Symbiosis

www.symbiosisonlinepublishing.com
ISSN Online: 2374-8362

Research Article | Journal of Advanced Research in Biotechnology | Open Access

Structural and Functional Analysis of Ferritin Heavy Chain Subunit in *Oryzias latipes*

Ajit Tiwari[1*], A D Upadhyay[1], Himanshu Priyadarshi[2], A K Roy[1], Rumpi Ghosh[1], Suresh Yambem[2] and Dibyajyoti Uttameswar Behera[2]

[1]*Bioinformatics Centre, College of Fisheries, CAU (I), Lembucherra, Tripura*

[2]*Dept of Fish Genetics and Reproduction (FGR), College of Fisheries, CAU (I), Lembucherra, Tripura*

Received: June 25, 2019; Accepted: July 12, 2019; Published: July 22, 2019

***Corresponding author:** Ajit Tiwari, Bioinformatics Centre, College of Fisheries, CAU (I), Lembucherra, Tripura.*

Abstract

In North East region of India, iron toxicity is one of the major problems in culture fisheries. To overcome this challenge, it is necessary to identify the role of ferritin protein as an iron detoxificant and store in fishes. The Heavy chain in a ferritin protein possess di-Fe binding site in the fourth helix that interacts with oxygen. In this paper attempt has been made to study the structure and function of ferritin heavy chain subunit of *Oryzias latipes* from amino acid sequence. Physicochemical characterization by Expasy ProtParam form tools reveals that the protein is acidic, unstable and hydrophilic. A hydropathy scale showed two peak with significant score above the threshold value (0 to + value) but TMHMM conclude that there were one transmembrane domain within protein. The secondary structures contain alpha helix (56.50%), extended strands (10.73%) and coiled region (32.77%). The query sequence shows homology to the selected template (structure of mouse heavy chain modified ferritin by X-ray diffraction technique) with maximum % identity. To analyse the phylogenetic relationship, ML tree was constructed between *Oryzias latipes* and *Oryzias melastigma* for the ferritin heavy sub unit along with Cyprinus carpio as out crossed. Two distinct clads formation was observed between same genus species and the out crossed species. In protein-protein interaction analysis via STRING 10.0 tool, two enriched pathways of KEGG, six Inter Pro domains, one PFAM protein domain, one Uniprot keywords and eleven functional parameters of network analysis were identified in *Oryzias latipes*. The overall investigation reveals the structural features and their association in detoxification and iron homeostasis.

Keywords: Ferritin Heavy Chain; *Oryzias latipes*; Physicochemical Characterization; Homology Modeling

Introduction

Iron, required as a trace element in living organisms, forms metallo protein in conjugation with different proteins. However, excess iron in an aquatic ecosystem become toxic and acts as a catalyst in the Fenton reaction generating free radicals which are harmful to fishes. [1] Ferritin is the ubiquitous protein which stores in a nontoxic and reversible form and central to iron metabolism [2]. Thus, it has an important role in iron storage and detoxification [3, 4, 5]. Ferritin is structurally highly conserved protein from bacteria to human which emphasize their biological importance [6]. It forms a hollow shell with a cavity of 80 A1 diameters that can store up to 4,500 Fe (III) atoms as a biomineral. In eukaryotes, it consists of 24 protein subunits with molecular weight of 450 kDa [7, 8, 9]. In mammals, ferritin molecules of heavy (H) and light (L) chain subunits having molecular masses of 21 and 19 kDa. Heavy chains are important for Fe (II) oxidation, whereas Light chains assist in iron nucleation, mineralization, and long-term storage [9]. The Heavy and Light chain subunits co-assemble in different ratios to form a protein shell of 24 subunits capable of acquiring iron atoms. Key features that differentiate di-Fe binding site in H chain that interacts the fourth helix oxygen and is responsible for the ferroxidase activity of the protein [10]. For example, H-chains are abundant in heart tissues that involved in a rapid exchange of iron [11]. The dual role of ferritin in sea bass, with functional involvement in both iron metabolism and immune response, was also reported by the researcher [12].

Ferritin heavy chain subunit with the variable sequence is generally used for determining phylogenetic relationships among different organisms. It is considered to be useful in determining relationships within families and genera. A Comparative analysis generates evolutionary relationship and new classification schemes. The vertebrates mitochondrial DNA are more polymorphic and more useful for the identification of species and can evolve faster in comparison to nuclear DNA [13].

The Medaka fish, scientific name *Oryzias latipes*, is a small, bony, laying an egg in fresh water, native to Asian countries. It occurs, coastal waters having high adaptability and it collected in wide range especially from brackish, mangrove swamps, acidic freshwater, forest streams, canals, rice field, basins of rivers, pools, and oxbows [14]. The Medaka is a model organism used in the area of genomics, genetics, disease model, sex determination, reproduction and evolution. In present study, the structural model, protein-protein interaction and physicochemical properties of ferritin H chain protein sequence (accession number XP_020569048.2.) of *Oryzias latipes* were analyzed to determine the structural and functional role in Fishes.

Material and Methods

Sequence retrieval

The amino acid sequence of Ferritin heavy subunit of model fish *Oryzias latipes* was retrieved from National Center for Biotechnology Information (NCBI) database having the accession number XP_020569048.2. It was verified by peptide search in the UniProt KB and found entry no. H2LMW5.

Physicochemical properties and Secondary structure prediction

The physicochemical properties including molecular weight, theoretical pI, the % of the amino acid composition, total number of residue in negative and positive form, extinction coefficient, estimated half-life, atomic composition, aliphatic index, instability index grand average of hydropath city of the ferritin heavy chain protein of *Oryzias latipes* [15] was estimated using Expasy ProtParam server.

The secondary structures of the ferritin heavy chain subunit were predicted by PSIPRED [16] and GORIV methodology [17].

Prediction via ProtScale

The protein sequence was input in the FASTA format and the amino acid scale selected was Hphob /Kyte & Doolittle [18] with the window size of 19 as detection of hydrophobic, membrane-spanning domains is best suited at this 30 window. The ProtScale tool is available at the bioinformatics resource portal ExPASy (Expert Protein Analysis System).

Prediction via TMHMM

TMHMM is an online tool used to predict membrane protein topology based on a hidden Markov model [19]. It predicts transmembrane helices and can discriminate between soluble and membrane proteins with high degree of accuracy. It can correctly predict 97-98 % of the transmembrane helices. The FASTA sequence of the query protein was used as input with all the parameters set to default. TMHMM Server version 2.0 was used (Figure 5.0).

