

## Phylogenetic assessment and *in silico* characterization of cytochrome b protein of three alpheid shrimps

Muthusamy THANGARAJ

*Annamalai University, Faculty of Marine Sciences, Centre of Advanced study in Marine Biology, Parangipettai, Tamil Nadu 608 502, India; coralholder@yahoo.com*

### Abstract

Cytochrome b (cyt b) is one of the cytochrome proteins involved in electron transport in the respiratory chain of mitochondria. Cyt b is the only gene among the cytochrome complex coded by mitochondrial DNA. It is the most widely used gene for phylogenetic assessment and inter species variation studies. Here, the amino acid sequence of cyt b in three snapping shrimps such as, *Alpheus lobidens*, *A. randali*, *A. bellulus* was analysed and the results showed higher similarity in *A. lobidens* and *A. randali* as reflected in the phylogenetic tree. This study describes the applications of bioinformatics tools to predict the physico-chemical characters of cyt b protein. This protein was composed of least percentage of Cys (0.8%) and highest percentage of Leu (13.8%). The maximum molecular weight (MW) was predicted as 42.62 KDa in *A. randali*. The theoretical pI value was ranged from 8.35 to 8.36 and confirmed that cyt b was alkali in nature. The instability index value was in the range of 42.29 to 46.94 which showed the protein was more stable. The secondary structure of this protein was primarily composed of  $\alpha$ -helixes and random coil, revealing the stable structure. The comparative modelling was performed by Swiss model where the 3-D crystal structure of bovine cyt bc1 (6haw1.c) was used as template. Ramachandran plot analysis showed that most of the amino acids (>92%) falling on the favoured region. Seven conserved motifs were identified by MEME analysis. The modelled 3-D structure of this protein was validated by PROCHECK and QMEAN. The transmembrane protein topology and helix probability curve was predicted by TMHMM server. Protein-protein interactions was analysed by STRING tool and found the network of cyt b with related proteins. The results of this study may provide valuable insights into fundamental characteristics of cyt b in Alpheid shrimps.

**Keywords:** cyt b; phylogeny; physicochemical characters; Ramachandran plot; snapping shrimp

### Introduction

Cytochrome b is one of the proteins occurs in eukaryotic mitochondria. Its main role is electron transport with the help of heme group and is the major subunit of transmembrane cytochrome bc1 and b6f complexes (Esposti *et al.*, 1993). Cyt b is an integral membrane protein made up of about 400 amino acids composed of eight transmembrane sections (Berry *et al.*, 2000). Though, the cyt b gene is highly conserved in some group of vertebrates, this is used in interspecific molecular systematic studies (Chen *et al.*, 2009). Mutation in cyt b may lead some adverse effects in various animals and reduces their fitness (Singh *et al.*, 2012).

*Received: 07 Feb 2022. Received in revised form: 10 Mar 2022. Accepted: 11 Mar 2022. Published online: 14 Mar 2022.*

From Volume 13, Issue 1, 2021, Notulae Scientia Biologicae journal uses article numbers in place of the traditional method of continuous pagination through the volume. The journal will continue to appear quarterly, as before, with four annual numbers.

The snapping shrimp, (Family: Alpheidae) is one of the most taxonomically and ecologically diverse groups of coral-reef fauna and sponge-dwelling shrimp with over 500 species worldwide, and many species are symbionts with a diverse range of larger host taxa such as corals, anemones, sponge, fishes and mangroves (Chace, 1988). Alpheid shrimps are commonly distributed in tropical and shallow waters and some species occur in deeper waters (Anker *et al.*, 2015). In Alpheidae family, about 4 genera and 23 species were recorded in India and many of these species may be cryptic species and need taxonomic revision (Jha *et al.*, 2019). When comparing to the genomic data, the proteomic information in Alpheid shrimps is very meagre till date.

Comparative modelling or homology modelling is a type of methodology to build an atomic-resolution model of a target protein by its amino acid sequence (Martí-Renom *et al.*, 2000). Comparative modelling can construct high-quality structural models, if the target and template are more similar. This method becomes more attractive in the field of bioinformatics because the insights of the 3-D structure of a protein would be a helpful support to understand the structural and functional aspects of a particular protein (Gupta *et al.*, 2009). Since there is no crystal structure or predicted structure of cyt b in Alpheid shrimps is available, this study was planned to predict the 3-D structure of cyt b in *Alpheus lobidens*, *A. randali* and *A. bellulus* and to share some important findings about this protein.

## Materials and Methods

### *Sequence retrieval and phylogenetic analysis*

Amino acid sequence of cyt b in *Alpheus lobidens*, *A. randali* and *A. bellulus* were retrieved from the National Centre for Biotechnology Information (<https://www.ncbi.nlm.nih.gov/protein>) in FASTA format. The genetic variation between the three species, phylogenetic tree using neighbour joining algorithm were analysed in MEGA 4.0 software (Tamura *et al.*, 2007).

### *Physicochemical characterization*

The physicochemical properties such as amino acid composition, isoelectric point (pI), molecular weight (MW), total number of positive (+R) and negative (-R) residues, instability index (II), aliphatic index (AI), and GRAVY (grand average of hydropathicity) values were identified by ExPASy - ProtParam. (<http://web.expasy.org/protparam/>) (Gasteiger *et al.*, 2005).

