World Journal of Fish and Marine Sciences 7 (2): 87-97, 2015
ISSN 2078-4589

DOI: 10.5829/idosi.wjfms.2015.7.2.92185

Insilico Characterization of Cytochrome C Oxidase I Protein of Himalayan Snow Trout, *Schizothorax richardsonii*: 3-D Structure Prediction

[1]*Suresh Chandra,* [2]*Sourabh Jain,* [3]*Javaid I. Mir and* [1]*Birendra K. Singh*

[1]Department of Zoology, Kumoun University, Nainital, Uttarakhand, India
[2]National Research Centre on Plant Biotechnology, (ICAR), New Delhi, India
[3]Directorate of Coldwater Fisheries Research, (ICAR), Bhimtal-263136, Nainital, Uttarakhand, India

Abstract: The Cytochrome c oxidase (COI) gene is the terminal member of the respiratory chain catalyzing the reduction of dioxygen to water by ferrocytochrome C. The present investigation includes insilico three-dimensional (3D) structure prediction of coldwater fish Indian snowtrout, *Schizothorax richardsonii* (Gray, 1832). The tertiary structures of COI were predicted using the comparative modelling method. The suitable template for comparative modelling protein databank (PDB IDs: 1V54|N|) has been selected on the basis of basic local alignment search tool (BLASTp). The target-template alignment, model building, loop modelling and evaluation have been performed in Modeller 9.10. The tertiary structure of COI is α-helix structure connected by loops, which forms a compressed complex maintained by two disulfide bridges. The resultant 3D models are verified by ConSurf and ProCheck programmes. Further, active site prediction was done using Qsitefinder. The computational models could be of use for further evaluation of molecular mechanism of function, structure and properties of protein COI. However, homology modelling of snow trout will give more insight on DNA barcoding.

Key words: Homology Modelling · Cytochrome C Oxidase Subunit-I · Ramchandran Plot · Coldwater Fish

INTRODUCTION

The Cytochrome c oxidase (COI) region of mitochondrial DNA is the most studied region of the fish mitochondrial genome. Cytochrome c oxidase is the terminal oxidase of cell respiration, a process that reduces molecular oxygen to water with the electrons from cytochrome c, coupled to pumping protons from the matrix side of the mitochondrial membrane toward the cytosolic side (intermembrane space) [1, 2]. This enzyme contains two iron sites and two copper sites (Fe_a, Fe_{a3}, CU_A and CU_B) in addition to zinc and magnesium sites [1, 2]. The protein moiety is composed of 13 different polypeptide subunits [3, 4] three encoded by mitochondrial genes and ten by nuclear genes [5]. Because of its physiological importance and the intriguing reaction catalyzed, this enzyme has been studied as one of the most important subjects in bioenergetics since its discovery [6].

Comparative modelling of proteins (unknown or little known structure) is done to build a three-dimensional model on one or more related proteins of known structure [7]. Prediction of proteins using computational tools becoming more important for deducing the biological functions involved in the mechanism via structure-function relationship. Structure knowledge is essential for all areas of protein research such as enzyme kinetics, ligand-protein binding studies, gene characterization and construction, structure based molecule design and rational designing of proteins [8]. In addition, these models can speed up the process of experimental structure determination by molecular replacement phasing in X-ray crystallography.

Schizothorax richardsonii (Actinopterygii: Cyprinidae) is a cold water fish, classified as vulnerable (VU) in India by IUCN [9]. The distribution of the cyprinid species is confined to the Himalayan and sub Himalayan streams and rivers of India. Beside this species is

Corresponding Author: Suresh Chandra, Department of Zoology, Kumoun University, Nainital, Uttarakhand, India.

distributed in Bhutan, Nepal, Pakistan and Afghanistan [10]. The studies indicated that it is abundantly distributed along the Himalayan foothills and commonly found, recent observations over the last 5 to 10 years (www.iucnredlist.org) point out drastic declines in its wild population due to the loss of habitat, introduction of alien species, damming, pollution, poisoning and destructive fishing, these factors not only destroyed the breeding and feeding grounds but also caused havoc to the biodiversity of this important fishery. While in some areas the declines are more than 90%, the overall reduction is inferred to be less than 50% with similar rates predicted in the future [9].

