Biological Forum – An International Journal **14(4): 811-818(2022)**

ISSN No. (Print): 0975-1130
ISSN No. (Online): 2249-3239

In silico Structural Prediction and Functional Analysis of Cold Tolerant Glycerol-3-Phosphate Dehydrogenase Protein of Rainbow Trout

Ubaid Qayoom[1]*, Zahoor Mushtaq[2] and Pabitra Barik[3]
[1]ICAR-National Bureau of Fish Genetic Resources, Lucknow (Uttar Pradesh), India.
[2]Centurion University of Technology and Management, Paralakhemundi (Odisha), India.
[3]Dau Shri Vasudev Chandrakar Kamdhenu Vishwavidyalaya (Chhattisgarh), India.

(Corresponding author: Ubaid Qayoom)*
(Received 05 September 2022, Accepted 28 October, 2022)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: The present study aimed to investigate the glycerol-3-phosphate dehydrogenase (GPDH) protein of the rainbow trout (*Oncorhynchus mykiss*). Although the protein plays a crucial role in the cold tolerance pathway of fishes, especially rainbow trout, its structural and functional annotation remains to be studied. The present study was designed to enlist some of the physiochemical and functional properties of GPDH protein and generate information about its three-dimensional structure. The primary structure evaluation inferred that it is an acidic and hydrophobic protein with a molecular weight of 39.9 kDa. Functional analysis by InterPro classified the protein into NAD(P)-binding domain family having a biological role in the redox processes. The $SOSUI_{GramN}$ server indicated the cytoplasmic subcellular localization without any signal sequence. The three-dimensional model built using the Swiss-Model server was evaluated using QMEAN, ProQ, and ProSA programs. The predicted 3D structure will offer a good foundation for functional analysis of experimental origin crystal structures. The closest interacting protein was found to be belonging to the Glycerol-3-phosphate acyltransferase family having a score of >0.7 via the STRINGv 11.5 server. Overall, the present study of structural characterization of the trout glycerol-3-phosphate dehydrogenase will facilitate an improved understanding of its enzymatic actions as well as its role in the cold tolerance of rainbow trout.

Keywords: *Oncorhynchus mykiss*, GPDH, Homology modelling, Bioinformatics.

INTRODUCTION

Rainbow trout (*Oncorhynchus mykiss*), a freshwater salmonid, is one of the most promising cultivable fish species in cold waters and has considerable scope for its expansion in the upland regions. The culture of rainbow trout has been in vogue for more than a century, making it one of the longest-farmed fish in aquaculture. Trout is a native ofthe western USA, Western Canada, and North Western Mexico but has been introduced into more than 80 countries, including India (Welcomme 1988) for culture and game fisheries. It has tremendous potential for domestic consumption in cold climate geographic zones as well as foreign export. World farmed-trout production increased from 4,400 tonnes in 1950, to approximately 0.96 million tonnes, in 2018 with Iran being the largest producer (FAO 2020). In India, trout fisheries and culture have become economically significant, especially in the Himalayan regions, which are widely known for producing table-sized trout in high demand. Trout culture in India has flourished in states like Jammu and Kashmir, Himachal Pradesh, Arunachal Pradesh, and Sikkim, owing to the fitting temperature ranges (5 - 18°C) along with ample cold-water assetsfor the culture of this high-value fish (DCFR 2015).

Because fishes are ectothermic animals, the temperature of the environment majorly controls physiological and metabolic phenomena like growth and breeding (Brett 1979). During evolution, fishes have adapted to a wide range of temperatures, evident by their distribution around the globe across temperature gradients. These temperature acclimation routes include cold and heat stress with a relationship to spatial and temporal variabilities. Some of these pathways include the synthesis of antifreeze proteins (AFPs; DeVries and Cheng 2005) and antifreeze glycoprotein (AFGPs; Harding *et al*., 2003), formation of low-temperature tubulin (Guderley, 2004), and lack of haemoglobin (Hemmingsen, 1991). Antarctic fishes have primarily developed proteins acting as antifreeze entities to prevent the freezing of fish in severely cold temperatures. Another reported pathway involves the deposition of glycerol to push downthe freezing threshold of body fluids without causing any physiological alarm. Driedzic and Short (2007) reported the extreme accumulation of glycerol in rainbow smelt by mobilization of liver-stored glycogenand a significant increase of glycerol-3-phosphate dehydrogenase (GPDH) gene expression by low temperature alone. GPDH genes have been previously classified in endemic Himalayan fishes and their role in

stress-related responses has been studied (Barat *et al*., 2012).