### *Secondary structure and conserved motif prediction*

The secondary structure was predicted using SOPMA (Self-Optimized Prediction Method with Alignment) (Geourjon and Deléage, 1995) tool by the default parameters (window width, 17; similarity threshold, 8; and number of states, 4). All domains and conserved protein motifs were analysed using a protein BLAST in the MEME (Multiple Em for Motif Elicitation) (<http://meme-suite.org/doc/fasta-format.html>) (Bailey *et al.*, 2009).

### *3-D structure construction and evaluation*

The modelling of the three-dimensional structure of the proteins was performed by Swiss model (Arnold *et al.*, 2006). The suitable template protein was selected by performing PDB-BLAST for the 3-D structure from the protein data bank. The PDB-BLAST result showed 50 templates and out of which, the 3-D structure of bovine cyt bc1 (6haw1.c) crystal structure showed more similarity (66.31%). For homology modeling, the Swiss-model server (<https://swissmodel.expasy.org/>) (Bordoli *et al.*, 2009) was used to model the cyt b structure using 6haw1.c as a template. The resulting model was saved as .pdb file and visualized by Swiss-pdb viewer (Guex and Peitsch, 1997). The stereochemical property ( $\phi$  and  $\psi$  distributions) and structural consistency of the predicted model was evaluated by Ramachandran plot analysis in RAMPAGE (<http://mordred.bioc.cam.ac.uk/~rapper/rampage.php>) (Lovell *et al.*, 2003). The predicted 3-D model was

validated by PROCHECK (Laskowski *et al.*, 1996), QMEAN (Quantitative Model Energy Analysis) (Benkert *et al.*, 2011) and QMEAN-Disco (Studer *et al.*, 2020). The transmembrane protein topology and helix probability curve was predicted by TMHMM server V.2.0 (<http://www/cbs.dtu.dk/services/TMHMM.2.0>)

#### *Tertiary structure analysis*

For tertiary structure of *A. lobidens* cyt b was analysed by PyMol (Schrodinger and DeLano, 2020). The quinone reduction (Qi) site residues such as His202, Lys228, and Asp229 were identified. The heme binding residues such as His84 and His183, His98 and His197 were verified and highlighted.

#### *Protein interaction analysis*

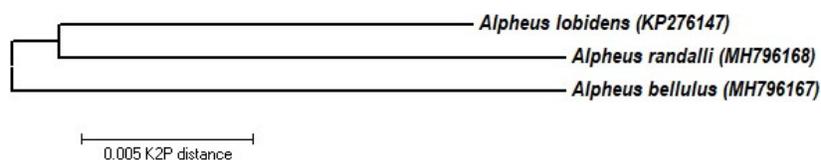
Protein-protein interactions are a central part of the cellular network and are known to have various impacts. To know the interaction about cyt b of *A. lobidens* with other closely related proteins, STRING v 10.0 (<http://string-db.org/>) (Szklarczyk *et al.*, 2015) Server was used. Cyt b of *A. lobidens* was selected as a query sequence and functional protein association network was generated.

## Results and Discussion

The K2P genetic distance between the three species is shown in Table 1. The maximum K2P distance (0.032) was found between *A. randalli* and *A. bellulus*, the mean genetic distance was 0.030. Figure 1 shows the phylogenetic relationship based on cyt b protein sequences using Neighbor - Joining Algorithm. The tree shows a significant phylogenetic relationship among the studied shrimp species.

**Table 1.** K2P genetic distance in Alpheid shrimps based on cyt b sequence

Species	<i>Alpheus lobidens</i>	<i>Alpheus randalli</i>	<i>Alpheus bellulus</i>
<i>Alpheus lobidens</i>	****		
<i>Alpheus randalli</i>	0.027	****	
<i>Alpheus bellulus</i>	0.030	0.032	****
Mean value	0.030		



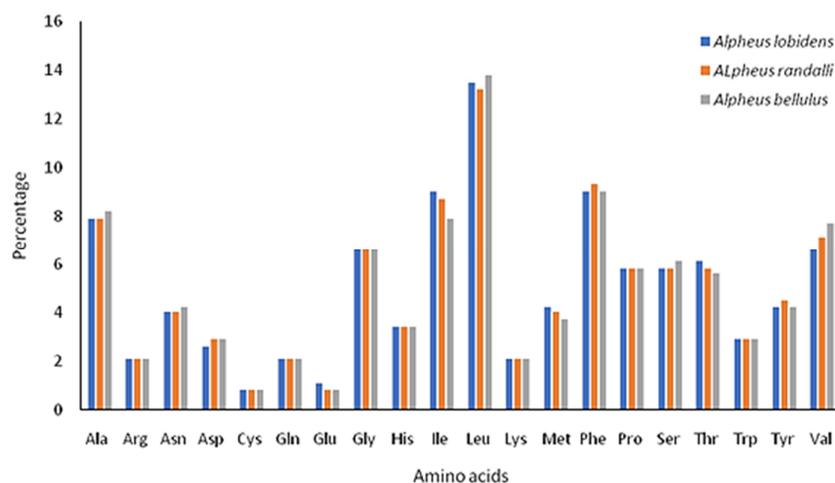
**Figure 1.** Neighbour joining tree shows the phylogenetic relationship of Alpheid shrimps