Tertiary structure prediction

Tertiary structure of ferritin was analyzed with a template search for the query protein through PDB sum database [20]. This result was cross verified by SWISSMODEL/Workspace, which displayed sequence identity with the query sequence [21]. By homology modeling, 3-D structure of ferritin heavy subunit was predicted through Swiss model [22, 23], Phyre 2 [24] and pymol software [25] and Raptor X [26]. The predicted structure of the ferritin heavy subunit was validated through Ramachandran plot by utilizing rampage server [27].

Phylogenetic analysis

Phylogenetic analysis of fth based on amino acid sequences was carried out using the software Molecular Evolutionary Genetic Analysis (MEGA; version 7) [28]. Sequences were aligned by Muscle method, and the evolutionary relationship was established among different fish species by Neighbor-joining (NJ) with complete gap deletion, Poisson substitution model, rates among sites (uniform rates), the pattern among lineages (Homogeneous) and 1000 bootstrap replications.

Protein-protein interactions

The Search Tool for the Retrieval of Interacting Genes/ Proteins (STRING 10.0) database (http://string-db.org/) was used to predict the interacting proteins [29]. The database contains information from different sources, including experimental repositories, computational prediction methods, and public text groups.

Results

Physicochemical properties

A physicochemical property of Ferritin heavy chain subunit of *Oryzias latipes* was analyzed by the Expasy ProtParam server. It was 177 amino acid long proteins with the estimated molecular weight 20880.39 kDa respectively. The isoelectric point (pI) of ferritin protein was 5.54 which revealed that ferritin heavy chain was acidic. The amino acid composition showed the maximum presence of Leucine (11.3%) and minimum presence of Tryptophan (1.1%) (Table 1). The total number of negatively charged and positively charged residues of Ferritin are (Asp+

Table 1: Amino acid composition of Ferritin heavy chain subunit of *Oryzias latipes*

Amino acids	No. s	Percentage
Ala (A)	10	5.6%
Arg (R	10	5.6%
Asn (N)	11	6.2%
Asp (D)	13	7.3%
Cys (C)	5	2.8%
Gln (Q)	12	6.8%
Glu (E)	17	9.6%
Gly (G)	9	5.1%
His (H)	10	5.6%
Ile (I)	6	3.4%
Leu (L)	20	11.3%
Lys (K)	11	6.2%
Met (M)	6	3.4%
Phe (F)	8	4.5%
Pro (P)	3	1.7%
Ser (S)	9	5.1%
Thr (T)	3	1.7%
Trp (W)	2	1.1%
Tyr (Y)	7	4.0%
Val (V)	5	2.8%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%

Glu)-30 and (Arg+ Lys)-21, respectively. The formula of ferritin heavy subunit is C912H1399N263O280S11 with 2865 atoms in total. The extinction coefficient of the protein calculated at 280 nm in water (M-1 cm^{-1}) which presented the values of Abs 0.1% (=1g/l) 1.038 and extinction coefficient 21680 and, assuming all pairs of Cys residues form cystines; but the extinction coefficient becomes 21430 and Abs 0.1% (=1g/l) is 1.026 when assumed all Cys residues reduced. The N-terminal sequence of the protein is Methionine (Met). The estimated half-life of Ferritin was 30 hours. The estimated instability index (ll) of the Ferritin was 49.37 which classify the protein as unstable. Aliphatic index 71.13 of the protein measures its considerable thermo stability along with the relative volume occupied by aliphatic side chains. The value -0.823 of the GRAVY indicates the hydrophilic nature of the protein (Table 2).

Table 2: The result of primary structure analysis and secondary structure prediction of ferritin heavy chain protein in fish *Oryzias latipes*

S. No.	Structure element	
Scientific Name		Oryzias latipes
Abbreviation		Orl_fth
Primary structure analysis		Percent
Tools	**Parameters**	
ProtParam	Number of amino acid (aa)	177
	Molecular weight (Mw)	20880.39
	Theoretical iso electric point (pI)	5.54
	Total Number of negatively charged residues (Asp + Glu)	30
	Total Number of positively charged residues (Arg + Lys)	21
	Instability index	49.37
	Aliphatic index	71.13
	GRAVY	-0.823
Secondary structure prediction		
Tools	**Parameters**	
PSIRPRED, GOR IV	Alpha helix	56.50
	3_{10} helix	0.00
	Beta bridge	0.00
	Extended strand	10.73
	Beta turn	0.00
	Bend region	0.00
	Random coil	32.77
	Ambiguous states	0.00
	Other states	0.00

Prediction of Transmembrane Segments

Prediction via ProtScale

The presence of a transmembrane segment in the protein ferritin heavy chain in *Oryzias latipes* was confirmed by plotting the hydropathy index. A hydropathy scale which is based on the hydrophobic and hydrophilic properties of the 20 amino acids is used. A moving window was used to determine the summed hydropathy at each point in the sequence (Y coordinate). These sums are then plotted against their respective positions (X coordinate).The resulting plot revealed the relative hydrophobicity of segments of the protein (Fig 1). A window size of 19 amino acid was used to plot the index. The image seen in figure 2 is the hydrophobicity plot returned by ProtScale using the Kyte& Doolittle Scale [30]. Since a large window size of 19 was selected for finding transmembrane domains, values above 1.6 were considered to be significant [30].

The peaks in the plot are predicted to be the potential transmembrane regions present within the protein over a span of 177 amino acids. The higher the peak, the higher is the hydrophobicity of the region which indicates those regions are buried in the non-polar phase of the lipid membrane, which can therefore be said to be transmembrane regions. It can be seen from the plot that there are two peaks with significant score above the threshold value. The highest score was observed for the first peak which means it is the most hydrophobic and this region also contains the most number of amino acids than the

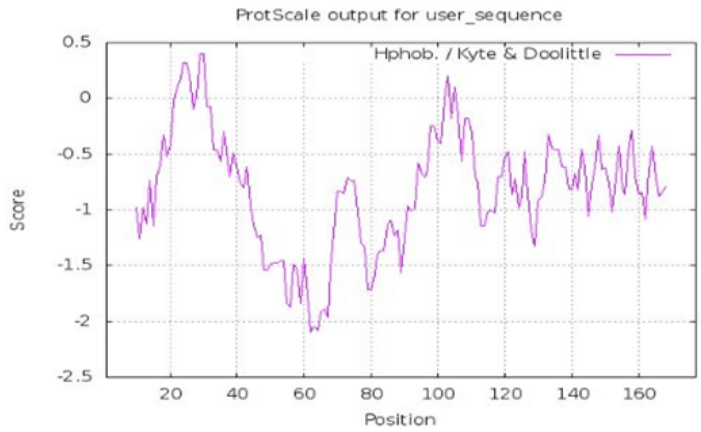

Figure 1: Protscale output of ferritin heavy chain subunit in *Oryzias latipes*

other three peaks because the base of the peak was wider than the other peak. Thus it can be concluded that there were two transmembrane regions with one smallest peak lie between 0 to 0.2 with least number of amino acid within the protein according to graphical representation.