Now a days several *in-silico* approaches provide great prospects for the investigation of proteins to hasten experimental attempts and amplify scientific notions. These approaches pave way for economical ways to comprehend the physicochemical and structural properties of proteins for fruitful designing of biological experiments within a short range of time along with their functional description. Homology modelling helps in designing and assessing biological experiments involving proteins when three-dimensional (3D) structures from physical methods are not available. This is due to the large size of many proteins for experimental analysis like nuclear magnetic resonance (NMR) while others are difficult or impossible to be crystallized for the X-ray diffraction method. These reasons make *in silico* protein modelling the sole way to obtain structural information. Computational prediction of 3D structures provides us with important insights into the molecular mechanism of the involved protein. Therefore, there is an obvious knowledge gap that demands computational methods for protein structure prediction (Schwede *et al*., 2003). With this background and the importance of GPDH protein in fishes, the present study was devisedto analyse rainbow trout GPDH protein by subjecting it to bioinformatic analysis to assess its physicochemical properties, predict secondary and tertiary structure, predict motifs, predict interacting proteins and analyse functionality.

MATERIAL AND METHODS

Ethics statement. The present study did not perform any studies with human participants or animals.

Similarity search and physicochemical properties. The sequence of trout GPDH protein was accessed in FASTA format (Accession no. A0A8C7WA10) from UniProt (Wang *et al*., 2021), a comprehensive public domain resource for protein sequence and annotation data,and used for further analysis. To obtain the close homologs for the respective protein, the sequence similarity tool BLASTp (Altschul *et al*., 1990) was utilized. BLAST-p is an NCBI program that returns the closest protein sequences available in the database for a given query sequence. The functional analysis and classification of the GPDH amino acid sequence were performed by the InterPro resource server (Blum *et al*., 2021) which identifies families, superfamilies, domains, and functional sites of the query protein. The physicochemical characteristics of GPDH including isoelectric point (pI), molecular weight (MW), the total number of positive and negative residues, extinction coefficient (ε), instability index (II), aliphatic index (AI) and grand average hydropathy (GRAVY) were estimated using the Prot Paramtool of Expasy (Gasteiger *et al*., 2005).

Secondary structure and functional analysis. SOPMA (Self-Optimized Prediction Method with Alignment) was employed to estimate the secondary structure (2º) topographies of trout GPDH (Geourjon and Deleage 1995). SOSUI server (Hirokawa *et al*., 1998) was used to predict transmembrane regions. Computational methods including CYS_REC(http://www.softberry.com) and CYSPRED (Fariselli *et al*., 1999) servers were used for ascertaining the presence of any disulphide bonds (SS), which are indicative of functional relationship and the stability of proteins. Potential phosphorylation and glycosylation sites of trout GPDH were predicted using NetPhos3.1a (Blom *et al*., 1999) and NetNGlyc-1.0 servers (Gupta and Brunak 2001) respectively.

Modelling and validation

The modelling of the 3D structure of the trout GPDH protein was performed by the homology modelling program Swiss-Model (Waterhouse *et al*., 2018). Swiss-Model server is an automated comparative modelling tool for protein structures allowing a user to get a 3D structure automatically after submitting an amino acid sequence. After the selection ofa template using BLAST and HHbits, a raw model is made based on a rigid fragment assembly methodology. The quality of the predicted model and precision of the generated structure were gauged by Ramachandran plot analysis using PROCHECK (Laskowski *et al*., 1993) and QMEAN (Benkert *et al.,* 2011) methods. Additionally, The QMEANDisco (Studer *et al.,* 2020) was used to assess the local quality of the built model.

Homology modelling was also performed using a template structure from the Protein Data Bank (PDB; Berman *et al*., 2000) via a BLASTp search (Altschul *et al*., 1990). The modelled 3D structure was evaluated using the online servers ProQ (Wallner and Elofsson 2003) and ProSA (Wiederstein and Sippl 2007). ProQ is a neural network-based predictor for the model quality with outputs in the form of LG score and MaxSub. On the other hand, ProSA calculates an inclusive quality score (z-score) for a specific input PDB structure with scores outside the characteristic range of native proteins containing potential errors.

Protein-protein interactions. To generate an interaction network of trout GPDH, the STRINGv11.5 server (Search Tool for the Retrieval of Interacting Genes/Proteins; Szklarczyk *et al*. (2021) was used. STRING database contains a large number of known and predicted protein interactions covering 67,592,464 proteins from 14,094 organisms. These interactions include physical and functional associations and are derived from genomic studies, high-through put transcriptomic and proteomic experiments, and existing information. Protein scores greater than 0.6 were included in the results.