Fish mitochondrial DNA (mtDNA) all have a similar genomic organization [11, 12] and are similar to other vertebrates including humans. Many parts of mtDNA such as those coding for protein genes or regulatory part as the control region are used as genetic markers for measurement of intra species and inter species diversity. This quality is because of an increased mutation rate for mtDNA, relative to nuclear DNA, which results in an accumulation of many base substitutions over a long period of time, providing tools for taxonomic, evolutionary and phylogenetic research [13-15]. COX/COI (E.C. 1.9.3.1.) is the terminal member of the respiratory chain catalyzing the reduction of dioxygen to water ferrocytochrome C. Biological identification of species through DNA barcodes has become popular in recent years mt genome of animals represent a better target for analysis then the nuclear genome because of its lack of introns and its limited exposure to recombination and its haploid mode of inheritance, COI likely possess a greater range in phylogenetic signal than any other mitochondrial gene [16].

This study describes proton pumping mechanisms in COI of *S. richardsonii* and its validation by comparative protein modeling by COI gene from bovine. Further, this study is planned to analyse the structure function relationship of COI gene that was sequenced and characterized. A three-dimensional structure of COI from *S. richardsonii* was built on the basis of comparative homology using the software Modeller. Prediction of conserved residues and functional/active sites in modelled three-dimensional structure of protein was done by using ConSurf and Q-SiteFinder software, respectively.

MATERIALS AND METHODS

Sample Collection: In total, 5 specimens of *Schizothorax richardsonii* were collected from Alaknanda River of pauri Garhwal (Uttarakhand). The Muscle samples were collected and preserved in 95% v/v ethanol. All specimens were fixed in 10% v/v formalin in the field as a voucher.

DNA Extraction, Amplification and Sequencing: Genomic DNA was extracted with the 50mg muscle by using the phenol-chloroform extraction by Sambrook [17] and a fragment of the COI gene (655bp) was amplified and sequenced using the universal primers: F1 (5'-TCAACCAACCACAAAGACATTGGCAC-3'), R1 (5'-TAGACTTCTGGGTGGCCAAAGAATCA-3'), F2 (5'-TCGACTAATCATAA AGATATCGGCAC-3'), R2 (5'-ACTTCAGGGTGACCGAAGAATCAGAA-3'). The sequences were submitted to NCBI GenBank (FJ170777 to FJ170781).

Sequence Analysis: The DNA sequences having Phred value above 20 (99% accuracy) were considered as good quality and selected for analysis. Reverse sequences were reversed and complimented, so as to make it parallel to forward sequence and then both are aligned. Full length consensus sequences were made from forward and reverse strands for all samples and aligned using ClustalW [18]. The consensus sequences were blasted in NCBI GenBank for the nearest similar sequence matches. All sequences were aligned with ClustalW, a module integrated in software MEGA4 [19]. All the nucleotide sequences were converted to amino acid sequences using ExPASy by defining the reading frame. The resultant protein is 218 amino acid in length. The most common amino acid sequence, chosen for predicting sequence of this target protein was submitted to NCBI (Acc. No. FJ170777) UNIPROT ID (B5U491). It was ascertained that structure of COI in coldwater fishes is unavailable and therefore, the work of tertiary structure prediction through homology modelling was undertaken.

Three-Dimensional Computer Modelling: The comparative modelling of COI gene from *S. richardsonii*, accession number, FJ170777, includes following stages: (a) the target protein sequence was obtained from NCBI-Genpept database and subsequently submitted to BLASTp [20]. The database chosen for BLASTp was PDB [21], which resulted in the identification of 1V54|N| PDB as suitable template for creating full atom three-dimensional model of COI gene from *S. richardsonii*. (b) The three dimensional structure of target protein was predicted by using MODELLER 9.10 [22]. It is used for homology or comparative modeling of protein three-dimensional structures [22, 23]. The user provides an alignment of a sequence to be modeled with known related structures

and MODELLER automatically calculates a model containing all non-hydrogen atoms. MODELLER implements comparative protein structure modeling by satisfaction of spatial restraints [24, 25].

Recognition of Errors in Three-Dimensional Structures of Proteins: ProSA-web server [26] is widely used to check three dimensional models of protein structures for potential errors. Its range of application includes error recognition in experimentally determined structures [27- 29], theoretical models [30- 32] and protein engineering [33].

Analysis on the Crucial Residues of Theoretical Model of COI protein from *S. richardsonii*: The theoretical models were submitted to ConSurf [34], an automated web based server for the identification of functional region in proteins. The conservation grades are projected onto the molecular surface of these proteins to reveal the patches of highly conserved residues that are often important for biological function.