Prediction via TMHMM

The TMHMM is much more advanced with more detailed and better graphical representation of the predicted transmembrane regions. The graphical representation produced by TMHMM revealed the putative transmembrane regions within the target protein (Fig 2)

It can be seen from fig.2, the number of predicted transmembrane helices is 1which confirms the results from ProtScale. The expected number of amino acids in transmembrane helices is 1.94794. The expected number of amino acids in transmembrane helices in the first 60 amino acids of the protein is 1.94794. The total probability that the N-term is on the cytoplasmic side of the membrane is 0.025531.The region of the first transmembrane helix is from 21 to 38 amino acids. Also, the graphical representation in fig 2 shows peaks indicating one transmembrane domain "INLELYASYVYLSMGYYF". The blue lines indicate the region of the protein that is inside the membrane whereas the pink lines indicate the regions that are outside the membrane. The red lines indicate the transmembrane regions.

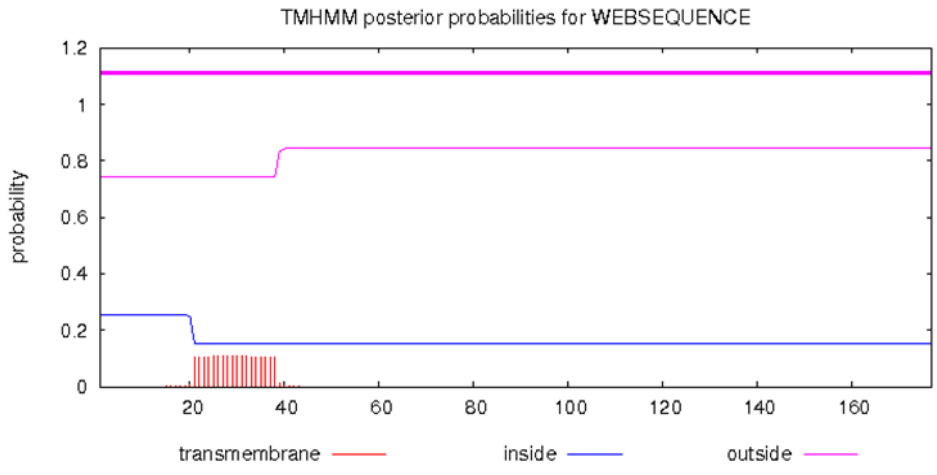

Figure 2: Transmembrane prediction of ferritin heavy chain subunit in *Oryzias latipes* via TMHMM

Secondary structure prediction

The secondary structure prediction of Ferritin heavy chain subunit was done by PSIPRED server and GOR IV server. The obtained results revealed the presence of alpha helix 56.50%, extended strands 10.73% and coiled region 32.77% (Table 2). Good helix former Ala, Glu, Leu and Met were 29.9% of total amino acid while poor helix former Pro, Gly, Tyr, and Ser were 15.9 % of total amino acid.

Tertiary Structure Prediction

Template identification was done by comparative modeling. It is usually initiated by the searching of sequence for known protein structures where target sequence used as a query in the PDB database to locate the sequences that so remotely related with render construction of a reliable comparative model. It is generally done by the comparative study of the target sequence with the sequence of each of the structures in the given database [31].

Template search for the query protein of ferritin heavy chain subunit was performed through PDB sum database which presented 338 hits. The template 3wnw (A) was identified showing 82.5% sequence similarity along with Z-score of 1225.7 from the PDB sum database (Table 3). The selected template contains a structure of mouse h-chain modified ferritin by X-ray diffraction technique (2.24Å). It was verified by SWISSMODEL/ Workspace possessing pdb code 3wnw.1.A with 82.46 % sequence identity with the query sequence (Table 3)

The structure predicted by SWISS-MODEL with the alpha helix 56.50%, extended strands and coiled region further visualized by Pymol and raptor X. The modeled structure has GMQE score0.94And QMEAN score 0.26 (Table 4). The predicted structure (Fig 3.1 and Fig 3.2) of the ferritin heavy chain subunit was validated through the Ramachandran plot (phi/psi). The stereo chemical analysis of RAMPAGE server showed the number of residues in the favored region is 95.6%, the allowed region are 4.4%, and Outlier region are0% respectively (Fig 4).

Table 3: Summary of the result obtained by Swiss Model using as templates the homologous structure

Model No.	Template	Sequence Identity	Oligo state	Source	Method	Resolution	Length	Header	Description
1	3wnw	82.5	Hom24mer	PDBsum	X-ray	2.24 A	172	Oxidoreductase	Structure of mouse h-chain modified ferritin
2	3mnw	82.46	Hom24mer	Swiss-model	X-ray	2.24 A	172	Oxidoreductase	Structure of mouse h-chain modified ferritin

Table 4: QMEAN score for Ferritin heavy chain modelled protein

Parameters	Value for the predicted model
Cβ interaction energy	2.88
All-atom pairwise energy	3.10
Solvation energy	2.16
Torsion angle energy	-1.00
QMEAN-score	0.26
GMQE	0.94

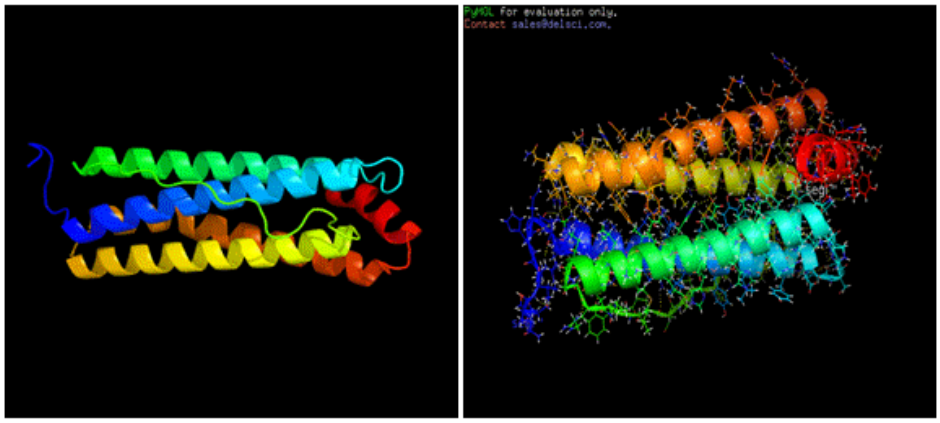

Figure 3.1: Structure view of ferritin heavy chain subunit of *Oryzias latipes*
Figure 3.2: Structure view of polar contacts with side chain, main chain in ferritin heavy chain subunit of *Oryziae latipes*
Formal charges: sum = -7.0
Count atoms: 2782 atoms