RESULTS AND DISCUSSION

BLAST-P hit table results revealed that the *O. mykiss* GPDH is most similar to the nearest relatives of the fish within the genus (Table 1). The selection was based on the lowest E-values and most identity percentile. The E-value integrates factors like sequence length, subject, and query similarity scores as well as database size to express the number of hits expected to get by chance.

The primary structure analysis was performedusing Expasy's ProtParam tool and different physicochemical parameters were computed. The molecular weight of the trout GPDH is 39.9 KDa with 365 amino acid residues. The computed isoelectric point (pI) of 6.68

indicates the acidic nature of trout GPDH. At the isoelectric point, a molecule does not carry any net charge and coincides with the pH where the protein is least soluble. Thus, at this point, the protein is most likely to precipitate out and shows total immobility in an electrophoresis system. The total number of negatively (Asp+Glu) and positively (Arg+Lys) charged residues in the trout GPDH were 42 and 40 respectively. The estimated extinction coefficient (EC) was 14440 $M^{-1}cm^{-1}$ with all Cys residues reduced and 15065 $M^{-1}cm^{-1}$ with all pairs of Cys residues forming cystines. The value is moderate due to the lesser number of cysteine, tryptophan and tyrosine residues in trout GPDH. The computed EC value will help in the quantitative study of GPDH interactions with other proteins and ligands in the solution. The computed instability index (II) was 27.19 signifying the protein as stable. This index indicates the stability of the protein *in vitro* with values smaller than 40 signifying a stable protein (Guruprasud *et al*., 1990). The aliphatic index (AI), defined as the relative size of a protein occupied by aliphatic side chains (A, V, I and L), was 102.58. It is regarded to have a positive influence on the increased heat stability of proteins (Ikai 1980); the higher the AI, the higher the stability of the corresponding protein. GRAVY represents the hydrophobicity value of the protein and is calculated by dividing the sum of the hydropathy values of all residues in the chain and their total number. In general, the lesser GRAVY value indicates better-quality interaction between the protein and water. The estimated GRAVY of trout GPDH was 0.100 suggesting that itmight be a hydrophilic (globular) protein owing to the presence of non-polar amino acids (Kyte and Doolittle 1982). The low GRAVY value also hints toward the cytoplasmic distribution of the trout GPDH rather than being membranous. The overall amino acid composition of the retrieved trout GPDH is presented in Table 2. Isoleucine is the most frequent amino acid present, followed by leucine and alanine. The presence of 15 aspartic acid residues in proteins is critical for interaction with the solvents, promoting protein stabilization and serving as binding sites for metal ions and charged ligands (Maniccia *et al*., 2009). In addition, 11 (3.0%) Cys residues present in trout GPDH imply the formation of disulphide bonds in its final 3D structure.

In nature, practically all of the proteins are structurally or functionally composed of one or more specific regions called domains. The varied combinations and arrangements of unique domains give rise to the diverse range of protein families found (Bateman *et al*., 2004). In the present study, functional analysis by InterPro classified the protein into NAD(P)-binding domain family having a biological role in the oxidation-reduction process. As with other dehydrogenases (Lesk 1995), the binding of the NAD to trout GPDH is expected to be involving numerous hydrogen bonds and van der Waals interactions. NAD-dependent GPDH catalyzes the reversible reduction of dihydroxyacetone phosphate (DHAP) to glycerol-3-phosphate in glycolysis (Otto *et al*., 1980). It is a cytoplasmic protein, active as a homodimer, each monomer containing an N-terminal NAD binding site. Otto *et al*. (1980) predicted the secondary structure of rabbit GPDH from its amino acid sequence and observed consistency with the properties of the NAD binding domain in other dehydrogenases in the first half of the sequence. The latter half of the sequence exhibited resemblances with the catalytic domain of glyceraldehyde-3-phosphate dehydrogenase (GAPDH). This dual-domain structure of trout GPDH allows favourable conditions for the specific interaction with B-nicotinamide and the substrate as glycerol 3-phosphate.