Prediction of Ligand Binding Sites of Theoretical Model of COI protein from *S. richardsonii*: Q-SiteFinder uses the interaction energy between the protein and a simple Vander Waals probe to locate energetically favourable binding sites. Energetically favourable probe sites are clustered according to their spatial proximity and clusters are then ranked according to the sum of interaction energies for the sites within each cluster [35].

RESULTS AND DISCUSSION

To explore the role of Cytochrome c oxidase I (COI) isolated from *S. richardsonii*, in aerobic cellular respiration, reducing dioxygen to water in a process coupled with the pumping of protons across the mitochondrial inner membrane, 3D structure was modeled to spot the active site and residues which are taking part in the mechanism. The *S. richardsonii* COI isolate (FJ170777) was used as target by using protein sequence similarity search against three-dimensional structural data banks. BLASTp was used as protein sequence similarity search engine which accepts input as protein sequence. Conserved domain sequence was submitted to this search engine and picked out its homologs with the help of PDB. Results of BLASTp revealed that COI from *S. richardsonii*, has 94.5% sequence similarity score with the PDB id 1V54|N| which was taken as template.

Three-dimensional structure of modeled COI from *S. richardsonii* reveals that like COI from bovine (1V54) located mainly in the transmembrane domain, it consists of 6 transmembrane helices, without any large extramembrane part [36], (Fig. 1). This subunit is cylindrical and is oriented perpendicularly to the membrane surface. Secondary structure assessment shows that COI from *S. richardsonii* encompasses α1 residues 1-24 (LYLLFGAWAGMVGTALSLLIRAE); α2 residues 34-69 (DQIYNVVVTAHAFVMIFFMVMPIMIGGFGNWLVPLM); α3 residues 78-96 (PRMNNMSFWLLPPSFLLLLASSM); α4 residues 125-153 (KRVHW), α5 residues 175-196 (FVWSVMITAVLLLLSLPVLAAGITMLLTD),α6residues 212-217 (ILYOHLF), similar to COI from bovine (1V54) (Fig. 2). In both the structures, same residues are participating in the α helices and no deviation whatsoever was found, apart from some similar residues (Fig. 2).

Earlier studies in bovine COI revealed that the shortest network and perhaps the most effective, as an electron transfer path between Cu_A and heme a, is a hydrogen bond system including His^{204} of subunit II, a ligand to Cu_A, a peptide bond between Arg^{438} and Arg^{439}

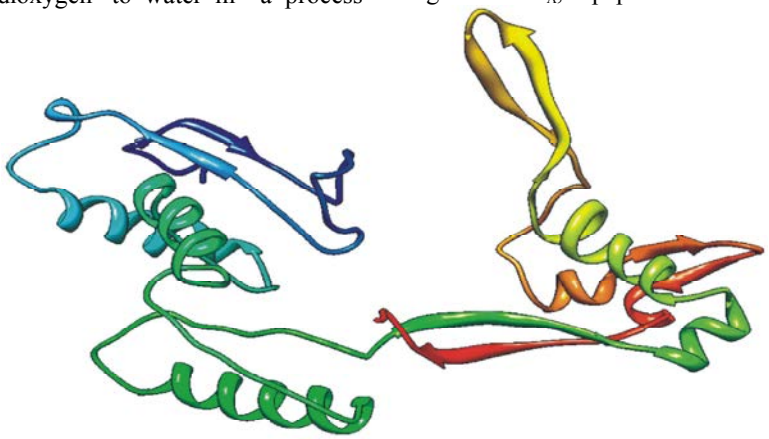

Fig. 1: Modeled structure of mtCOI from *S. richardsonii* by taking 1V54N as template. The representation is silhoutte, rounded ribbon view preset in CHIMERA tool.

