: Comprehensive Structural and Functional Annotations for Genome Sequences. *Nucleic Acids Res.* 43 (Database issue), D376–D381. doi:10.1093/nar/gku947
- Smith, S. G., Mahon, V., Lambert, M. A., and Fagan, R. P. (2007). A Molecular Swiss Army Knife: OmpA Structure, Function and Expression. *FEMS Microbiol. Lett.* 273 (1), 1–11. doi:10.1111/j.1574-6968.2007.00778.x
- Smith, S., Witkowski, A., and Joshi, A. K. (2003). Structural and Functional Organization of the Animal Fatty Acid Synthase. *Prog. Lipid Res.* Jul 42 (4), 289–317. doi:10.1016/s0163-7827(02)00067-x
- Soumya, M. P., and Nampoothiri, K. M. (2021). An Overview of Functional Genomics and Relevance of Glycosyltransferases in Exopolysaccharide Production by Lactic Acid Bacteria. *Int. J. Biol. Macromol* 184, 1014–1025. doi:10.1016/j.ijbiomac.2021.06.131
- Stancik, L. M., Stancik, D. M., Schmidt, B., Barnhart, D. M., Yoncheva, Y. N., and Slonczewski, J. L. (2002). P^H-dependent Expression of Periplasmic Proteins and Amino Acid Catabolism in *Escherichia coli*. *J. Bacteriol.* 184 (15), 4246–4258. doi:10.1128/jb.184.15.4246-4258.2002
- Süssmuth, R. D., and Mainz, A. (2017). Nonribosomal Peptide Synthesis—Principles and Prospects. *Angew. Chem. Int.* 56, 3770–3821. doi:10.1002/anie.201609079
- Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., et al. (2021). The STRING Database in 2021: Customizable Protein-Protein Networks, and Functional Characterization of User-Uploaded Gene/measurement Sets. *Nucleic Acids Res.* 49 (D1), D605–D612. doi:10.1093/nar/gkaa1074
- Ting, L., Williams, T. J., Cowley, M. J., Lauro, F. M., Guilhaus, M., Raftery, M. J., et al. (2010). Cold Adaptation in the marine Bacterium, *Sphingopyxis Alaskensis*, Assessed Using Quantitative Proteomics. *Environ. Microbiol.* 12 (10), 2658–2676. doi:10.1111/j.1462-2920.2010.02235.x
- Tribelli, P. M., and López, N. I. (2018). Reporting Key Features in Cold-Adapted Bacteria. *Life (Basel)* 13 (1), 8. doi:10.3390/life8010008
- Tusnády, G. E., and Simon, I. (2001). The HMMTOP Transmembrane Topology Prediction Server. *Bioinformatics* 17 (9), 849–850. doi:10.1093/bioinformatics/17.9.849
- Walton, T. A., and Sousa, M. C. (2004). Crystal Structure of Skp, a Prefoldin-like Chaperone that Protects Soluble and Membrane Proteins from Aggregation. *Mol. Cel.* 15 (3), 367–374. doi:10.1016/j.molcel.2004.07.023
- Wang, Y. (2002). The Function of OmpA in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 292 (2), 396–401. doi:10.1006/bbrc.2002.6657
- Wilson, M. M., and Bernstein, H. D. (2016). Surface-Exposed Lipoproteins: An Emerging Secretion Phenomenon in Gram-Negative Bacteria. *Trends Microbiol.* 24 (3), 198–208. doi:10.1016/j.tim.2015.11.006
- Yu, C. S., Chen, Y. C., and Lu, C. H. (2006). Hwang JK. Prediction of Protein Subcellular Localization. *Proteins* 64 (3), 643–651. doi:10.1002/prot.21018
- Zeller, T., and Klug, G. (2006). Thioredoxins in Bacteria: Functions in Oxidative Stress Response and Regulation of Thioredoxin Genes. *Naturwissenschaften.* Jun 93 (6), 259–266. doi:10.1007/s00114-006-0106-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ijaq, Chandra, Ray and Jagannadham. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.