Computational characterisation, homology modelling and structure-based functional annotation of very low-density lipoprotein receptor of *Salmo trutta fario*

ABSTRACT

The present study was designed to enlist the physiochemical and functional properties of the very low-density lipoprotein receptor of *Salmo trutta*, and provide information about its three-dimensional structure and interaction with partner molecules using standard bioinformatic tools. VLDL receptor is a large flexible protein with a molecular weight of around 95.9 kDa, relatively unstable, and hydrophobic with a primary transmembrane helix from N-786 to C-808. All of the 70 Cysteine residues (except four) are in a disulphide bonding state. Secondary structure analysis shows that most of the protein has a predominant random coiled configuration. Six Epidermal growth factor-like domains and two EGF-like-Ca^{2+} binding domains were predicted. The protein plays a crucially important role in various metabolic pathways including vitellogenesis in fishes. Understanding the structural and functional properties of the VLDL receptor will facilitate a better understanding of its molecular dynamics and the designing of experimental procedures.

Keywords: Expasy, SOSUI, Fisheries, Swiss-Model, Vitellogenesis

Comment [t1]: Why did you pick interest in this fish? Could you provide its common name alongside the scientific name please?

1. INTRODUCTION

Fish eggs contain a huge yolk mass as a source of nutrients for developing embryos and the growth of larvae. The major precursors of yolk mass accumulation are the liver-originated diverse vitellogenins (VtgAa, VtgAb, VtgC etc) in a hormonally controlled and regulated process. During this process, estrogen triggers the synthesis of vitellogenins in the liver, which are then transported to oocyts via the blood, internalized and stored in the ooplasm [1-4].

In fishes, while many studies have been conducted on different aspects of the vitellogenesis process, the internalization and subsequent uptake of the very low-density lipoproteins (VLDLs) including Vtgs have not been described in detail. Earlier reports [5,6] revealed that salmonids possess VtgC along with A-type Vtg genes. Additionally, specific Vtg receptors (Vtgrs) are involved in the selective ovarian uptake of different Vtgs types via Vtgr-mediated endocytosis in fishes. These receptors have been identified and classified as low-density lipoprotein receptor (LDLR) family. Over the past two decades, several studies have identified an ever-increasing number of low-density lipoprotein receptor (LDLR) family members. Although these receptors have several common structural features, their functional roles appear to be quite diverse attributed to the wide range of ligands [7,8]. The common structural features in this family include 40-residue binding repeats, epidermal growth factor precursor-type repeats, consensus tetrapeptide, and signals in the cytoplasmic region for receptor internalization. LR8 or VLDLR, one of the members containing a single ligand-binding domain with eight ligand-binding repeats [9], was previously identified and designated as the Vtgr in teleosts [10]. Davail *et al*., [11] characterized an 826-residue type-I membrane receptor, belonging to the low-density lipoprotein receptor (LDLR) superfamily. The receptor had a striking similarity to the mammalian so-called very low-density

lipoprotein receptors as both have amino-terminal ligand-binding domains consisting of a cluster of 8 cysteine-rich repeats.

The major drawbacks of experimental methods for the characterization of any protein are the involvement of high cost and time consumption. Nowadays several *in-silico* approaches and online servers provide great prospects for the investigation of proteins to hasten experimental approaches. Computational tools deliver researchers a cost-effective way to understand physicochemical and the structural properties of a protein along with their functional description. The amino acid sequence provides most of the information required for determining and characterizing the molecule's function and physical and chemical properties. Homology modelling is of great interest in designing and evaluating biological experiments involving proteins when no experimental three-dimensional (3D) structures are available. Computational prediction of 3D structures can give us valuable insights regarding the molecular mechanism of the involved protein. Therefore, there is an obvious knowledge gap that demands computational methods for protein structure prediction [12]. With this background, the present study was undertaken to describe the structural features and to understand the molecular function of ovarian the VLDL receptor in brown trout.