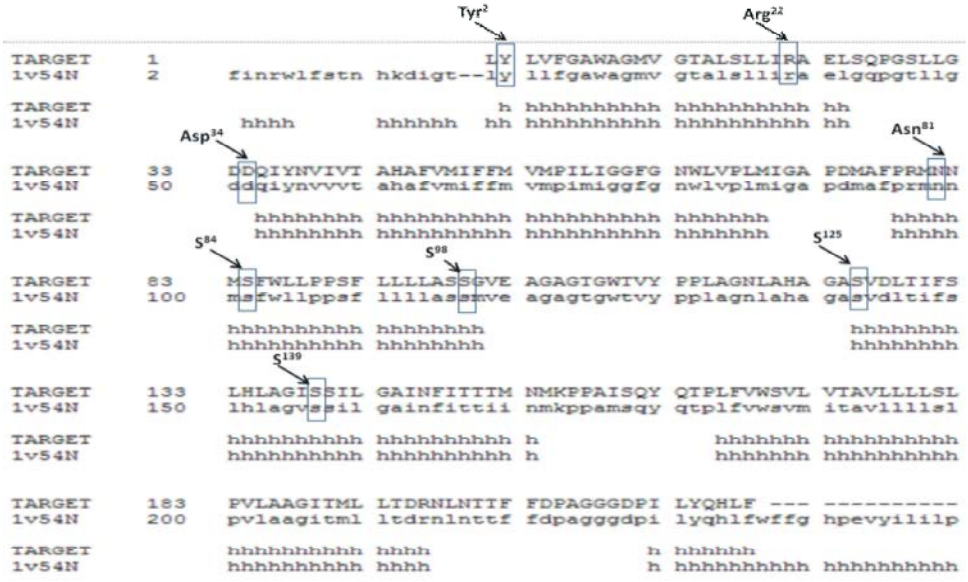

Fig. 2: Secondary structure assessment of modeled COI (target) with 1V54N. Conserved (functional) residues are highlighted. The position of the residues is written according to the modeled structure. The helices patterns are also shown in the figure.

of subunit I and a propanoite of heme a [37]. Similar hydrogen bonds have been found in the bacterial enzyme [38]. In case of COI from *S. richardsonii*, the hydrogen bond system could found to be conserved with the presence of Arg^{438} and Arg^{439} taking into account the similarity score and as the sequence is partial. These amino acids may control the electron transfer through the hydrogen bond systems or enhance the electron transfer through space between the redox sites [37].

Proton Pumping Mechanisms: Cytochrome c oxidase reduces O_2 to H_2O coupled with the proton pumping, making channels for protons, H_2O and O_2 a prerequisite to this function. Mutagenesis results suggest two types of proton channels, one for protons for H_2O formation and the other for proton pumping [39, 40]. Protons are transferred most effectively through hydrogen bonds, particularly in a hydrophobic environment such as the transmembrane region of this enzyme. The network was proposed in study on bovine COI as the candidates for the proton channels in the enzyme [37]. The main part of a network involves residues in helices III and IV of subunit I. At the entrance of this network are His^{503} and Asn^{11}, both in extended loops and hydrogen-bonded to a fixed water. The channel terminates at Ser^{142} connected to Ser^{115} with a possible hydrogen bond structure. The Ser^{115} could form a new hydrogen bond with Ser^{142} when proton release to the cytosolic side is required and thus serve as a system for preventing undesirable back flow of protons. This channel contains four other possible hydrogen-bond structures between Tyr^{19}, a fixed water hydrogen-bonded to Asn^{98}, Ser^{101} and Ser^{156}. If a conformational change around the almost completely conserved residue, Tyr^{19}, is induced by redox or ligand binding reactions coupled with a pK change of at least one of the amino acids in the network, it could provide unidirectional proton transfer. All the amino acids in the network are well conserved, although in some species Ser or Thr are replaced with alanine or glycine. Because of the small size of these residues, they could form cavities to contain water and still allow proton transfer. In case of COI from *S. richardsonii* the residues (Ser^{98}, Ser^{84}, Ser^{125}, Tyr^{2}, Asn^{81}, Ser^{139}) are found to be well conserved (Fig. 2) and thus, may be same functional residues are involved exhibiting similar proton pumping mechanism in *S. richardsonii*.

However, in the mutagenesis study [36], the mechanism for proton-pumping in bovine COI has been modified and is summarized as while heme *a* is in the oxidized state, Arg^{38} remains predominantly protonated even under the positive charge influence from heme *a*, because water molecules in the matrix space are accessible to Arg^{38} via the water channel. Asp^{51} is buried inside the protein and is protonated. Upon heme *a* reduction, which removes the net positive charge in heme *a*, Asp^{51} is exposed to the intermembrane space and the capacity of