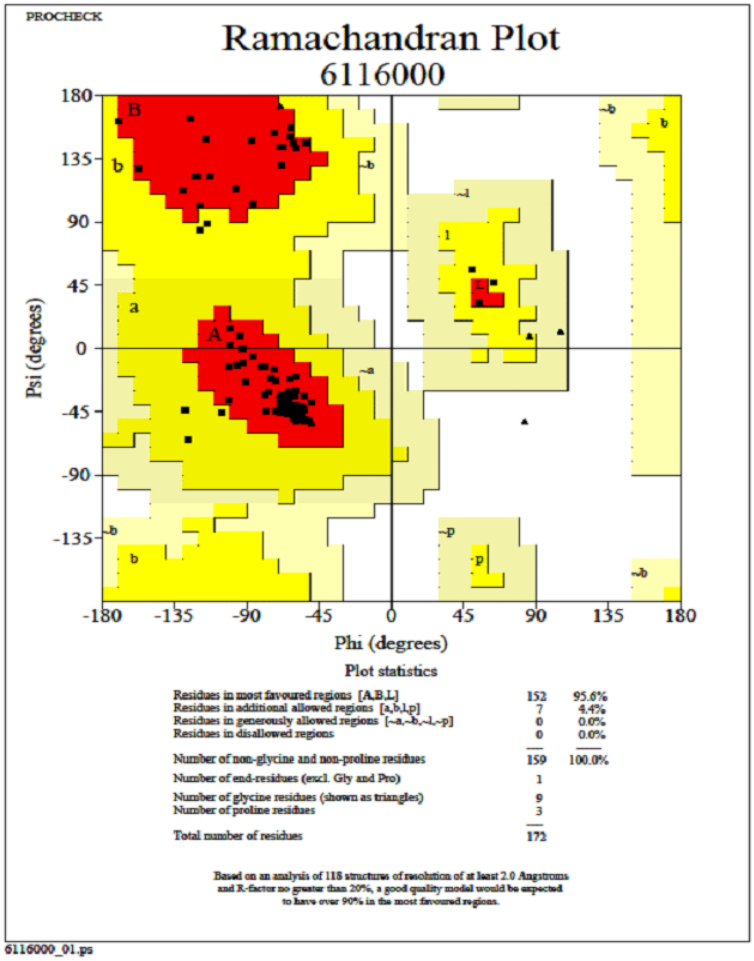

Figure 4: Ramachandran plot of the predicted structure of Ferritin heavy chain subunit of *Oryzias latipes*

Ferritin consist of 24 peptide subunits which form two types of channels where these subunits intersect; the 3-fold channel was found polar and the 4-fold channel was non polar. (The residues that line the channels determine the polarity of the channel. The electro negativity of Iron (Fe) (1.83 eV) and magnesium (1.31 eV) indicate electro negativity difference of 0.52 eV. The polar covalent bond was found. The sum of formal charges was -7 showing highly electronegative. For small fractions of charges, we use the symbols δ+ and δ−. Polar molecules have slightly opposite charges on opposite ends of the molecule or a dipole. When the Fe (III) in the crystalline mineral was reduced to Fe (II), the iron becomes solvated and ferritin releases the solvated iron, Fe (H2O)6 2+, through the 3-fold polar channel. Hence, it can control the amount of available iron in the body, preventing iron disorders like anemia and iron overload. [32]

Phylogenetic analysis

Ferritin heavy subunits of 2 species belonging to same genera (*Oryzias latipes* and *Oryzias melastigma*) were analyzed with *Cyprinus carpio* (Out crossed) for the same gene (Fig 5). Maximum likelihood (ML) tree was constructed for the analysis of distantly related sequences. An overall genetic distance based on K2P distance of 0.162 was found. GC content, Conserved, Singleton and Variable siteswere shown in (Table 5, and 6). The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model. The tree with the highest log likelihood (-1200.11) is shown. The tree was drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 3 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 534 positions in the final dataset.

Functional interaction network analysis

In order to predict the interacting proteins, *Oryzias latipes* ferritin heavy chain protein was applied to the STRING 10.0 tool as the model fish samples. The database utilized here to determine the interaction of known and predicted protein. The interactions occur in the form of direct and indirect contacts. Two enriched KEGG pathways, six inter pro protein domain, one PFAM protein domain, one Uniprot keywords and eleven functional parameters of network analysis were determine by the protein-protein interaction analysis, enlisted in (Table 7 and 8) along with their