Comment [t2]: This phrase seems incorrect. Please can you correct it

2. MATERIAL AND METHODS

2.1 Sequence retrieval and physicochemical properties

The sequence of brown trout VLDL receptor protein was retrieved in FASTA format (Accession no. A0A673XME1) from UniProt [13], a comprehensive public domain resource for protein sequence and annotation data, and used for further analysis. The functional analysis of the VLDL receptor sequence and classification was done by the InterPro resource server [14] which identifies protein families, domains, and functional sites. The physicochemical properties were computed from the primary structure of the VLDL receptor where parameters including theoretical isoelectric point (pI), molecular weight (MW), the

Comment [t3]: What was the reason for retrieving and using VLDL receptor and not any other receptor in brown trout?

Comment [t4]: Replace with were used for further analysis

total number of positive and negative residues, extinction coefficient (ε), instability index, aliphatic index and grand average hydropathy (GRAVY) were computed using the Expasy's ProtParam prediction server [15].

2.2 Structural analysis

SOPMA (Self-Optimized Prediction Method with Alignment) was employed for estimating the secondary structural (2^0) features of trout VLDL receptor protein [16]. SOSUI server [17] was used for the identification of transmembrane regions. Computational methods including CYS_REC (http://linux1.softberry.com/berry.phtml?topic) and CYSPRED [18] servers were also used for determining disulphide bonds (SS), which are essential in determining the functional linkage and the stability of a particular protein. Potential phosphorylation and glycosylation sites of the protein were studied using NetPhos3.1a [19] and NetNGlyc servers (http://www.cbs.dtu.dk/services/NetNGlyc/) respectively.

The three-dimensional protein structures provide valuable insights into the molecular basis of protein function, allowing an effective design of experiments, such as site-directed mutagenesis, studies of disease-related mutations, or the structure-based design of specific inhibitors. The modelling of the 3D structure of the trout VLDL receptor was performed by the homology modelling program Swiss-Model [20]. Swiss-Model server is for automated comparative modelling of 3D protein structures allowing users to submit a sequence and get back a structure automatically. After the selection of a template using BLAST and HHbits, a raw model is built based on a rigid fragment assembly approach. The quality of the predicted model and the accuracy of the generated structure were evaluated using Ramachandran plot analysis using PROCHECK [21] and QMEAN [22] methods. The local quality of the model was evaluated based on QMEANDisco [23].

Homology modelling was also done by using a template structure from PDB [24] through BLASTP search [25]. The modelled 3D structure was evaluated using the online

servers ProQ [26] and ProSA [27]. ProQ is a neural network-based predictor that predicts the quality of a protein model in the form of LGscore and MaxSub based on several structural features. On the other hand, ProSA calculates an overall quality score (z-score) for a specific input structure. If the calculated score is outside the characteristic range of native proteins, the structure probably contains errors.

2.3 Protein-protein interactions

To generate an interaction network of the trout VLDL receptor, the STRINGv11.5 server (Search Tool for the Retrieval of Interacting Genes/Proteins; [28] was used. STRING is a database 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 context, high-throughput experiments, conserved expression, and previous knowledge. Protein scores greater than 0.6 were included in the results.

3. RESULTS AND DISCUSSION

The primary structure analysis was performed using Expasy's ProtParam tool and different physicochemical parameters were computed. The molecular weight of the trout VLDL receptor is 95.9 kDa with 863 amino acid residues. The computed isoelectric point (pI) of 4.83 indicates the highly acidic nature of the trout VLDL receptor. At the isoelectric point, a molecule carries no charges or the negative and positive charges are equal and coincide with the pH where the protein is least soluble. Thus, at this point, the protein is most likely to precipitate out and the mobility in an electro-focusing system is zero. The total number of negatively (Asp+Glu) and positively (Arg+Lys) charged residues were 125 and 84 respectively. The estimated extinction coefficient (EC) was 141195 M^{-1} cm^{-1} at 280 nm measured in water, assuming all pairs of Cys residues form cystines. EC is a measure of the strength of absorbs light absorption at a specific wavelength (generally 280 nm) and is an intrinsic property of a protein. Hence, to precisely determine protein concentration, it is