The functional characterization of trout GPDH protein including the prediction of a transmembrane helix and SS bonding pairs was also performed. The SOSUI functional analysis server used for the identification of transmembrane regions along with their consequent lengths classified the trout GAPDH as a soluble protein found free in cellular compartments such as the cytoplasm, nucleus, or endoplasmic reticulum. This was validated by the results of the $SOSUI_{GramN}$ server that indicated the cytoplasmic subcellular localization and absence of signal sequence. The secondary structure of GPDH was predicted using SOPMA which predicts amino acids for a three-state description of the secondary structure (alpha-helix, beta-sheet, and coil) prediction (Geourgon and Deleage 1995). The predicted secondary structure features using default parameters are listed in Table 3. The table shows that the trout GPDH secondary structure is dominated by Alpha helices (47.12%) formed when a polypeptide chain twists around on itself to form a rigid cylinder. The random coils might be due to the presence of hydrophobic proline in the protein that upsets the 2º structure by generating kinks in the chain (Saleem and Rajput 2020). This generates a conformation where the amino acid residues are oriented randomly while still being bonded to their neighbouring residues.

The online tool CYS_REC identifies the occurrence of SS bonds and possible bonding pairs among all cystein residues in a sequence. As an output, CYS_REC yields the number and location of cysteins and the pattern of cystein pairs in the given sequence. On the other hand, CYSPRED is a neural network-based predictor that characterises the bonding states of cysteine in proteins beginning from the residue sequence. While CYS_REC did not identify any SS bonds with high probability in trout GPDH, CYSPRED predicted only CYS4 residue in the SS bonding state. NetPhos3.1a predicted 19 potential phosphorylation sites suggesting the high activity of the GPDH protein (Fig. 1). The tool predicts Ser, Thr and Tyr phosphorylation sites in eukaryotic proteins using groups of neural networks. On the other hand, the NetNglyc-1.0 server identifies N-Glycosylation sites in proteins employing artificial neural networks to scan the sequence context of Asn-Xaa-Ser/Thr patches. NetNGlyc-1.0 tool did not predict any potential glycosylation sites in the trout GPDH protein.

Table 1: BLAST-P hit table results against rainbow trout GPDH protein.

Scientific Name	Total Score	E value	% Identity	Accession
Oncorhynchus tshawytscha	684	0	99.7	XP_024230975.1
Oncorhynchus kisutch	682	0	99.4	XP_020361607.1
Oncorhynchus nerka	681	0	99.4	XP_029519881.1
Oncorhynchus keta	681	0	99.11	XP_035597128.1
Oncorhynchus gorbuscha	681	0	99.11	XP_046199302.1
Salvelinus namaycush	678	0	98.51	XP_038864488.1
Salvelinus alpinus	677	0	98.51	XP_023846499.1
Salmo salar	677	0	98.21	XP_014001025.1

Table 2: Amino acid composition of Trout GPDH protein (in percentage) computed using ProtParam tool.

Amino acid	Number	Percentage
Alanine (A)	31	8.5%
Arginine (R)	9	2.5%
Asparagine (N)	19	5.2%
Aspartic Acid (D)	15	4.1%
Cysteine (C)	11	3.0%
Glutamine (Q)	9	2.5%
Glutamic Acid (E)	27	7.4%
Glycine (G)	29	7.9%
Histidine (H)	12	3.3%
Isoleucine (I)	33	9.0%
Leucine (L)	32	8.8%
Lysine (K)	31	8.5%
Methionine (M)	10	2.7%
Phenylalanine (F)	15	4.1%
Proline (P)	12	3.3%
Serine (S)	13	3.6%
Threonine (T)	19	5.2%
Tryptophan (W)	1	0.3%
Tyrosine (Y)	6	1.6%
Valine (V)	31	8.5%

Table 3: Calculated secondary structure features by SOPMA.

Feature	Residues covered	Percentage covered
Alpha helix (Hh)	172	47.12
Extended strand (Ee)	57	15.62
Beta turn (Tt)	33	9.04
Random coil (Cc)	103	28.22
Other states	0	0.00

3D protein structures and homology modelling delivervaluable insights into the molecular foundations of protein function, permitting an efficient design of experimentations. Different web-based homology modelling servers can predict the 3D structure of proteinsat various levels of complexity from their amino acid sequences. During evolution, the 3D protein structure is generally stable and undergoes much fewer changes than the associated sequences. This leads to the adaptation of closely related sequences to virtually matching structures and nevertheless more distantly related sequences fold into comparable tertiary and quaternary structures (Chothia and Lesk 1986). Thus, the homologous proteins have regions that maintain the same overall fold and regions where the foldsdiverge. In the present study, the modelling of the 3D structure of trout GPDH protein was executedby the homology modelling online package Swiss-Model. The final modelled structure was pictured by the Swiss PDB Viewer (Fig. 2).