the water channel increases so that water molecules are taken up from the matrix space, while the proton on Asp^{51} is released to the intermembrane space. Upon heme *a* oxidation, Asp^{51} moves back to the interior of the protein and the appearance of the net positive charge on heme *a* decreases the affinity of the formyl oxygen for the proton shared with Arg^{38}. This decreased affinity promotes proton transfer from Arg^{38} toward Asp^{51}. The propionate group hydrogen-bonded to the fixed water could also accelerate proton transfer along the hydrogen-bond network. The resulting deprotonated Arg^{38} extracts protons from water molecules in the water channel before the water channel capacity decreases to expel OH^-. The ability of the channel to rapidly undergo a reduction in water capacity and release OH^- immediately after water dissociation is expected to prevent reverse proton transfer. With the conservation of both Arg^{38} and Asp^{51} in the modelled COI as Arg^{21} and Asp^{34} (Fig. 2), the above said pumping mechanism involving low spin heme can also be responsible for proton-pumping in *S. richardsonii*, however, x-ray structures of the bovine and bacterial enzymes suggest the existence of a proton pump system driven by the low-spin heme apart from the O_2 reduction system in all cytochrome c oxidases [41-44]. Thus, the validation of proton pumping mechanism in fishes can be done only by confirmation through the availability of crystal structure.

Recognition of Errors in 3D Structure: The stereochemistry of the theoretical model of COI was done by PROCHECK, which is used for verification and evaluation of three-dimensional structure of proteins. The main output of PROCHECK is the Ramachandran plot. In the Ramachandran plot analysis, the residues were classified according to their regions in the quadrangle. The Ramachandran map for COI from *S. richardsonii* is represented in Fig. 3A. By the plot analysis, it can be noticed that more than 98.0% of the residues are in allowed regions, leading to a good validation for the model. The residues that are in bad quadrangles are a reflex from the protein used as template. A good quality model would be expected to have over 90% in the most favored regions.

A major problem in structural biology is the recognition of errors in experimental and theoretical models of protein structures. ProSA program (Protein Structure Analysis) is an established tool which has a large user base and is frequently employed in the refinement and validation of experimental protein structures and in structure prediction and modeling [45]. ProSA uses knowledge-based potentials of mean force to evaluate model accuracy [26]. After parsing the coordinates, the energy of the structure is evaluated using a distance-based pair potential [26, 46] and a potential that captures the solvent exposure of protein residues (26). From these energies, two characteristics of the inputs structure are derived and displayed on the webpage: its Z-score and a plot of its residues energies. The Z-score indicates overall model quality and measure the deviation of the total energy of the structure with respect to an energy distribution derived from random conformation [26, 47].

Z-score outside range characteristics for native proteins indicate erroneous structures. Groups of structures from different sources (X-ray, NMR) are distinguished by different colors. This plot can be used to check whether the Z-score of the protein in question is within the range of the proteins of similar size belonging to one of these groups. Fig. 3B shows the location of the Z-score. The value, -4.62 is in the range of native conformation. Overall, the residues energies are largely negative with the exception of some peaks in N-terminal part (Fig. 3C).

Analysis of Conserved Residues: The ConSurf server was used to extract information about important residues, which are of functional value. This server provides evolutionary related conservation scores for residues, which could be correlated with biological function. The continuous conservation scores of each of the amino acid position are available (Table 1) that are divided into a discrete scale of 9 grades for visualization purpose. Grade 1 contains the most variable positions and is colored light grey; grade 5 contains intermediately conserved position and colored white; and grade 9 contains the most conserved positions and is colored dark grey.

The color conservation grades are projected onto the three dimensional structure of the query protein and COI from Bovine (1V54). ConSurf server shows that 56 amino acid residues in COI from *S. richardsonii* and 67 amino acid residues in COI from Bovine are with high conservation score. In Table 1, amino acid residues (shown as bold), of COI from *S. richardsonii* and COI from Bovine shows the conservation score of 8 and 9.

From comparative analysis of residues, it is clear that most conserved residue with conservation score 8 and 9 and having functional value are conserved in COI from *S. richardsonii* and COI from Bovine. The residues involved in proton channels present in helices III and IV of subunit