The physicochemical properties of cyt b in three shrimps were predicted using ExPASy - ProtParam and the results are shown in Table 2. The amino acid sequence size was 378 and the amino acid composition of this protein is given in Figure 2. Cyt b was composed the least percentage amino acid of Cys (0.8%) and highest percentage of Leu (13.8%). Totally, 14 positive charged and 16 negatively charged amino acids were recorded in this protein. The maximum molecular weight (MW) was predicted as 42.62 KDa in *A. randalli*. The theoretical pI value was ranged from 8.35 to 8.36 and confirmed that cyt b was alkali in nature. The instability index value was in the range of 42.29 to 46.94 which showed the protein was stable. The aliphatic index (AI) of the proteins of thermophilic bacteria has been found to be higher and the index could be used as a measure of thermostability of proteins. This index value is directly proportional to the number of Ala, Ile, Leu and Val in a protein (Idicula-Thomas and Balaji, 2005). In this study, these four amino acids were more in number and AI values were ranged from 114.29 to 115.05 (Table 2). The GRAVY was calculated to measure the protein

hydrophilicity/hydrophobicity. Hydrophobicity is revealed by a positive GRAVY value while hydrophilicity is revealed by the negative value (Kyte and Doolittle, 1982). Here, the GRAVY value was found to be 0.764, 0.770 and 0.771 indicated that cyt b in three shrimps were hydrophobic in nature.

**Table 2.** Physicochemical parameters of cyt b

Parameters	<i>Alpheus lobidens</i>	<i>Alpheus randali</i>	<i>Alpheus bellulus</i>
Accession No.	KP276147	MH796168	MH796167
Number of amino acids	378	378	378
Negatively charged residues	14	14	14
Positively charged residues	16	16	16
Molecular weight	42.58KDa	42.62KDa	42.43KDa
Theoretical pI	8.36	8.35	8.36
The instability index	46.94	46.60	42.29
Aliphatic index	114.81	114.29	115.05
GRAVY	0.770	0.771	0.764



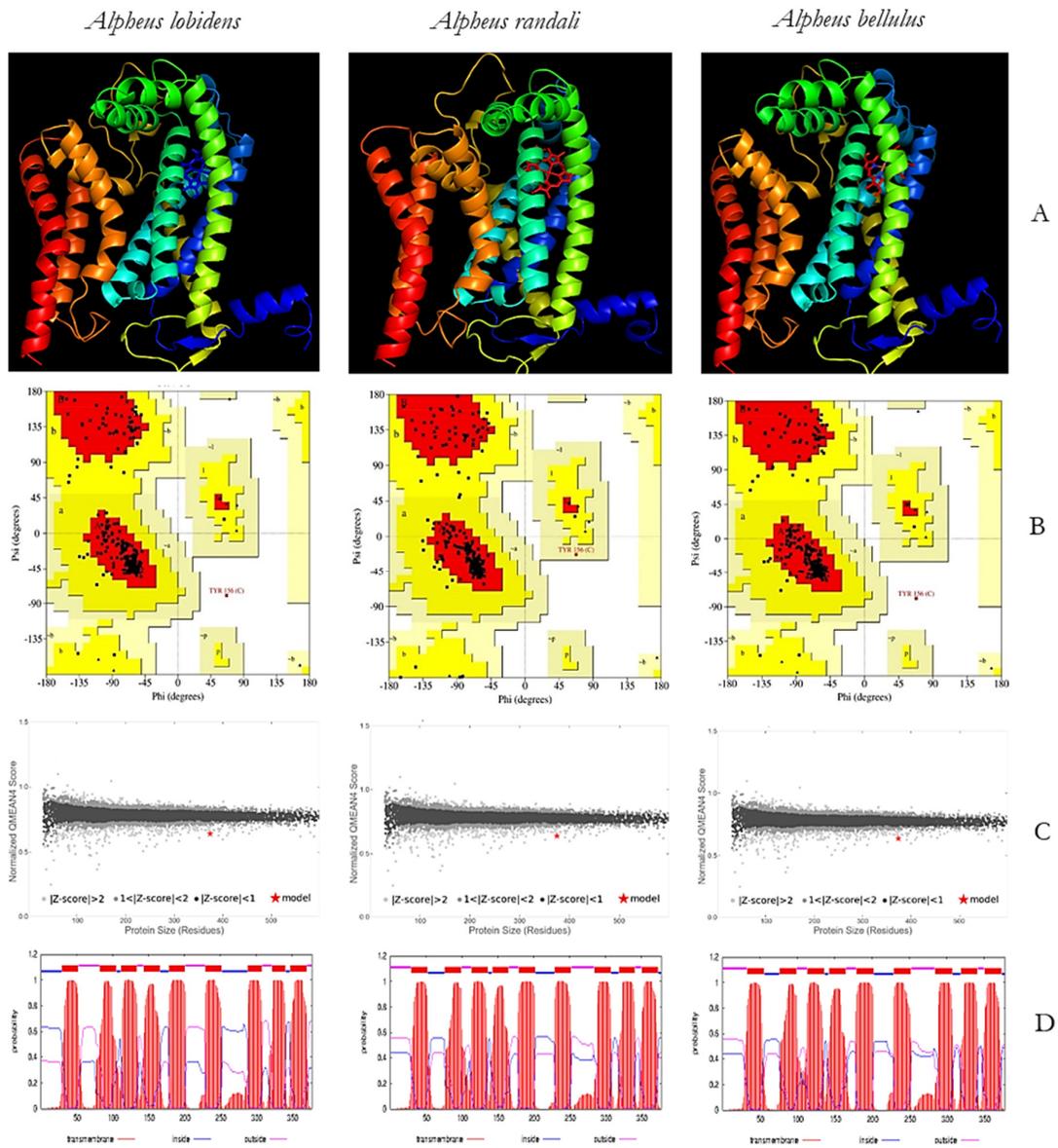
**Figure 2.** Amino acid composition of cyt b in three Alpheid shrimps