The quality of the trout GPDH predicted model and the correctness of the structure generated were evaluated using Ramachandran plot analysis using PROCHECK and the QMEAN method. Ramachandran's plot exhibits the geometry of a 3D structure by displaying residue-wise torsion angles. Ramachandran plot analysis of trout GPDH showed that 96.83% of all residues were in the most favoured quartet (> 90% cut-off for model reliability proposed by Lovell *et al*., 2003), with only 0.72% of residues in the outlier region demonstrating the suitability of the model (Fig. 3). Earlier many studies have reported the most favoured residues as 91.9% (Singh *et al*., 2009), 92.7 (Tran *et al*., 2015), 88.8% (Sahay *et al*., 2020) and 85.7% (Ashraf *et al*., 2022) for different proteins. The global quality of the model was assessed by QMEAN (Qualitative Model Energy Analysis) score given by the SWISS-MODEL server. The QMEAN value of 0.84 points towards thegood quality of the global 3D structure of trout GPDH (Fig. 4).

The local quality value of individual residues (Fig. 5) showed that most regions of the 3D structure were modelled with good quality except residues near 145 and 265 which had moderate quality values.

The modelled structure of GPDH was also validated by the verification server Protein Quality (ProQ), which validates protein models based on different parameters. The predicted LGscore of the model was 9.707 suggesting an extremely good model (LGscore >4). ProSA was used to check the generated 3D model of trout GPDH for potential errors. The program displays a z-score and a plot of residue energies of the input model. The z-score denotes overall model quality by measuring the deviation of the total energy of the built 3D structure from random structural conformations. As presented in Fig. 6a, the z-score for the predicted GPDH structure was on the scaleof scores for proteins of comparable sizes signifying a very reliable structure. In addition, Fig. 6b demonstrates a comparable energy plot for both the target and template structures with very fewer positive values. Earlier, many similar *in silico* homology modelling studies have been reported for different proteins in different species (Hossain 2012; Chang and Yang 2013; Enany 2014; Saleem and Rajput 2020).

Investigation of protein-protein interactions is critical to understand protein function and biology as the vast majority of proteins interact with other proteins for proper biological activity. In this regard, Protein-protein interaction databases have become a major resource for deciphering these biological networks and pathways in cells as experimental attempts are challenging concerning their cost, technicality and procedure. For functional protein association networks of trout GPDH, STRINGv11.5 was used for the prediction of interaction with other partners. Only the partner proteins with a score of more than 0.6 were included in the results. The nearest interacting proteins with the shortest nodes were found to be belonging to the *GPAT/DAPAT* (Glycerol-3-phosphate acyltransferase) family having a score of >0.7. These proteins contain *PlsC* domains predicted to enable acyltransferase activity. In addition, *LNS2* domains have been found in lipins and lipin homologues from *Saccharomyces cerevisiae* and are involved in one of the steps in triacylglycerol synthesis (Csaki *et al*., 2013). The other potential interacting proteins associated with trout GPDH protein are listed in (Fig. 7). Notably, many of these partners have not been characterized yet in rainbow trout depicting the scope for further research in the area. Elucidating these types of convoluted protein interactions will give vitalclues as to the working and molecular dynamics of novel proteins that govern cell behaviour.

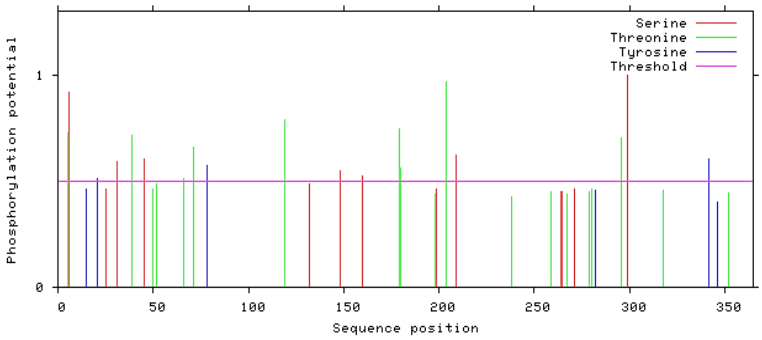

Fig. 1. Net Phos3.1a predicted phosphorylation sites in rainbow trout GPDH.

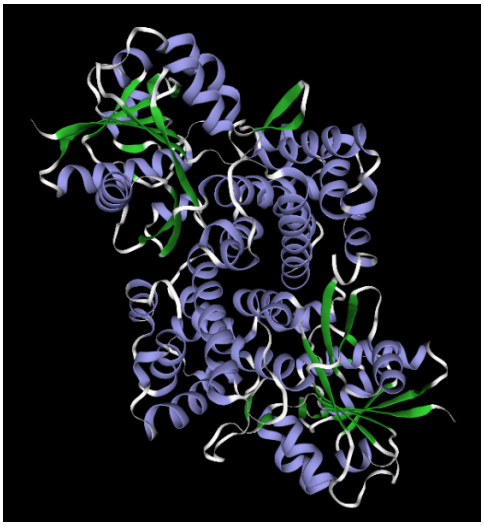

Fig. 2. 3D structure of rainbow trout GPDH. Secondary structure features are highlighted in different colours.