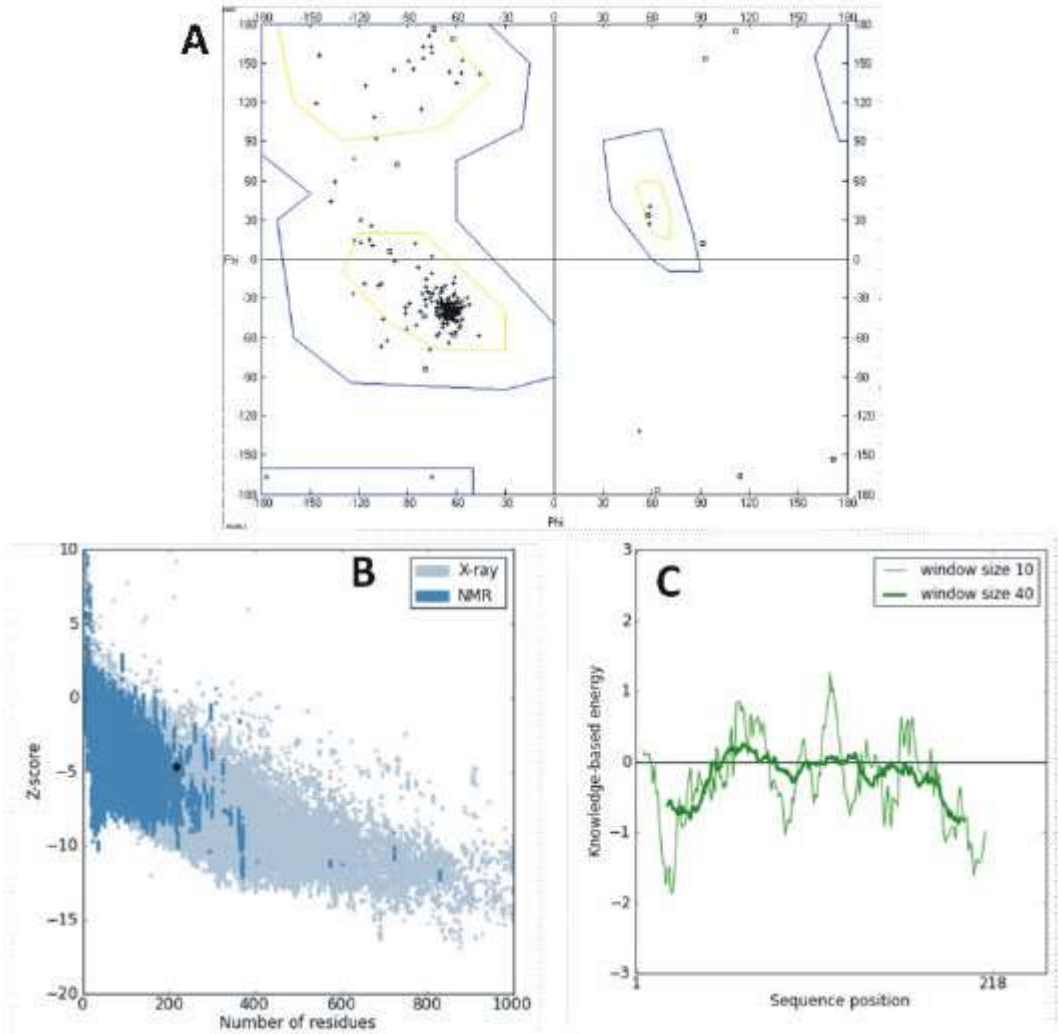

Fig. 3: (A): Ramachandran plot depicting ~99% residues in the allowed regions. (B) Z-score of the modeled structure (-4.62) lies in the range of native conformation. (C) Energy plot depicting largely negative residues energy.

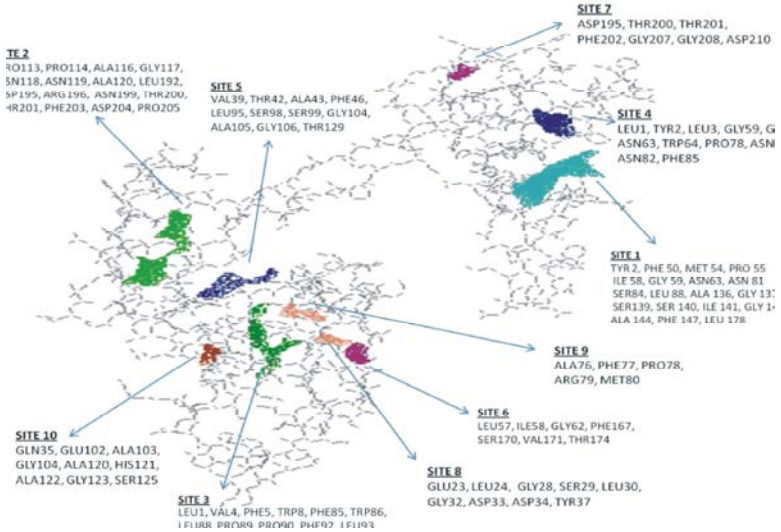

Fig. 4: Predicted Active sites by Qsite finder for the modeled structure of mtCOI from *S. richardsonii*