Comment [t5]: I suggest this should be deleted

fundamental to accurately determine the extinction coefficient. The EC value of the trout VLDL receptor is highly attributed to a large number of Cys, Trp, and Tyr residues. Additionally, the Instability index (II) provides an estimation of the stability of the protein in vitro where a protein whose instability index is smaller than 40 is predicted as stable and a value >40 predicts that the protein may be unstable [29]. The computed (II) of the trout VLDL receptor was 40.64 classifying the protein as unstable. The aliphatic index (AI), defined as the relative volume of a protein occupied by aliphatic side chains (A, V, I, and L) was 64.16. It is regarded as a positive factor for the increase of thermal stability of globular proteins [30]. The higher the aliphatic index, the higher will be the stability of the protein. The estimated half-life was 30 hours (mammalian reticulocytes, in vitro), >20 hours (yeast, in vivo) and >10 hours (*Escherichia coli*, in vivo). GRAVY represents the hydrophobicity value of the peptide or protein by calculating the sum of the hydropathy values of all the amino acids divided by the number of residues in the sequence. In general, the lesser the value is, the more superior the interaction takes place between a protein and water. The estimated GRAVY of the trout's VLDL receptor was -0.503, indicating hydrophilicity [31].

The trout's VLDL receptor has 863 amino acids. Aspartic Acid is the most abundant amino acid present, followed by cysteine and glycine. The presence of high aspartic acid residues in proteins is vital as it interacts with the solvent, stabilizing the protein's 3D structure in addition to being binding sites for metal ions and charged ligands [32]. The presence of 70 (8.1%) Cys residues in the protein indicate the presence of disulphide bonds in the final structure. The overall amino acid composition of trout VLDL receptor included Alanine (5.0%), Arginine (5.1%), Asparagine (6.4%), Aspartic Acid (8.3%), Cysteine (8.1%), Glutamine (4.3%), Glutamic Acid (6.1%), Glycine (7.9%), Histidine (1.9), Isoleucine (5.3%), Leucine (5.8%), Lysine (4.6%), Methionine (1.3%), Phenylalanine (3.1%), Proline (4.6%), Serine (7.0%), Threonine (5.3%), Tryptophan (2.3%), Tyrosine (2.1%) and Valine (5.4%).

The InterPro server classified the trout VLDL receptor into the low-density lipoprotein receptor (LDL receptor) superfamily. Earlier studies in other organisms/ species have also reported that the VLDL receptor is a member of (the LDL) family [33]. InterPro analysis suggested the presence of six Epidermal growth factor-like (EGF-like) domains and two EGF-like-Ca^{2+} binding domains in the trout VLDL receptor. Both domains are characterized by a sequence of about 30 to 40 amino acid residues with three internal disulfide bridges. Structurally both are similar and consist of a two-stranded β-sheet followed by a loop to a C-terminal short two-stranded sheet [34].

In addition, the functional characterization of trout VLDL receptor including transmembrane region identification, and prediction of disulphide bonding pairs was also performed. The SOSUI functional analysis server used for the identification of transmembrane helices along with their corresponding lengths classified the receptor as a membrane protein with one primary transmembrane helix from N-786 to C-808 (Figure 1a and 1b). The $SOSUI_{signal}$ server indicated the presence of a ten amino acid long (MAFVFCALQT) signal sequence at the N-terminal of the sequence. Additionally, $SOSUI_{GramN}$ showed the extracellular localization of the trout VLDL receptor.

The secondary structure was predicted using SOPMA for the states of alpha-helix, extended strand, beta-turn, and random coil [16] using default parameters (Windows width: 17, similarity threshold: 8, and the number of states: 4). The results revealed that the secondary structure of trout VLDL receptor is dominated by random coils (63.73%) followed by extended strand (21.67%), alpha helix (9.62%) and beta-turn (4.98%). All other secondary structure features such as 3_{10} helix, Pi helix, Ambiguous states, Bend region and Beta bridge were not found.