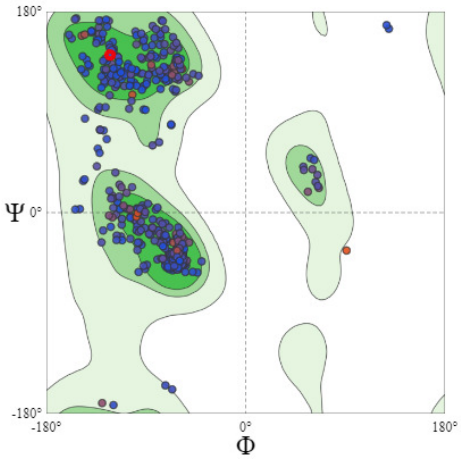

Fig. 3. Ramachandran plot analysis rainbow trout GAPDH 3D model structure.

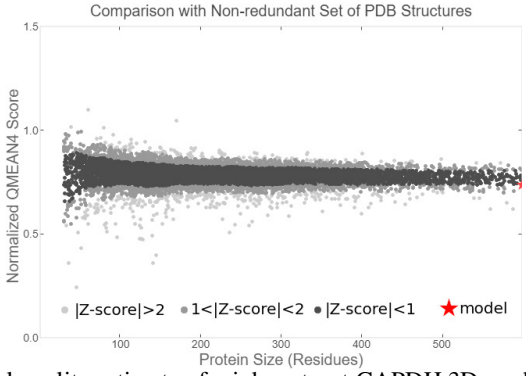

Fig. 4. Global quality estimate of rainbow trout GAPDH 3D model structure.

Fig. 5. Local quality estimate rainbow trout GAPDH 3D model structure.

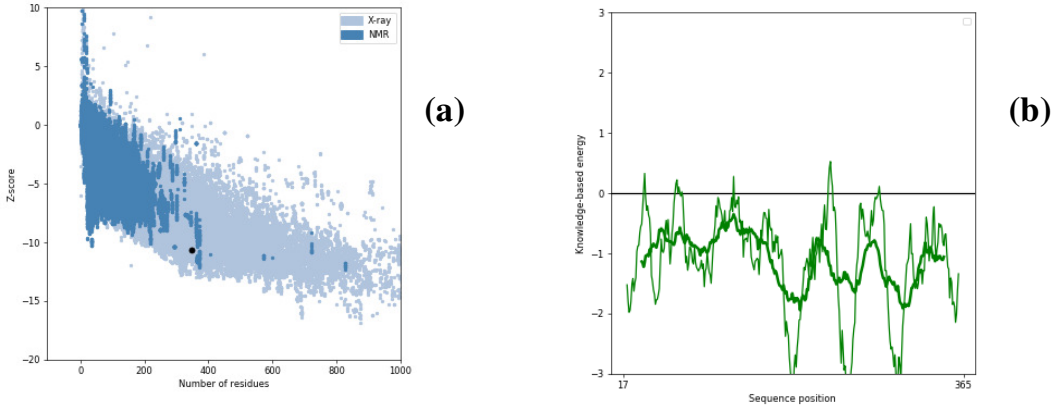

Fig. 6. ProSA output of rainbow trout GPDH model. **(a)** z-scores of all proteins in PDB resolved by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) corresponding to their lengths. The z-score of trout GPDH isemphasisedas a large black dot. **(b)** Energy plot (local model quality) of trout GPDH.

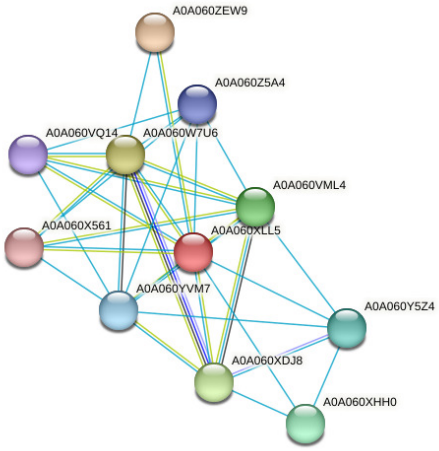

Fig. 7. Protein-protein interfacemap of trout GPDH (A0A060XLL5) by STRINGv11.5 server. The closest interacting protein is at the nearest node.