The secondary structure of cyt b was predicted by SOPMA and the results are given in Table 3. The secondary structure of this protein was primarily composed of  $\alpha$ -helices and random coil, revealed the stable structure. In *A. lobidens*, out of 378 amino acid residues, 37.30% residues formed  $\alpha$ -helix, 21.42% formed extended structure, 7.673% amino acids formed  $\beta$ -turn and 33.61% residues formed random coiled structure. In *A. randali*, among the total residues, 36.50% formed  $\alpha$ -helix, 21.16% formed extended, 7.67% formed  $\beta$ -turn and 34.67% residues formed random coiled structure. In *A. bellulus*, the cyt b protein showed 38.88%  $\alpha$ -helix, 20.63% extended structure, 7.14%  $\beta$ -turn and 33.35% random coiled structure. These results are in accordance with the earlier report of Gosh *et al.* (2020) in Indian major carp, *Labeo rohita*.

The transmembrane topology and helix probability curve of cyt b was predicted by TMHMM server and the result is presented in the Figure 3D. Here, the cyt b found to be a trans-membrane protein containing N-terminal signal peptide and 9 transmembrane helices. The domains and conserved motifs of this protein was analyzed by MEME and the results are depicted in Figure 4. Seven conserved motifs were observed in three protein sequences and the motifs width was ranged from 29 to 50. Similar type of transmembrane topology was reported in fishes (Ebenezer *et al.*, 2005; Gosh *et al.*, 2020) and lizard (Chen *et al.*, 2009).

**Table 3.** Secondary structure characters and Ramachandran plot computed with PROCHECK

Parameter	<i>Alpheus lobidens</i>	<i>Alpheus randali</i>	<i>Alpheus bellulus</i>
$\alpha$ helix	37.30%	36.50%	38.88%
Extended strand	21.42%	21.16%	20.63%
$\beta$ turn	07.67%	07.67%	07.14%
Random coil	33.61%	34.67%	33.35%
$\phi$ and $\psi$ distributions of amino acid residues			
In the most favoured region	94.20%	92.60%	95.10%
In additionally allowed region	05.50%	07.10%	4.60%
In disallowed region	0.30%	0.30%	0.30%



**Figure 3.** A. Three-dimensional structure; B. Ramachandran plot; C. QMEAN z-score; D. Transmembrane properties of cyt b in three Alpheid shrimps