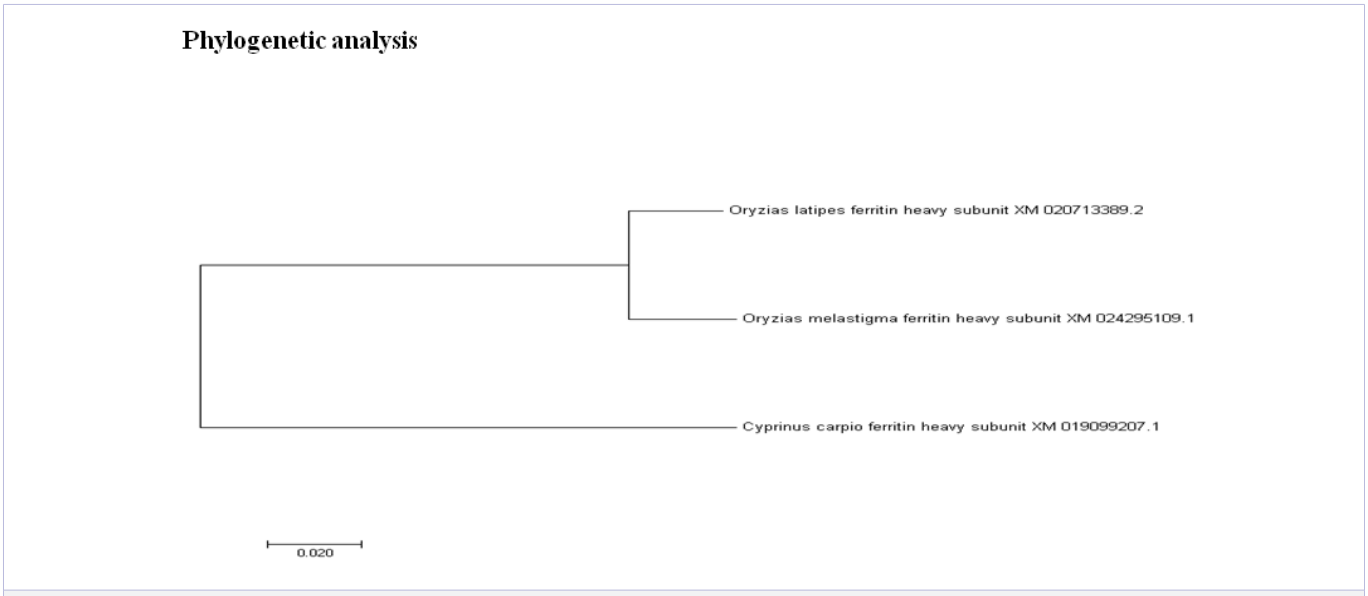

Figure 5: Maximum likelihood tree for *Phylogenetic* analysis

Table 5: Conserved, Singleton and Variable sites

Conserved site	Variable site	Singleton site	Zero fold degenerate site	Two fold degenerate site	Four fold degenerate site
423/534	16/77	111/534	339/534	111/534	52/534

Table 6: Nucleotide composition

T	A	G	C	GC1	GC2	GC3
20.9	27.6	27	24.5	53.2	40.5	60.7

Table 7: Characteristics of input protein *Oryzias latipes* ferritin heavy chain subunit (fth1, Accession number: XP_020569048.2) functional parameters predicted with STRING 10.0

S. No	Name	Functions	Score
1	ENSORL00000005872	Ferritin; Stores iron in a soluble, non-toxic, readily available form. Important for iron homeostasis. Iron is taken up in the ferrous form and deposited as ferric hydroxides after oxidation (177 aa)	
2	fth1	Ferritin; Stores iron in a soluble, non-toxic, readily available form. Important for iron homeostasis. Iron is taken up in the ferrous form and deposited as ferric hydroxides after oxidation	0.802
3	rps30	40S ribosomal protein S30; Finkel-Biskis-Reilly murine sarcoma virus (FBR-MuSV) ubiquitously expressed a; Belongs to the eukaryotic ribosomal protein eS30 family	0.604
4	Sod2	Superoxide dismutase; Destroys radicals which are normally produced within the cells and which are toxic to biological systems	0.591
5	ENSORL00000003269	Nuclear receptor co activator 4 (481 aa)	0.581
6	tfrc	Transferrin receptor 1a (764 aa)	0.578
7	fosl1	FOS-like antigen 1a (340 aa)	0.547
8	sf1	Uncharacterized protein; Splicing factor 1 (418 aa)	0.542
9	MAP3K11	Si-cabz01078036.1; Mitogen-activated protein kinase 11 (811 aa)	0.539
10	Fkbp2	Peptidylprolyl isomerase; FK506 binding protein 2 (139 aa)	0.539
11	ENSORL00000015509	Solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2 (554 aa)	0.530

Table 8: Observed KEGG pathways in protein-protein interaction of *Oryzias latipes* ferritin heavy chain subunit in **Fig 7**.

Index	ID	Term	Count in gene set	False discovery rate
KEGG Pathways				
1	ko04216	Ferroptosis	4 of 46	1.07e-07
2	ko04217	Necroptosis	2 of 122	0.0094
Uniprot Keywords				
3	KW-0409	Iron storage	2 of 3	5.66e-05
PFAM Protein Domain				
4	PF00210	Ferritin like domain	2 of 3	4.81e-05
INTERPRO Protein Domains and Features				
5	IPR014034	Ferritin, conserved site	2 of 3	8.49e-05
6	IPR012347	Ferritin like	2 of 3	8.49e-05
7	IPR009078	Ferritin like super family	2 of 7	8.49e-05
8	IPR009040	Ferritin like diiron domain	2 of 3	8.49e-05
9	IPR008331	Ferritin/DPS protein domain	2 of 3	8.49e-05
10	IPR001519	Ferritin	2 of 3	8.49e-05

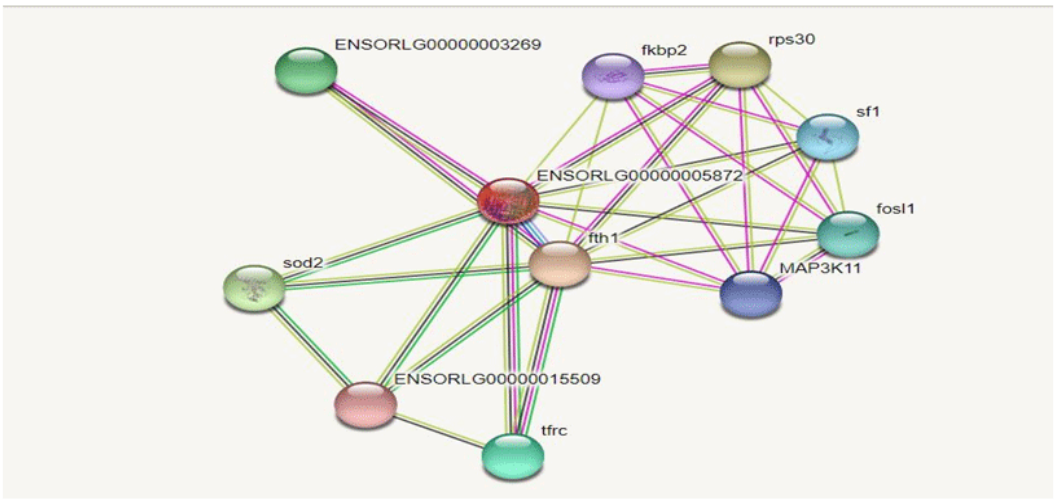

Figure 6: The interactive network view of predicted protein-protein interactions using STRING 10.0 tool. Network nodes are proteins, and the edges represent the predicted functional associations. The K means algorithm used to cluster the proteins in different groups. Inter-cluster edges represented by dashed lines. Small nodes: protein of unknown 3D structure; Large nodes: some 3 D structure is known or predicted. Extended lines: gene co-occurrence

function. The K- means algorithm used for the determination of protein clustering and their colors represent the input protein (Fig 6).