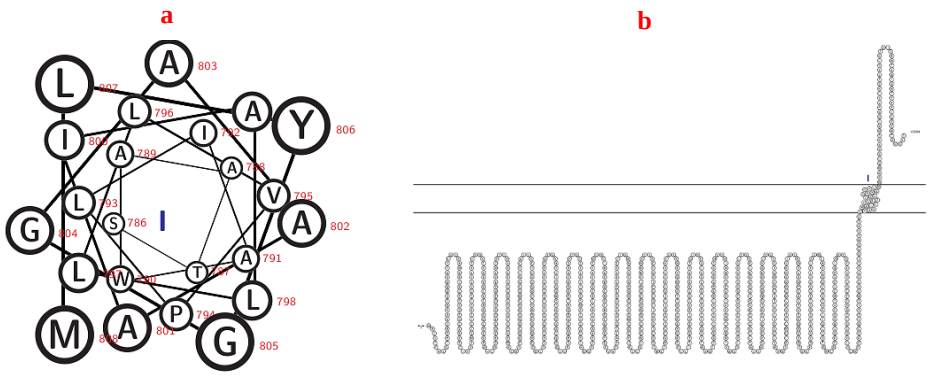

Figure 1. Representation of transmembrane helix of brown trout VLDL receptor. (a) Wheel plot (b) Snake-like plot.

The tool CYS_REC identifies the presence of SS bonds and possible bonding pairs among all Cys residues. CYS_REC locates the “SS bond” between the pair of cystein residues and if present, yields the position of cysteins, their total number, and the pattern of pairs in the protein sequence. On the other hand, CYSPRED is a neural network-based predictor that distinguishes the bonding states of cysteine in proteins starting from the residue chain. CYS_REC identified most of the cysteine residues in SS bond conformations with 42 residues having a high probability (>30). On the other hand, CYSPRED predicted only four Cys residues (Cys6, Cys360, Cys441 and Cys631) to be in a non-bonding SS state. Earlier, all four were found to be non-SS bound by CYS-REC also. NetPhos3.1a predicted 77 potential phosphorylation sites (serine, threonine, and tyrosine) suggesting the high activity of the VLDL receptor protein (Figure 2). NetPhos3.1a uses ensembles of neural networks to predict phosphorylation sites in eukaryotic proteins. On the other hand, the NetNglyc-1.0 server predicts N-Glycosylation sites in proteins using artificial neural networks that examine the sequence context of Asn-Xaa-Ser/Thr patches. NetNGlyc-1.0 tool analysis of trout VLDL

receptor predicted five potential glycosylation sites at residues 156, 207, 277, 319 and 549 (Figure 3).

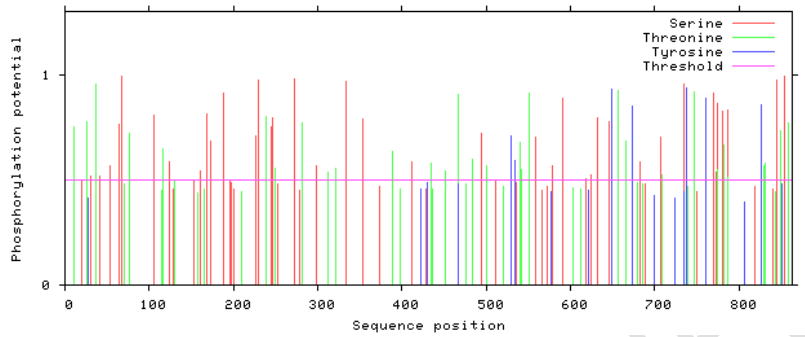

Figure 2. NetPhos3.1a predicted phosphorylation sites in the brown trout VLDL receptor.