Table 1: Conservation score of the modeled mtCOI and COI from B. taurus (1V54|N|) showing the significant score for functional residues. Highly conserved residues with score 7-9 are shown in Bold

	Modeled mtCOI		1V54 chainN			Modeled mtCOI		1V54 chainN			Modeled mtCOI		1V54 chainN	
Residue No.	ATOM	Score	ATOM	Score	Residue No.	ATOM	Score	ATOM	Score	Residue No.	ATOM	Score	ATOM	Score
1.	L	4	L	6	73.	P	9	P	9	145.	I	1	I	4
2.	Y	9	Y	9	74.	D	9	D	9	146.	N	9	N	9
3.	L	1	L	5	75.	M	9	M	9	147.	F	9	F	9
4.	V	1	L	1	76.	A	9	A	9	148.	I	9	I	8
5.	F	1	F	5	77.	F	9	F	9	149.	T	4	T	6
6.	G	9	G	9	78.	P	9	P	9	150.	T	9	T	9
7.	A	1	A	1	79.	R	9	R	9	151.	T	1	I	1
8.	W	6	W	5	80.	M	1	M	4	152.	M	1	I	4
9.	A	1	A	3	81.	N	9	N	9	153.	N	9	N	9
10.	G	5	G	5	82.	N	9	N	9	154.	M	9	M	9
11.	M	2	M	5	83.	M	7	M	7	155.	K	7	K	8
12.	V	1	V	1	84.	S	9	S	9	156.	P	1	P	1
13.	G	9	G	9	85.	F	9	F	9	157.	P	1	P	1
14.	T	4	T	6	86.	W	9	W	9	158.	A	1	A	1
15.	A	3*	A	4	87.	L	3	L	6	159.	I	2	M	2
16.	L	1	L	2	88.	L	8	L	8	160.	S	1	S	1
17.	S	9	S	9	89.	P	9	P	9	161.	Q	1	Q	1
18.	L	1	L	1	90.	P	9	P	9	162.	Y	3*	Y	1
19.	L	1	L	2	91.	S	5	S	6	163.	Q	1	Q	2
20.	I	9	I	8	92.	F	1	F	1	164.	T	1	T	1
21.	R	9	R	9	93.	L	1	L	1	165.	P	8	P	9
22.	A	1	A	1	94.	L	1	L	7	166.	L	9	L	9
23.	E	9	E	9	95.	L	7	L	8	167.	F	9	F	9
24.	L	9	L	9	96.	L	1	L	1	168.	V	5	V	7
25.	S	1	G	3	97.	A	1	A	1	169.	W	9	W	9
26.	Q	1	Q	1	98.	S	8	S	9	170.	S	5	S	6
27.	P	8	P	9	99.	S	5	S	5	171.	V	8	V	7
28.	G	7	G	7	100.	G	1	M	1	172.	L	1	M	1
29.	S	1	T	1	101.	V	1	V	4	173.	V	5	I	5
30.	L	1	L	1	102.	E	3*	E	5	174.	T	9	T	9
31.	L	5	L	8	103.	A	1	A	1	175.	A	7	A	8
32.	G	1	G	5	104.	G	9	G	9	176.	V	1	V	1
33.	D	1	D	3	105.	A	1	A	6	177.	L	8	L	8
34.	D	9	D	9	106.	G	9	G	9	178.	L	9	L	9
35.	Q	9	Q	8	107.	T	9	T	9	179.	L	7	L	8
36.	I	4	I	6	108.	G	9	G	9	180.	L	5	L	7
37.	Y	8	Y	7	109.	W	9	W	9	181.	S	9	S	9
38.	N	9	N	9	110.	T	9	T	9	182.	L	8	L	8
39.	V	5	V	7	111.	V	7	V	7	183.	P	9	P	9
40.	I	3	V	2*	112.	Y	9	Y	9	184.	V	9	V	9
41.	V	8	V	9	113.	P	9	P	9	185.	L	9	L	9
42.	T	9	T	9	114.	P	9	P	9	186.	A	9	A	9
43.	A	7	A	8	115.	L	9	L	9	187.	A	8	A	9
44.	H	9	H	9	116.	A	4	A	5	188.	G	9	G	8
45.	A	9	A	9	117.	G	1	G	1	189.	I	9	I	9
46.	F	7	F	7	118.	N	1	N	1	190.	T	9	T	9
47.	V	1	V	1	119.	L	1	L	1	191.	M	9	M	9
48.	M	9	M	9	120.	A	2	A	4	192.	L	9	L	9
49.	I	9	I	9	121.	H	9	H	9	193.	L	9	L	9
50.	F	9	F	9	122.	A	1	A	2	194.	T	9	T	9
51.	F	9	F	9	123.	G	9	G	9	195.	D	9	D	9
52.	M	8	M	8	124.	A	1	A	1	196.	R	9	R	9
53.	V	9	V	9	125.	S	8	S	7	197.	N	9	N	9