Protein model database

Finally the generated model for ferritin heavy chain subunit was successfully submitted in the Protein Model Database (PMDB) having the PMID: PM0081909.

Discussion

Ferritin is a conserved iron-binding protein involved in host defense and cellular iron metabolism and storage of toxic metal ions in most organisms. Iron metabolism including the regulation of its concentration and its detoxification [33]. As it had been reported in disk abalone [34], turbot [35], and red drum [36], purified recombinant ferritin is able to bind iron. This attribute explains ferritin's efficacy, via the iron with holding strategy, as part of the host innate immune response against microbial infections [37]. Ferritins are regarded as acute-phase proteins that respond to stress and inflammation. In lower vertebrates such as fish, ferritin synthesis is modulated by microbial infections and its expression increasing both in the liver and the brain, which is consistent with the decrease in serum iron and with the need to increase iron storage in order to make it unavailable for bacterial growth. [38, 39, 40]

In current study, the investigation of ferritin heavy chain protein in fish *Oryzias latipes* was accomplished with the use of bioinformatics tools and software. At first, primary structure analysis was done by computing following parameters of protein which are as sequence length, molecular weight, theoretical isolectric point (pI value), total number of negatively (Asp+Glu) and positively (Arg+Lys) charged residues, instability index, aliphatic index, and grand average of hydropathy (GRAVY) (Table 2). In current investigation, ferritin heavy chain subunits in *Oryzias latipes* were found acidic, unstable and hydrophilic. Moreover, the computed isoelectric point was 5.54. PI is a pH at which a protein carries no net charge. Computed values of instability index of FTH were 49.37. A protein whose instability index below 40 is predicted as stable, and above 40 is predicted as unstable. The aliphatic index of a protein defined as relative volume occupied by aliphatic side chains (valine, alanine, isoleucine, and leucine) [41]. It may be regarded as a positive factor for the increase of thermo stability of globular protein as it was in Table 2. As a result, high aliphatic index indicated structural stability. GRAVY value obtained below 0 in negative form that indicate the protein is hydrophilic in nature. The GRAVY value for a protein was calculated as the sum of hydropathy values of all the amino acids, divided by the number of residues in the sequence [42].

Before deducing the structure of the protein, it was necessary to compute the number of transmembrane regions of the protein. Both ProtScale and TMHMM predicted the protein to have transmembrane helices. These transmembrane domains are hydrophobic. TMHMM predicted one transmembrane helix in a sequence from 21 to 38 (Fig 2). Hydropathy plot of ferritin heavy chain subunit of *Oryzias latipes* showed two peaks above threshold value (0 to + value). The highest score was observed for the first peak which means it is the most hydrophobic and this region also contains the most number of amino acids than the other peaks because the base of the peak was wider than the other peaks (Fig 1). The hydrophobic force is simply that force, arising from the strong cohesion of the solvent, which drives molecules lacking any favorable interactions with the water molecules themselves from the aqueous phase [42]. In the case of the formation of the native structure from the random coil, this force participates in the reaction because hydrophobic side-chains, which are exposed to water in the extended coil, are removed to the interior of the protein during the folding of the native structure [43]. The native structure of a protein molecule will be that structure that permits the removal of the greatest amount of hydrophobic surface area and the smallest number of hydrophilic positions from exposure to water [43, 44].

In secondary structure of ferritin heavy chain subunit, alpha helix 56.50%, extended strands 10.73% and coiled region 32.77% were present (Table 2). Most of the residues have feeble but certain choices either for or contrary being in α-helix: Ala, Glu, Leu, and Met are good helix formers while Pro, Gly, Tyr, and Ser (Table 1) are very poor [45]. α-Helices were central to all the early attempts to predict secondary structure from amino acid sequence (e.g., [46, 47, 48, 49, 50] and they are still the characteristics that can be predicted with greatest accuracy [51, 52]. It has been reported that α-helices are more stable, robust to mutations and designable than β-strands in natural proteins [53] and also in artificial designed proteins [54].

The 3 D structure is the ultimate goal of protein structure prediction and it is essential to understand protein function. It is shown in Fig 3.1 and 3.2by Swiss Model, Phyre 2 server, and Raptor X. Homology modeling predicted the 3 Dimensional structure of the Ferritin heavy chain subunit of *Oryzias latipes*. The conformational analysis of protein structure was done by Swiss model server aligning the query sequence to the template sequence. The selected template contains a structure of mouse h-chain modified ferritin by X-ray diffraction technique (2.24Å) (Table 3). The score QMEAN estimated the model quality, and their full form is qualitative model energy analysis shown in (Table 4). This composite scoring function depicting the major geometrical aspects of protein structures. It was checked on several standard decoy sets including a molecular dynamics simulation decoy set as well as on a comprehensive dataset [55]. It shows a statistically significant improvement over all scales of quality and describing the ability of the scoring function to determine the native structure and recognize good and bad models [56]. The general understanding of ferritin structure is based on the human ferritin subunit [57], frog ferritin [58] and the E. coli ferritin [59]. The researcher study on the structure and functions of ferritin could not notice through, research review. With superposition and comparison, the ferritin structures of the human, frog, and E. coli was found in fishes.

In Phylogenetic analysis, Ferritin heavy chain subunits of 2 species belonging to same genera (*Oryzias latipes* and *Oryzias melastigma*) were analyzed with *Cyprinus carpio* (Out crossed) (Table 5, 6) for the same gene. According to the phylogeny tree, FTH were derived from an ancestor and evolved into different groups (Fig 5).

Genes which has involvement in related biological pathways are usually expressed cooperatively for their functions and their information on interaction is the key to understand biological systems at the molecular level. To further explore which genes are possibly regulated by FTH protein or pathway, a protein-protein interaction network was assembled (Table 7, 8). The protein-protein interaction network is an important component for the understanding of the cellular process at system-level. This network can be used to evaluate by filtering functional genomics data. It provides an instinctive platform for evolutionary, annotating, structural, and functional properties of a protein [60]. According to Table functional partners were identified. Partner FTH1 with the highest score (0.802) had functions including "Stores iron in a soluble", "non-toxic", "iron homeostasis" while partners ENSORL00000015509 with lowest score (0.530) had function including "proton-coupled divalent metal ion transporters", rest of the other obtained partners in this analysis had score between 0.604and 0.530, and functions including "40S ribosomal protein S30", "Superoxide dismutase", "Nuclear receptor co activator 4", "Transferrin receptor 1a", "FOS-like antigen 1a", "Splicing factor 1", "Mitogen-activated protein kinase kinasekinase 11" and "Peptidylprolyl isomerase; FK506 binding protein 2". Additionally,

two KEGG pathways viz "Ferroptosis", "Necroptosis",, six domains viz IPR014034, IPR012347, IPR009078, IPR009040, IPR008331, IPR001519 and one uniprot keywords including KW-0409 were identified in the network analysis.