CONCLUSION

In conclusion, genomic and proteomic sequencing projects have resulted in many protein sequences without any annotated functions or designated hypothetical/conserved proteins. The GPDH gene in rainbow trout plays a crucial role in cold stress physiology. In snow trouts of Himalayan rivers, Barat *et al*. (2012) characterized 1,012 bp of the GPDH gene from *Schizothorax richardsonii* and *S. niger* and showed a 19-fold increase in its expression in liver tissue at 5ºC compared to 15 ºC conditions suggesting a role in stress-linked responses in these endemic Himalayan fishes. However, the protein has not been physically characterized using NMR or X-ray crystallography. Consequently, it becomes imperative to utilise bioinformatic means to fathom these aspects of GPDH as well as other critical proteins in rainbow trout.

FUTURE SCOPE

Understanding the structural aspects of this protein will lead to a better evaluation of cold resistance in reported fishes. This may also have an impact on the transgenesis research in fisheries aimed to induce cold tolerance in otherwise warm water fishes.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403-410.

Ashraf, S., Nakkeeran, S., Saranya, N., Jothi, G., Gulsar Banu, J., Mohankumar, S., Saravanan, R., Mahendra, K. and Krishna Nayana, R. U. (2022). Computational analysis reveals 10-acetyl-9,10-dihydroacridine as a novel biomolecule from *Bacillus licheniformis* (MW301654) possessing nematicidal property against banana root knot nematode *Meloidogyne incognita*. *Biological Forum–An International Journal*, *14*(3), 363-370.

Blum, M., Chang, H. Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A. and Finn, R. D. (2021). The InterPro protein families and domains database: 20 years on. *Nucleic acids research*, *49*(D1), D344-D354.

Blom, N., Gammeltoft, S. and Brunak, S. (1999). Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of molecular biology*, *294*(5), 1351-1362.

Brett, J. R. (1979). Environmental factors and growth. *Fish physiology, vol. VIII. Bioenergetics and growth*, 599-677.

Benkert, P., Biasini, M., and Schwede, T. (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics*, *27*(3), 343-350.

Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S. and Eddy, S. R. (2004). The Pfam protein families database. *Nucleic acids research*, *32*(suppl_1), D138-D141.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H. and Bourne, P. E. (2000). The protein data bank. *Nucleic acids research*, *28*(1), 235-242.

Barat, A., Goel, C., Tyagi, A., Ali, S. and Sahoo, P. K. (2012). Molecular cloning and expression profile of snow trout GPDH gene in response to abiotic stress. *Molecular biology reports*, *39*(12), 10843-10849.

Chothia, C. and Lesk, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *The EMBO journal*, *5*(4), 823-826.

Chang, K. Y. and Yang, J. R. (2013). Analysis and prediction of highly effective antiviral peptides based on random forests. *PloS one*, *8*(8), e70166.

Csaki, L. S., Dwyer, J. R., Fong, L. G., Tontonoz, P., Young, S. G. and Reue, K. (2013). Lipins, lipinopathies, and the modulation of cellular lipid storage and signaling. *Progress in lipid research*, *52*(3), 305-316.

Driedzic, W. R. and Short, C. E. (2007). Relationship between food availability, glycerol and glycogen levels in low-temperature challenged rainbow smelt *Osmerus mordax*. *Journal of Experimental Biology*, *210*(16), 2866-2872.

DCFR (2015). Rainbow trout farming in India: R&D Perspectives. Bulletin No. 23,33.

DeVries, A. L. and Cheng, C. H. C. (2005). Antifreeze proteins and organismal freezing avoidance in polar fishes. *Fish physiology*, *22*, 155-201.

Enany, S. (2014). Structural and functional analysis of hypothetical and conserved proteins of *Clostridium tetani*. *Journal of infection and public health*, *7*(4), 296-307.

Fariselli, P., Riccobelli, P. and Casadio, R. (1999). Role of evolutionary information in predicting the disulfide-bonding state of cysteine in proteins. *Proteins: Structure, Function, and Bioinformatics*, *36*(3), 340-346.

FAO (2020). The state of world fisheries and aquaculture 2020. Sustainability in action. Rome, 266.

Gupta, R. and Brunak, S. (2001). Prediction of glycosylation across the human proteome and the correlation to protein function.

Guderley, H. (2004). Metabolic responses to low temperature in fish muscle. *Biological reviews*, *79*(2), 409-427.

Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server. *The proteomics protocols handbook*, 571-607.