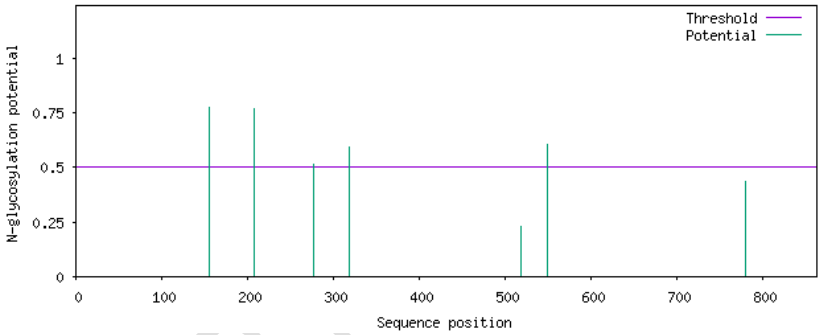

Figure 3. NetNGlyc-1.0 predicted glycosylation sites in the brown trout VLDL receptor.

3D protein structures provide valuable insights into the molecular basis of protein function, allowing an effective design of experiments. These models are of great interest for planning and analyzing biological experiments when no experimental three-dimensional structures are available. Different web-based homology modelling servers can predict the 3D structure of proteins at 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 similar sequences to practically identical structures and distantly related sequences still fold into similar structures [35]. Thus,

the homologous proteins have regions that retain the same general fold and regions where the folds differ. In the present study, the modelling of the 3D structure of trout VLDL receptor was performed by the homology modelling program Swiss-Model. The final modelled structures were visualized by Swiss PDB Viewer as shown in figure 4.

The quality of the trout VLDL receptor predicted model and the accuracy of the structure generated were evaluated using Ramachandran plot analysis which concluded that 68.54% of amino acids were favoured and 9.19% were Ramachandran outliers (Figure 5) suggesting that the model is moderately acceptable. The global quality of the model was evaluated by QMEAN (Qualitative Model Energy Analysis) score given by the Swiss-Model server. The QMEAN value of the structure was 0.57, validating the moderate quality of the overall structure (Figure 6a). The local quality value of individual amino acids also showed that most of the regions of the 3D structure were modelled with moderate quality (Figure 6b).

Comment [t6]: What is the standard/range for acceptability of the model please. Kindly state the standard for the purpose of comparism

Comment [t7]: What is the global and local standard for quality of acceptability please. This will help in justifying this result

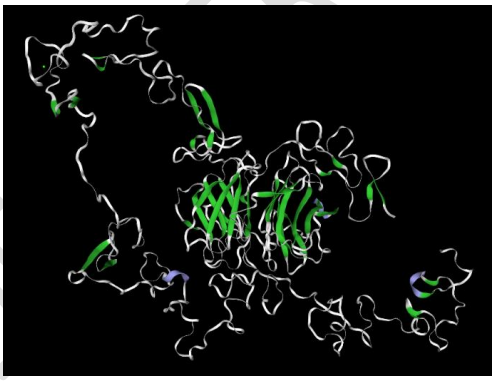

Figure 4. 3D structure of brown trout VLDL receptor. Secondary structure features are highlighted in different colours.

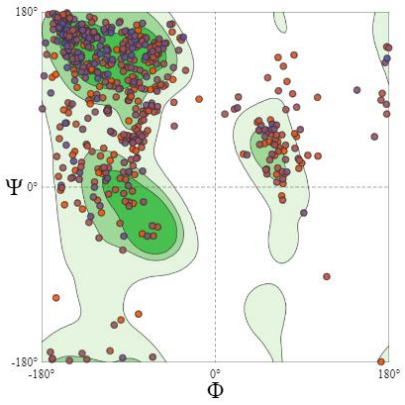

Figure 5. Ramachandran plot analysis of the brown trout VLDL receptor 3D model structure

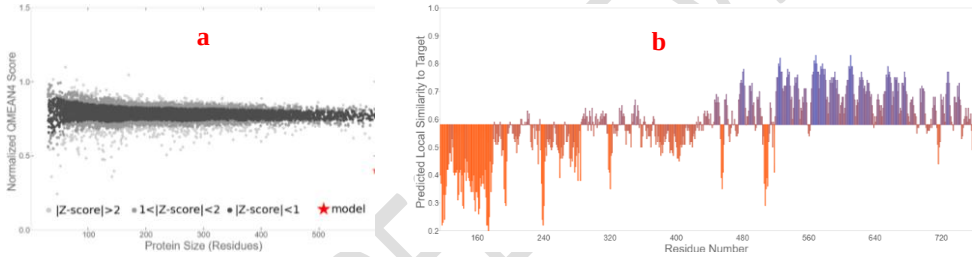

Figure 6. Quality estimate of brown trout VLDL receptor 3D model structure. (a) Global quality estimate (b) Local quality estimate