Bioinformatics can play an important role in the analysis and interpretation of genomic and proteomic data with the use of method and technologies from mathematics, statistics, computer science, physics, biology and medicine [61]. It can be a powerful tool to predict the function of a protein from its amino acid sequence and has revolutionized the studies of organisms metabolism [62]. In this research, bioinformatics analyses of ferritin heavy chain protein in model fish *Oryzias latipes* shows the homology with mouse ferritin heavy chain. The obtained data from the analysis of protein provide a background for bioinformatics studies of the function and evolution of the proteins of freshwater fishes. In future these analytical approaches will be used to characterize the structural and functional role of protein in most of the freshwater fishes.

Acknowledgements

The author is thankful to the Dean, College of Fisheries, Lembucherra, and Tripura for encouragement and moral support. This work has been carried out under DBT funded Project "Establishment of Bioinformatics Infrastructure Facility for Biology Teaching through Bioinformatics" and hence financial support by the Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi.

Compliance with Ethical Standards

Conflict of Interest

The author declares that they have no conflict of interest.

References

1. Rakshit Ameta, Anil K Chohadia, Abhilasha Jain, Pinki B Punjabi. Fenton and Photo-Fenton. Advanced Oxidation Processes for Waste Water Treatment. Emerging Green Chemical Technology. 2018;49-87.
2. Arosio P and Levi S. Cytosolic and mitochondrial ferritins in the regulation of cellular iron homeostasis and oxidative damage. Biochim Biophys Acta. 2010;1800(8):783-792. Doi: 10.1016/j.bbagen.2010.02.005
3. Crichton RR, Charloteauxwauters M. Iron transport and storage. Eur J Biochem. 1987;164(3):485–506.
4. Aisen P, Listowsky I. Iron transport and storage proteins. Annu Rev Biochem. 1980;49:357–393.
5. Anderson GJ, Frazer DM. Hepatic iron metabolism. Semin Liver Dis. 2005;25(4):420–432.
6. Andrews SC. The ferritin-like superfamily: evolution of the biological iron storeman from a rubrerythrin-like ancestor. Biochim Biophys Acta. 2010;1800(8):691-705. Doi: 10.1016/j.bbagen.2010.05.010
7. Aisen P and Listowsky I. Iron transport and storage proteins. Annu Rev Biochem. 1980;49:357–393.
8. Harrison PM and Arosio P. The ferritins: molecular properties, iron storage function and cellular regulation. Biochim Biophys Acta. 1996;1275(3):161–203.
9. Torti FM and Torti SV. Regulation of ferritin genes and protein. Blood. 2002;99(10):3505–3516.
10. Guo J H, Juan S H and Aust S D. Mutational analysis of the four alpha-helix bundle iron loading channel of rat liver ferritin. Arch Biochem Biophys. 1998;352(1):71-77.
11. Arosio P and Levi S. Cytosolic and mitochondrial ferritins in the regulation of cellular iron homeostasis and oxidative damage. Biochim Biophys Acta. 2010;1800(8):783–792. Doi: 10.1016/j.bbagen.2010.02.005
12. Neves JV, Wilson JM and Rodrigues PNS. Transferrin and ferritin response to bacterial infection: the role of the liver and brain in fish. Dev Comp Immunol. 2009;33(7):848–857. Doi: 10.1016/j.dci.2009.02.001
13. Kocher TD, Thomas WK, Meyer A, Edwards SV, Paabo S, Villablanca FX, Wilson AC, et al. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc Natl Acad Sci USA. 1989;86(16):6196–6200.
14. Lynne R. Parenti. A phylogenetic analysis and taxonomic revision of ricefishes, Oryzias and relatives (Beloniformes, Adrianichthyidae). Zoological Journal of the Linnean Society. 2008;154(3):494–610.
15. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Ron D Appel, Amos Bairoch, et al. Protein identification and analysis on the ExPASy Server. The Proteomics Protocols Handbook. 2005;571-607.
16. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389-3402.
17. Sen TZ, Jerigen R, Kloczkowski A. GOR V server for protein secondary structure prediction. Bioinformatics. 2005;21(11):2787-2788.
18. Kyte J and Doolittle R F. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology. 1982;157(1):105–132.
19. Krogh A, Larsson B, von Heijne G, Sonnhammer E L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305(3):567–580.
20. Laskowski RA. PDBsum: summaries and analyses of PDB structures. Nucleic. 2001;29(1):221-222.
21. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling. Bioinformatics. 2006;22(2):195-201.
22. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and SwissPdbViewer: a historical perspective. Electrophoresis. 2009;1:S162-S173. Doi: 10.1002/elps.200900140
23. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, et al. SWISS-MODEL: modeling protein tertiary and quaternary structure using evolutionary information. Nucleic Acid Research. 2014;42:W252-W258. Doi: 10.1093/nar/gku340
24. Wass MN, Kelley LA and Sternberg MJ. 3DLigandSite: predicting ligand-binding sites using similar structures. Nucleic Acids Research. 2010;38:W469-473. Doi: 10.1093/nar/gkq406
25. DeLano WL, Ultsch MH, de Vos AM, Wells JA. Convergent solutions to binding at a protein-protein interface. Science. 2000;287(5456):1279–1283.