Geourjon, C. and Deleage, G. (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics*, *11*(6), 681-684.

Guruprasad, K., Reddy, B. B. and Pandit, M. W. (1990). Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Engineering, Design and Selection*, *4*(2), 155-161.

Harding, M. M., Anderberg, P. I., and Haymet, A. D. J. (2003). 'Antifreeze'glycoproteins from polar fish. *European Journal of Biochemistry*, *270*(7), 1381-1392.

Hossain, M. M. (2012). Fish antifreeze proteins: Computational analysis and physicochemical characterization. *International Current Pharmaceutical Journal*, *1*(2), 18-26.

Hemmingsen, E. A. (1991). Respiratory and cardiovascular adaptations in hemoglobin-free fish: resolved and unresolved problems. In *Biology of Antarctic fish* (pp. 191-203). Springer, Berlin, Heidelberg.

Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998). SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics (Oxford, England)*, *14*(4), 378-379.

Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *The Journal of Biochemistry*, *88*(6), 1895-1898.

Kyte, J. and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology*, *157*(1), 105-132.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*, *26*(2), 283-291.

Lesk, A. M. (1995). NAD-binding domains of dehydrogenases. *Current opinion in structural biology*, *5*(6), 775-783.

Lovell, S. C., Davis, I. W., Arendall III, W. B., De Bakker, P. I., Word, J. M., Prisant, M. G. and Richardson, D. C. (2003). Structure validation by Cα geometry: □, ψ and Cβ deviation. *Proteins: Structure, Function, and Bioinformatics*, *50*(3), 437-450.

Maniccia, A. W., Yang, W., Johnson, J. A., Li, S., Tjong, H., Zhou, H. X. and Yang, J. J. (2009). Inverse tuning of metal binding affinity and protein stability by altering charged coordination residues in designed calcium binding proteins. *PMC biophysics*, *2*(1), 1-16.

Otto, J., Argos, P. and Rossmann, M. G. (1980). Peredriction of Secondary Structural Elements in Glycerol-3-Phosphate Dehydrogenase by Comparison with Other Dehydrogenases. *European Journal of Biochemistry*, *109*(2), 325-330.

Saleem, A. and Rajput, S. (2020). Insights from the in silico structural, functional and phylogenetic characterization of canine lysyl oxidase protein. *Journal of Genetic Engineering and Biotechnology*, *18*(1), 1-11.

Singh, B., Gupta, A. and Potukuchi, S. M. (2009). Comparative modeling and analysis of 3-D structure of Hsp 70, in *Cancer irroratus*. *Biological Forum - An International Journal, 1*(2): 1-4.

Sahay, A., Piprodhe, A. and Pise, M. (2020). In silico analysis and homology modeling of strictosidine synthase involved in alkaloid biosynthesis in *Catharanthus roseus*. *Journal of Genetic Engineering and Biotechnology*, *18*(1), 1-6.

Studer, G., Rempfer, C., Waterhouse, A. M., Gumienny, R., Haas, J. and Schwede, T. (2020). QMEANDisCo-distance constraints applied on model quality estimation. *Bioinformatics*, *36*(6), 1765-1771.

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S. and von Mering, C. (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*, *49*(D1), D605-D612.

Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic acids research*, *31*(13), 3381-3385.

Tran, N. T., Jakovlić, I. and Wang, W. M. (2015). In silico characterisation, homology modelling and structure-based functional annotation of blunt snout bream (*Megalobrama amblycephala*) Hsp70 and Hsc70 proteins. *Journal of Animal Science and Technology*, *57*(1), 1-9.

Welcomme, R. L. (Ed.). (1988). *International introductions of inland aquatic species* (Vol. 294). Food and Agriculture Org. pp318.

Wang, Y., Wang, Q., Huang, H., Huang, W., Chen, Y., McGarvey, P. B. and UniProt Consortium. (2021). A crowdsourcing open platform for literature curation in UniProt. *PLoS Biology*, *19*(12), e3001464.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R. and Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic acids research*, *46*(W1), W296-W303.

Wallner, B. and Elofsson, A. (2003). Can correct protein models be identified? *Protein science*, *12*(5), 1073-1086.

Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic acids research*, *35*(suppl_2), W407-W410.

How to cite this article: Ubaid Qayoom, Zahoor Mushtaq and Pabitra Barik (2022). *In silico* Structural Prediction and Functional Analysis of Cold Tolerant Glycerol-3-Phosphate Dehydrogenase Protein of Rainbow Trout. *Biological Forum – An International Journal, 14*(4): 811-818.