The modelled structure was also validated by the verification server Protein Quality (ProQ), which validates protein models based on different validation parameters. The predicted LG score of the model was 9.237 suggesting an extremely good model (LG score >4). ProSA was used to check the three-dimensional model of trout VLDL receptor for potential errors. The program displays two quality measures of the input structure; z-score and a plot of its residue energies. The z-score indicates overall model quality and measures the deviation of the total energy of the structure from an energy distribution derived from random conformations. As shown in Figure 7a, the Z-score for the VLDL receptor was well

within the range of scores typically found for proteins of similar large size, indicating a reliable structure. The energy plot shows the local model quality by plotting energies as a function of amino acid sequence position. In general, positive values correspond to problematic or erroneous parts of a model. Figure 7b displays a comparable energy plot for both the target and template structures. A similar type of in silico homology modelling has also been reported for different proteins in different species [36-39].

Comment [t8]: insert "this has been reported in silico..."

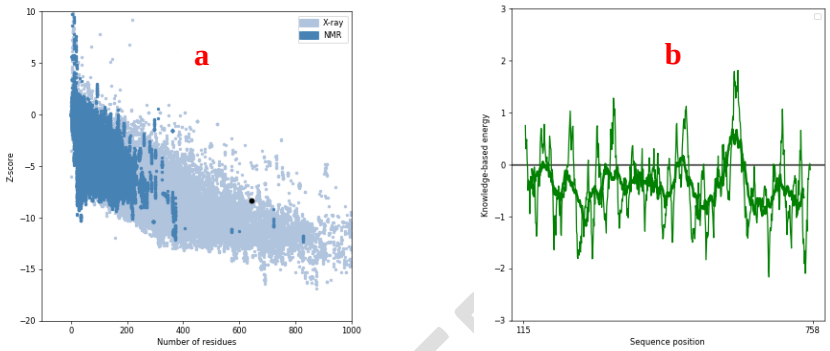

Figure 7. ProSA analysis of *Salmo trutta* VLDL receptor model. **(a)** z-scores of all protein chains in PDB determined by X-ray crystallography (light blue) or NMR spectroscopy (dark blue) with respect to their length. The z-score of the *Salmo trutta* VLDL receptor model is highlighted as a large black dot. **(b)** Energy plot (local model quality) of *Salmo trutta* VLDL receptor

All proteins function by binding to other proteins or molecules, and many methods exist for studying protein-protein interactions and identifying potential protein partners. 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. For functional protein association networks of brown trout VLDL receptor, STRINGv11.5 was used for the prediction of interaction with other partners. Only the partner proteins with a score of more than 0.4 were included in the results.

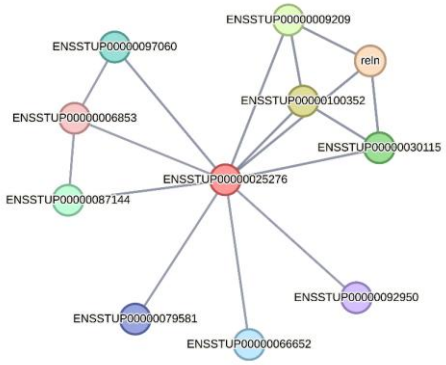

Figure 8. Protein-protein interaction map of trout VLDL receptor (ENSSTUP00000025276) by STRING web server. The closest interacting protein has the shortest node.