26. Sheng Wang, Wei Li, Shiwang Liu and Jinbo Xu. RaptorX-Property: a web server for protein structure property prediction. Nucleic Acids Res. 2016; 44(Web Server issue):W430–W435. Doi: 10.1093/nar/gkw306
27. Lovell c, Davis IW, Arendall WB, De Bakker PIW, Word JM, Prisant MG, Richardson JS, et al. Structure validation by Calpha geometry: phi, psi, and Cbeta deviation. Proteins. 2003;50(3):437-450.
28. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013;30(12):2725–2729. Doi: 10.1093/molbev/mst197
29. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43:D447–D452. Doi: 10.1093/nar/gku1003
30. Kyte J and Doolittle R F. A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology. 1982;157(1):105–132.
31. Andras Fiser. Template-Based Protein Structure Modeling. Methods Mol Biol. 2010;673:73–94. Doi: 10.1007/978-1-60761-842-3_6
32. Rachel Casiday and Regina Frey. Iron in Biology: Study of the Iron Content in Ferritin, The Iron-Storage Protein. Iron Use and Storage in the Body: Ferritin and Molecular Representations. Department of Chemistry, Washington University St. Louis, MO. 1999;63130.
33. Zhang L, Sun W, Cai W, Zhang Z, Gu Y, Chen H, Ma S, et al. Differential response of two ferritin subunit genes (VpFer1 and VpFer2) from Venerupisphilippinarum following pathogen and heavy metals challenge. Fish Shellfish Immunol. 2013;35(5):1658- 1662. Doi: 10.1016/j.fsi.2013.07.030
34. De Zoysa M and Lee J. Two ferritin subunits from disk abalone (Haliotis discus discus): cloning, characterization and expression analysis. Fish Shellfish Immunol. 2007;23(3):624-635.
35. Zhang X, Wei W, Wu H, Xu H. Gene cloning and characterization of ferritin H and M subunits from large yellow croaker (Pseudosciaenacrocea). Fish Shellfish Immunol. 2010;28(5-6):735-742. Doi: 10.1016/j.fsi.2009.11.027
36. Hu YH, Zheng WJ and Sun L. Identification and molecular analysis of a ferritin subunit from red drum (Sciaenopsocellatus). Fish Shellfish Immunol. 2010;28(4):678-686. Doi: 10.1016/j.fsi.2010.01.001
37. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976;72:248-254.
38. Torti FM and Torti SV. Regulation of ferritin genes and protein. Blood. 2002;99(10):3505–3516.
39. Orino K, Lehman L, Tsuji Y, Ayaki H, Torti SV, Torti FM. Ferritin and the response to oxidative stress. Biochem J. 2001;357:241–247.
40. Rogers JT, Bridges KR, Durmowicz GP, Glass J, Auron PE, Munro HN. Translational control during the acute phase response. Ferritin synthesis in response to interleukin-1. J Biol Chem. 1990;265(24):14572–14578.
41. Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem. 1980;88(6):1895–1898.
42. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1):105–132. Doi: 10.1016/0022-2836(82)90515-0
43. Levitt M, Chothia C. Structural patterns in globular proteins. Nature. 1976;261(5561):552-558. doi: 10.1038/261552a0
44. Bigelow CC. On the average hydrophobicity of proteins and the relation between it and protein structure. J. Theor. Biol. 1967;16(2):187-211.
45. Michael levitt, jonathangreer. Automatic identification of secondary structure in globular proteins. J. Mol. Bid. 1977;114(2):181-293. doi: 10.1016/0022-2836(77)90207-8
46. Lippert T, Rarey M. Fast automated placement of polar hydrogen atoms in protein-ligand complexes. Journal of Cheminformatics,. 2009. doi: 10.1186/1758-2946-1-13
47. Davies, DR. A correlation between amino acid composition and protein structure J. Mol. Biol. 1965;5(6):809-822. Doi: 10.1016/S0006-3495(65)86753-4
48. Guzzo AV. The influence of amino acid sequence on protein structure. Biophys. J. 1965;5(6):809–822. doi: 10.1016/S0006-3495(65)86753-4
49. Prothero JW. Correlation between the distribution of amino acids and alpha helices. Biophys. J. 1966;6(3);367–370. doi: 10.1016/S0006-3495(66)86662-6
50. Cook DA. The relation between amino acid sequence and protein conformation.J. Mol. Biol. 1967;29(1):167-171. doi: 10.1016/0022-2836(67)90188-x
51. Chou PY, Fasman G. Conformational parameters of amino acids in helical, β-sheet, and random coil regions calculated from proteins. Biochemistry. 1974;13(2):211-222. doi: 10.1021/bi00699a001
52. Lim, V.I. Algorithm for prediction of alpha helical and Beta structural regions of globular proteins. J. Mol. Biol. 1974;88(4):873-894. doi: 10.1016/0022-2836(74)90405-7
53. Abrusan, G, Marsh JA. Alpha helices are more robust to mutations than beta strands. PLoS Computational Biology. 2016;12(12):1–16. doi: 10.1371/journal.pcbi.1005242
54. Rocklin GJ, Tamuka M. Chidyausiku,Inna Goreshnik, Alex Ford, Scott Houliston et al. Global analysis of protein folding using massively parallel design, synthesis, and testing. Science. 2017;357 (6347):168–175. doi: 10.1126/science.aan0693.
55. Acharjee Samik, Upadhyay Datta Anil, Roy Kumar Ajit. Insilico analysis of Myostatin protein of Labeocalbasu. MedCrave. 2017;3(5):422–428. doi: 10.15406/jabb.2017.03.00080
56. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics. 2011;27(3):343-350. doi: 10.1093/bioinformatics/btq662
57. Lawson DM, Artymiuk PJ, Yewdall SJ, Smith JM, Livingstone JC, Treffry A, Luzzago A et al. Solving the structure of human H ferritin by genetically engineering intermolecular crystal contacts[J]. Nature. 1991;349(6309):541-544.
58. Ha Y, Shi D, Small W, Theil EC, Allewell NM. Crystal structure of bullfrog M ferritin at 2.8 A resolution: analysis of subunit interactions and the binuclear metal center[J]. Journal of Biological Inorganic Chemistry. 1999;4(3),:243-256.
59. Macedo S, Romao CV, Mitchell E, Matias PM, Liu MY, Xavier AV et al. The nature of the di-iron site in the bacterioferritin from Desulfovibriodesulfuricans[J]. Natural Structural Biology. 2003;10(4):285-290. doi: 10.1038/nsb909
60. Samin Seddigh, Darabi Maryam. Functional, structural, and

phylogenetic analysis of mitochondrial cytochrome b (cytb) in insects, Mitochondrial DNA Part A. DNA Mapping, Sequencing, and Analysis. 2017;29(2):236-249. doi: 10.1080/24701394.2016.1275596

61. Romano P, Giugno R, Pulvirenti A. Tools and collaborative environments for the bioinformatics research. Briefings in Bioinformatics. 2011;12(6);549-561. doi: 10.1093/bib/bbr055

62. Seddigh S, Darabi M. Structural and phylogenetic analysis of α-glucosidase protein in insects. Biologia. 2015;70(6):812–825. doi: 10.1515/biolog-2015-0096