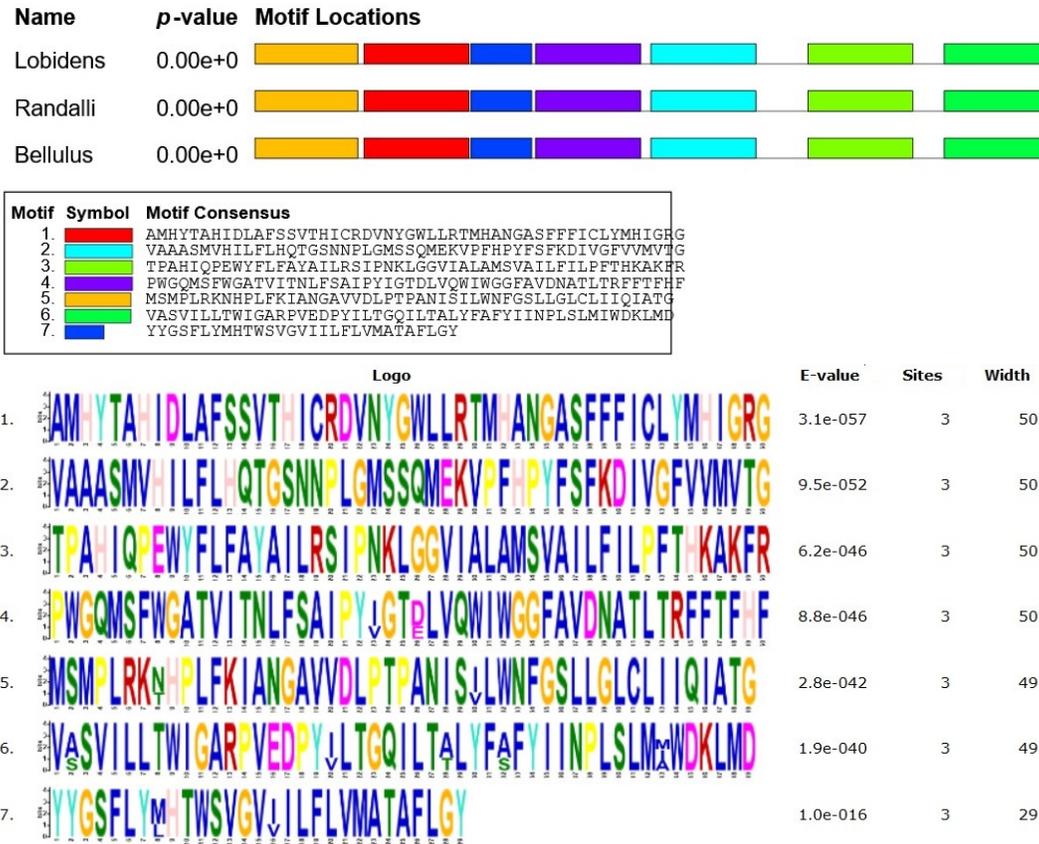


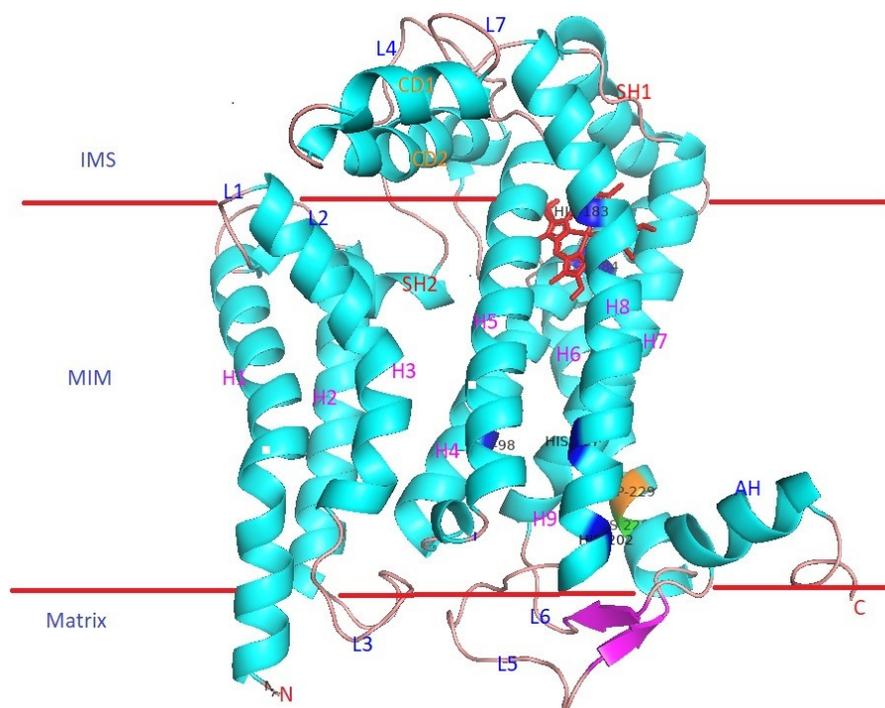
Figure 4. Conserved motif structure of cyt b in three Alpheid shrimps

The modeling of 3-D structure of the protein was performed by homology modeling program, Swiss-Model. The final modeled structure was visualized by Swiss-PDB Viewer and shown in Figure 3A. For model validation, the Ramachandran plot and PROCHECK were analysed. The geometrical and structural consistency of the predicted model was evaluated by different approaches. Ramachandran plot is a way to visualize backbone dihedral angles  $\psi$  against  $\phi$  of amino acid residues in protein structure. The  $\phi$  and  $\psi$  distributions of the Ramachandran map generated by non-glycine, non-proline residues were summarized in Table 3. Figure 3B. In *A. lobidens*, 94.2% of the amino acids were found in the favoured region, 5.5% were in the additionally allowed region and only 0.3% were in the outlier region. In *A. randali*, 92.6% were in the favoured region, 7.1% found in the additionally allowed region and 0.3% were in the outlier region. These findings reflect the results of previous studies on *Carangoides equula* (Ebenezer *et al*, 2005). In the case of *A. bellulus*, 95.1% were in the favoured region, 4.6% were in the additionally allowed region and the remaining residues were in the outlier region. Gly residue having a hydrogen atom as its variable (R) group and it can provide much flexibility for adjacent residues for conformational changes. Therefore, it is not surprising that Gly plays a crucial role in the structure and function of any protein (Yan and Sun, 1997). In this study, the highest amount of Gly residues (25) were detected among the three shrimps cyt b protein. The QMEAN score values are given in Table 4, which shows the quality of the modeled protein. The QMEAN-Z score values for c- $\beta$ , interaction, packing, torsion, S-S agreement and QMEAN<sub>4</sub> were confirmed the structural confidence.