The network analysis revealed very low-density lipoprotein particle receptor activity, response to insecticide, clathrin-dependent synaptic vesicle endocytosis and phosphatidylcholine-sterol o-acyltransferase activity. Earlier Mitchel *et al*., [40] explored the possibilities of targeting Tick VLDL receptors specifically via novel vaccines or insecticides as a control measure. Moreover, clathrin-dependent synaptic vesicle endocytosis has been well explained in insects, unlike fishes [41, 42]. Potential interacting proteins associated with the trout VLDL receptor are listed in (Figure 8). Notably, many of these partners have not been characterized yet in brown trout depicting the scope for further research in the area. Elucidating these types of complex protein interactions will give important clues as to the function of novel proteins that govern cell behaviour.

Comment [t9]: The discussion is too scanty. More relevant references have to be cited with a detailed discussion please.

4. CONCLUSION

In this study, the very low-density lipoprotein receptor that plays an important role in the lipid metabolism of brown trout was selected. Physicochemical characterization was performed by computing theoretical isoelectric point (pI), molecular weight, the total number

of positive and negative residues, extinction coefficient, instability index, aliphatic index, and grand average hydropathy (GRAVY). Functional analysis of this receptor was performed using the SOSUI server which predicted the transmembrane helix. Secondary structure analysis revealed the dominancy of random coils followed by extended strand, alpha helix and beta-turn. The modelling of the 3D structure was performed by the Swiss-Model and validated using protein structure checking tools ProQ and ProSA. Homology-derived models can be used in a wide range of applications and serve as a keystone for functional analysis of experimentally derived crystal structures.

ETHICAL APPROVAL

This article does not contain any studies with human participants or animals.

REFERENCES

1. Wallace RA. Vitellogenesis and oocyte growth in nonmammalian vertebrates. Oogenesis. 1985;127-177.
2. Selman K, Wallace RA. Cellular aspects of oocyte growth in teleosts. Zoological science. 1989;6(2):211-231.
3. Hiramatsu N, Matsubara T, Weber GM, Sullivan CV, Hara A. Vitellogenesis in aquatic animals. Fisheries science. 2002;68(sup1):694-699.
4. Patiño R, Sullivan CV. Ovarian follicle growth, maturation, and ovulation in teleost fish. Fish Physiology and Biochemistry. 2002;26(1):57-70.
5. Buisine N, Trichet V, Wolff J. Complex evolution of vitellogenin genes in salmonid fishes. Molecular Genetics and Genomics. 2002;268(4):535-542.
6. Mushirobira, Y., Mizuta, H., Luo, W., Morita, Y., Sawaguchi, S., Matsubara, T., Hiramatsu N, Todo T, Hara A. Changes in levels of dual vitellogenin transcripts and proteins in cutthroat trout Oncorhynchus clarki during ovarian development. Nippon Suisan Gakkaishi. 2013;79(2):175-189.

7. Bujo H, Hermann M, Kaderli MO, Jacobsen L, Sugawara S, Nimpf J, Yamamoto T, Schneider WJ. Chicken oocyte growth is mediated by an eight ligand binding repeat member of the LDL receptor family. The EMBO journal. 1994;13(21):5165-5175.
8. Elkin RG, MacLachlan I, Hermann M, Schneider WJ. Characterization of the Japanese quail oocyte receptor for very low density lipoprotein and vitellogenin. The Journal of nutrition. 1995;125(5):1258-1266.
9. Bujo H, Lindstedt KA, Hermann M, Dalmau LM, Nimpf J, Schneider WJ. Chicken oocytes and somatic cells express different splice variants of a multifunctional receptor. Journal of Biological Chemistry. 1995;270(40):23546-23551.
10. Hiramatsu N, Chapman RW, Lindzey JK, Haynes MR, Sullivan CV. Molecular characterization and expression of vitellogenin receptor from white perch (Morone americana). Biology of reproduction. 2004;70(6):1720-1730.
11. Davail B, Pakdel F, Bujo H, Perazzolo LM, Waclawek M, Schneider WJ, Le Menn F. Evolution of oogenesis: the receptor for vitellogenin from the rainbow trout. Journal of lipid research, 1998;39(10):1929-1937.
12. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic acids research. 2003;31(13):3381-3385.
13. Wang Y, Wang Q, Huang H, Huang W, Chen Y, McGarvey PB *et al*. A crowdsourcing open platform for literature curation in UniProt. PLoS Biology. 2021;19(12):e3001464.
14. Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, *et al*. The InterPro protein families and domains database: 20 years on. Nucleic acids research.2021;49(D1):D344-D354.
15. Gasteiger E, Hoogland C, Gattiker A, Wilkins,MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook. 2005;571-607.

16. Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics. 1995;11(6):681-684.
17. Hirokawa T, Boon-Chieng S, Mitaku S. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics (Oxford, England), 1998;14(4):378-379.
18. Fariselli P, Riccobelli P, Casadio R. Role of evolutionary information in predicting the disulfide-bonding state of cysteine in proteins. Proteins: Structure, Function, and Bioinformatics. 1999;36(3):340-346.
19. Blom N, Gammeltoft S, Brunak S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. Journal of molecular biology. 1999;294(5):1351-1362.
20. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, *et al*. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research. 2018;46(W1):W296-W303.
21. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. Journal of applied crystallography. 1993;26(2):283-291.
22. Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics, 2011;27(3):343-350.
23. Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T. QMEANDisCo-distance constraints applied on model quality estimation. Bioinformatics. 2020;36(6):1765-1771.
24. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, *et al*. The protein data bank. Nucleic acids research.2000;28(1):235-242.

25. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of molecular biology. 1990;215(3):403-410.

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

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

28. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, *et al*. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic acids research. 2021;49(D1):D605-D612.

29. Guruprasad K, Reddy BB, Pandit MW. 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. 1990;4(2):155-161.

30. Ikai A. Thermostability & aliphatic index of globular proteins. The Journal of Biochemistry. 1980;88(6):1895-1898.

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

32. Maniccia AW, Yang W, Johnson JA, Li S, Tjong H, Zhou HX, Shaket LA, Yang JJ. Inverse tuning of metal binding affinity and protein stability by altering charged coordination residues in designed calcium binding proteins. PMC biophysics. 2009;2(1):1-16.

33. Takahashi S, Sakai J, Fujino T, Hattori H, Zenimaru Y, Suzuki J, Miyamori I, Yamamoto TT. The very low-density lipoprotein (VLDL) receptor: characterization

and functions as a peripheral lipoprotein receptor. Journal of atherosclerosis and thrombosis. 2004;11(4):200-208.

34. Davis CG. The many faces of epidermal growth factor repeats. The New biologist. 1990;2(5):410-419.
35. Chothia C, Lesk AM. The relation between the divergence of sequence and structure in proteins. The EMBO journal. 1986;5(4):823-826.
36. Chang KY, Yang JR. Analysis and prediction of highly effective antiviral peptides based on random forests. PloS one. 2013;8(8):e70166.
37. Hossain MM. Fish antifreeze proteins: Computational analysis and physicochemical characterization. International Current Pharmaceutical Journal. 2012;1(2):18-26.
38. Enany S. Structural and functional analysis of hypothetical and conserved proteins of Clostridium tetani. Journal of infection and public health. 2014;7(4):296-307.
39. Saleem A, Rajput S. Insights from the in silico structural, functional and phylogenetic characterization of canine lysyl oxidase protein. Journal of Genetic Engineering and Biotechnology. 2020;18(1):1-11.
40. Mitchell III, R.D., Sonenshine, D.E. and Pérez de León, A.A., 2019. Vitellogenin receptor as a target for tick control: a mini-review. Frontiers in physiology, 10, p.618.
41. Sappington, T.W. and Raikhel, A.S., 1998. Molecular characteristics of insect vitellogenins and vitellogenin receptors. Insect biochemistry and molecular biology, 28(5-6), pp.277-300.
42. Van Antwerpen, R., Daphne, Q.D.P. and Ziegler, R., 2005. Accumulation of lipids in insect oocytes. In Reproductive Biology of Invertebrates, Vol. 12, Part B (pp. 281-304). CRC Press.