The predicted result shows that the cyt b protein of *A. lobidens* spans the mitochondrial membrane with nine transmembrane (TM) helices with both the N- and the C-terminus are located in the mitochondrial matrix. In the tertiary structure, the nine helices (H1-H9) are arranged in two helical bundles, one consisting of three helices (H1-H3) and the other having six helices (H4-H9) (Figure 5).

**Table 4.** Structural validation by QMEAN

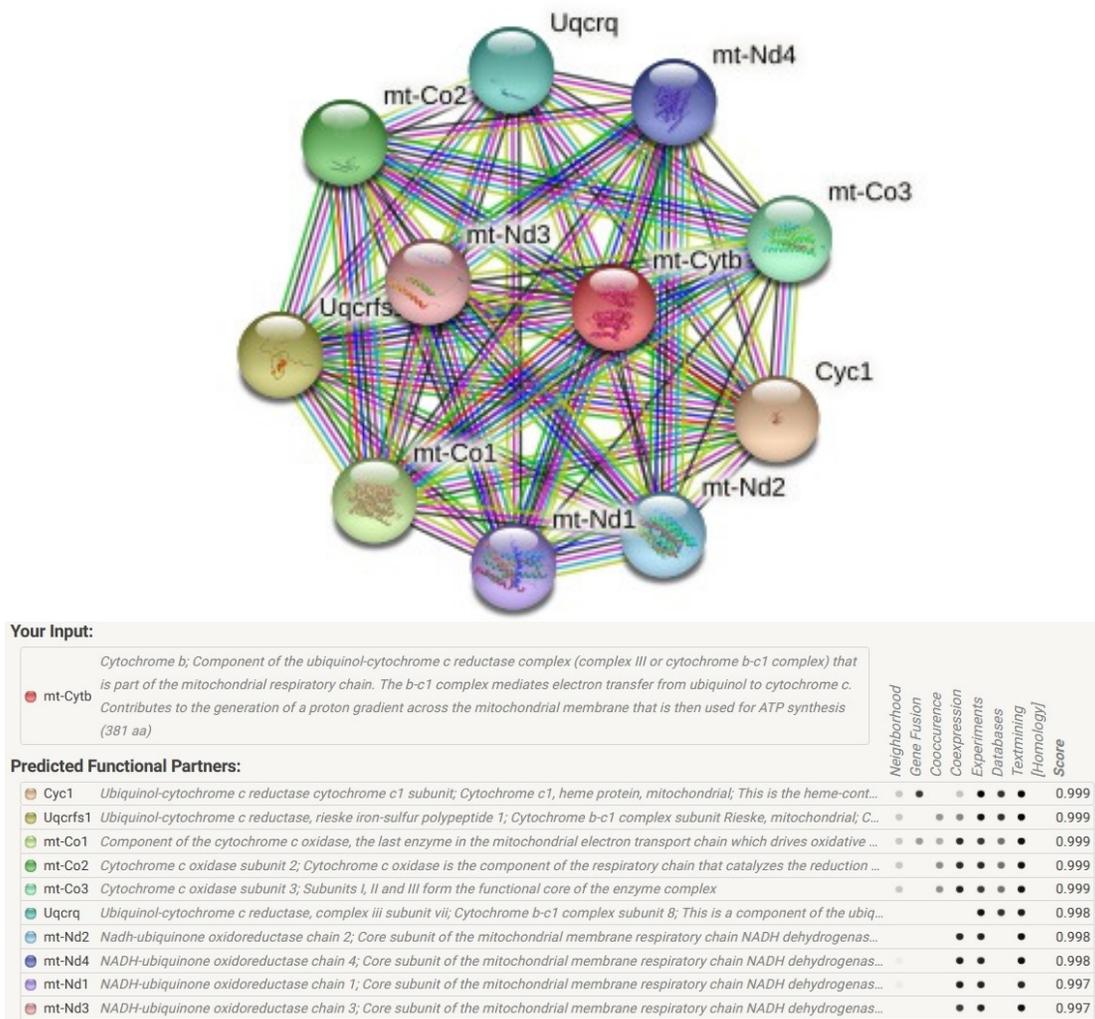
Seaweeds	<i>Alpheus lobidens</i>	<i>Alpheus randali</i>	<i>Alpheus bellulus</i>
GMQE	0.82	0.81	0.81
QMEAN Disco Global	0.74 ± 0.05	0.73 ± 0.05	0.74 ± 0.05
QMEAN score value	-3.45	-3.51	-3.52
Function terms	QMEAN Z-score value		
C-β	-3.9544	-4.9342	-4.0476
Interaction	-1.2228	-0.5978	-1.1120
Packing	0.2614	0.9365	0.2567
Torsion	-2.9554	-3.0991	-3.0038
SS-Agreement	0.8282	0.7550	0.8025
QMEAN4	-3.4504	-3.5190	-3.5098

**Figure 5.** Ribbon model of the cytb structure of *A. lobidens*

The demographic model is oriented with the matrix at the bottom, the intermembrane space (IMS) on the top, and the mitochondrial inner membrane (MIM) in the middle, which is roughly delineated with the two parallel lines. The conserved amino acids His84, His98, His183, His197, His202 are highlighted with blue colour, Lys228 green colour and Asp229 orange colour.

The transmembrane (TM) helices were connected by seven extra-membrane loops (L1-L7), including four long loops (L3, L4, L5, L6) and three short loops (L1, L2, L7). Near to the H1 and H7 transmembrane helices, two small helices (SH1, SH2) were found. Between the loops L4 and L7, two short helices forming a hairpin arrangement, namely, CD1 and CD2. Among these seven loops, L3, L5, L6 are in the matrix side, while the remaining four loops (L1, L2, L4, L7) are on the intermembrane space (IMS) side. In addition, a small helix (AH) was located very close to N-terminal. The cytb contains two bound hemes and two ubiquinol/ubiquinone (Qo/Qi) binding sites. The heme binding amino acids are His84, His183, His98 and His197, these four histidines are highly conserved in nature (Gao *et al.*, 2003). The four long loops maybe most

important to the function of cyt b, as they are the primary participants in the formation of the quinol oxidation (Qo) and quinone reduction (Qi) site. In Qi site, the residues such as, His202, Lys228, and Asp229 are highly conserved (Chen *et al.*, 2009). Gao *et al.* (2003) explained that the His202 binds to the carbonyl oxygen of the bound ubiquinone through a water molecule; Asp229 interacts with the other carbonyl oxygen via another water molecule that was stabilized by Lys228. Hacker *et al.* (1993) reported that, if any mutation occurs in these three residues may decrease the rate of quinone reduction. This suggests that these three residues play a crucial role in Qi binding site. In this study, it was noticed that, two short helices (CD1, CD2) existed between the loop L4 and L7. According to previous report by Chen *et al.* (2009) these L4 and L7 are the predominant portion for ISP (iron-sulfur protein) binding site, and the instability of these loops could affect the function of these two short helices (CD1, CD2). The protein-protein interaction (PPI) network of cyt b of *A. lobidens* is shown in Figure 6. By this PPI network, it clearly demonstrates, cyt b protein is interacted more with other mitochondrial proteins such as Nd1, Nd2, Nd3, Nd4, COI, COII and COIII.



**Figure 6.** The protein-protein interaction (PPI) network of cyt b of *A. lobidens* built by STRING database. Line thickness indicates the strength of the interaction. The lengths of the connectors indicate the farness or closeness of the flow of information. Redline: presence of fusion evidence, Greenline: neighbourhood evidence, Blueline: Co-occurrence evidence, yellow line: text mining evidence, Light blue line: database evidence, Blackline: co-expression evidence.

## Conclusions

Bioinformatics studies on protein, nucleic acid and other biomolecules are very supportive to solve many problems in almost all bio research fields. Advances in computing tools offer the opportunity to analyze the functional, physical and chemical properties of gene and gene product. Experiments used by sophisticated instruments to characterize a protein are an investment of high cost and time consumption. By molecular modeling technique, it is easy to design new drugs and molecules which accurately bind to targets. Comparative modeling is also a powerful method in the field of bioinformatics because the knowledge of 3-D structure of a protein would be an invaluable aid to delineate the structural and functional details of a particular protein. Though the Alpheids are hiding in burrows and not involved in active swimming they may conserve much energy. The hypothetical structures of cyt b in alpheid shrimps may provide a good basis for experimental analysis related to their oxygen consumption and energy utilization.

## Authors' Contributions

The author read and approved the final manuscript.

## Ethical approval (for researches involving animals or humans)

Not applicable.

## Acknowledgements

The author thanks the Ministry of Environment, Forests and Climate Change, New Delhi for the financial assistance in the form of research project (F.No.22018-30/2019-CS (Tax)). He also grateful to The Dean and The Director, CAS in Marine Biology and Authorities of Annamalai University for providing the necessary facilities and encouragement.

## Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

## References

- Anker A, Komai T, Marin IN (2015). A new echiuran associated snapping shrimp (Crustacea: Decapoda: Alpheidae) from the Indo-West Pacific. *Zootaxa* 3914:441-455. <https://doi.org/10.11646/zootaxa.3914.4.4>
- Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-Ma web-based environment for protein structure homology modelling. *Bioinformatics* 22:195-201. <https://doi.org/10.1093/bioinformatics/bti770>
- Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Noble WS (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37(2):202-208. <https://doi.org/10.1093/nar/gkp335>
- Benkert P, Biasini M, Schwede T (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27:343-350. <https://doi.org/10.1093/bioinformatics/btq662>
- Berry EA, Guergova-Kuras M, Huang LS, Crofts AR (2000). Structure and function of cytochrome bc complexes. *Annual Review of Biochemistry* 69:1005-1075. <https://doi.org/10.1146/annurev.biochem.69.1.1005>

- Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T (2009). Protein structure homology modelling using SWISS-MODEL workspace. *Nature Protocols* 4:1-13. <https://doi.org/10.1038/nprot.2008.197>
- Chace Jr. FA (1988). The caridean shrimps (Crustacea: Decapoda) of the Albatross Philippine Expedition, 1907-1910, Part 5: Family Alpheidae. *Smithsonian Contributions to Zoology* 466:1-99.
- Chen QL, Tang XS, Yao WJ, Lu SQ (2009). Bioinformatics analysis the complete sequences of cytochrome b of *Takydromus sylvaticus* and modeling the tertiary structure of encoded protein. *International Journal of Biological Sciences* 5(6):596-602. <https://doi.org/10.7150/ijbs.5.596>
- Ebenezer NS, Uma Maheswari M, Joyce Priyakumari C (2005). Homology modeling of cytochrome b from *Carangoides equula*. *International Journal of Science and Research* 6(2):92-95.
- Esposti MD, De Vries S, Crimi M, Ghelli A, Patarnello T, Meyer A (1993). Mitochondrial cytochrome b: evolution and structure of the protein. *Biochimica et Biophysica Acta* 1143(3):243-271. [https://doi.org/10.1016/0005-2728\(93\)90197](https://doi.org/10.1016/0005-2728(93)90197)
- Gao XG, Wen XL, Esser L, Quinn B, Yu L, Yu CA, Xia D (2003). Structural basis for the quinone reduction in the bc(1) complex: A comparative analysis of crystal structures of mitochondrial cytochrome bc(1) with bound substrate and inhibitors at the Qi site. *Biochemistry* 42(30):9067-9080. <https://doi.org/10.1021/bi0341814>
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005). Protein identification and analysis tools on ExPASy server. In: Walker JM (Ed). *Proteomics Protocol Handbook*. Humana Press, pp 571-607.
- Geourjon C, Deléage G (1995). SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics* 11:681-684. <https://doi.org/10.1038/s41598-017-07064-4>
- Ghosh R, Upadhyay AD, Roy AK, Tiwari A (2020). Structural and functional analysis of cytochrome b protein of Indian major carps *Labeo rohita*. *Journal of Entomology and Zoology Studies* 8(2):540-550.
- Guex N, Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18(15):2714-2723. <https://doi.org/10.1002/elps.1150181505>
- Gupta A, Pal SK, Narayan J, Potukuchi SM (2009). *In silico* structure prediction of p27Sj, a novel protein in St John's Wort, that suppresses expression of HIV-1 genome. *Biofrontiers* 1:32-35.
- Hacker B, Barquera B, Crofts AR, Gennis RB (1993). Characterization of mutations in the cytochrome-b subunit of the bc(1)-complex of *Rhodobacter sphaeroides* that affect the quinone reductase site (Qc). *Biochemistry* 32(16):4403-4410. <https://doi.org/10.1021/bi00067a033>
- Idicula-Thomas S, Balaji PV (2005). Understanding the relationship between the primary structure of proteins and their amyloidogenic propensity: clues from inclusion body formation. *Protein Engineering, Design and Selection* 18:175-180. <https://doi.org/10.1093/protein/gzi022>
- Jha PN, Purushothaman P, Chinnadurai S, Renjith RK, Remesan MP, Baiju MV, Edwin L (2019). Occurrence of *Alpheus euphrosyne* de Man, 1897 (Crustacea: Decapoda: Alpheidae) from the South-Eastern Arabian Sea, India. *Indian Journal of Geo Marine Sciences* 48(11):1695-1698.
- Kyte J, Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 157:105-142.
- Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR* 8:477-486. <https://doi.org/10.1007/bf00228148>
- Lovell SC, Davis IW, Arendall WB, Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2003). Structure validation by C $\alpha$  geometry:  $\phi$ ,  $\psi$  and C $\beta$  deviation. *Proteins: Structure Function and Bioinformatics* 50(3):437-450
- Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Sali A (2000). Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure* 29:291-325. <https://doi.org/10.1146/annurev.biophys.29.1.291>
- Schrodinger L, DeLano W (2020). PyMol. Retrieved from <http://www.pymol.org/pymol>
- Singh M, Gupta A, Lakra WS (2012). *In silico* 3-D structure prediction of cytochrome b protein of sisorid catfish *Glyptothorax ngapang*. *Indian Journal of Biotechnology* 11:156-162.
- Studer G, Rempfer C, Waterhouse AM, Gumienny G, Haas J, Schwede T (2020). QMEANDisCo - distance constraints applied on model quality estimation. *Bioinformatics* 36:1765-1771. <https://doi.org/10.1093/bioinformatics/btz828>

- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, ... Mering CV (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* 43:D447-D452.
- Tamura K, Dudley J, Nei M, Kumar M (2007). MEGA4: Molecular evolutionary genetics analysis (MEGA) software v. 4.0. *Molecular Biology and Evolution* 24:1596-1599. <https://doi.org/10.1093/molbev/msm092>
- Yan BX, Sun YQ (1997). Glycine residues provide flexibility for enzyme active sites. *Journal of Biological Chemistry* 272:3190- 3194. <https://doi.org/10.1074/jbc.272.6.3190>



The journal offers free, immediate, and unrestricted access to peer-reviewed research and scholarly work. Users are allowed to read, download, copy, distribute, print, search, or link to the full texts of the articles, or use them for any other lawful purpose, without asking prior permission from the publisher or the author.



**License** - Articles published in *Norulae Scientia Biologicae* are Open-Access, distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) License.

© Articles by the authors; SHST, Cluj-Napoca, Romania. The journal allows the author(s) to hold the copyright/to retain publishing rights without restriction.