Table 1: Continued

	Modeled mtCOI		1V54 chainN			Modeled mtCOI		1V54 chainN			Modeled mtCOI		1V54 chainN	
Residue No.	ATOM	Score	ATOM	Score	Residue No.	ATOM	Score	ATOM	Score	Residue No.	ATOM	Score	ATOM	Score
54.	M	9	M	9	126.	V	9	V	9	198.	L	1	L	1
55.	P	9	P	9	127.	D	9	D	9	199.	N	9	N	9
56.	I	2	I	4	128.	L	1	L	5	200.	T	9	T	9
57.	L	5	M	6	129.	T	1	T	5	201.	T	6	T	7
58.	I	8	I	9	130.	I	9	I	9	202.	F	9	F	9
59.	G	9	G	9	131.	F	9	F	9	203.	F	9	F	9
60.	G	9	G	9	132.	S	9	S	9	204.	D	9	D	9
61.	F	8	F	9	133.	L	9	L	9	205.	P	9	P	9
62.	G	9	G	9	134.	H	9	H	9	206.	A	1	A	3
63.	N	9	N	9	135.	L	7	L	7	207.	G	9	G	9
64.	W	9	W	8	136.	A	9	A	9	208.	G	9	G	9
65.	L	8	L	8	137.	G	9	G	9	209.	G	9	G	9
66.	V	1	V	5	138.	I	1	V	1	210.	D	9	D	9
67.	P	9	P	9	139.	S	9	S	9	211.	P	9	P	9
68.	L	6	L	6	140.	S	9	S	9	212.	I	6	I	6
69.	M	9	M	9	141.	I	9	I	9	213.	L	9	L	9
70.	I	5	I	6	142.	L	8	L	8	214.	Y	5	Y	6
71.	G	8	G	8	143.	G	7	G	8	215.	Q	9	Q	9
72.	A	8	A	9	144.	A	8	A	8	216.	H	9	H	9
										217.	L	9	L	9
										218.	F	9	F	9

I of Bovine COI (Ser^{142}, Ser^{115}, Asn^{98}, Tyr^{19}, Ser^{101} and Ser^{156}) are conserved in modeled COI too (Ser^{98}, Ser^{85}, Ser^{125}, Tyr^{2}, Asn^{81}, Ser^{139}). Further, residues which have been identified as key residues in the low spin heme proton pumping mechanism Arg^{38} and Asp^{51} in Bovine COI [36] are located close to the protein surface in the modelled COI as Arg^{21} and Asp^{34} and depict high conservation values. Thus, the findings are analogous with the proposed proton pumping mechanisms [36, 37].

Prediction of Ligand Binding Sites: The function of a protein is defined by the interaction it makes with other proteins and ligands. Computational methods for the detection and characterization of functional sites on protein have increasingly become an area of interest [48]. There is at least one successful prediction in the top three predicted sites in 90% of proteins tested when using Q-SiteFinder. Generally, ligand binding site prediction method analyzes the protein surface for pockets. The ligand binding sites are usually in the largest pocket. The pockets are defined only by energetic criteria. The method calculates the Vander Waals interaction energies of probe with protein. Probes with favorable interaction energies are retained and clusters of these probes are ranked according to their total interactions energies. The energetically most favorable cluster is then ranked first. The results of QSiteFinder studies are shown in Fig. 4. Most favorable binding sites contain amino acids with high conservation residue score. These residues may be involved with main binding site.

Hebert, proposed that a single gene sequence would be sufficient to differentiate all (or) at least the vast majority of animal species and proposed the use of mitochondrial DNA gene cytochrome c oxidase subunit 1 COI as a global bio identification system for animals popularly known as DNA bar coding, protein coding cytochrome c oxidase subunit (COI) gene are responsible well conserved proven to be robust evolution marker for determining inter specific relationship [16, 49, 50] worked on homology model for plant cytochrome P450S. They suggested that homology modelling represents a reliable and relatively rapid alternative method for analyzing structure-function relationships and predicting-substrates for many P450S (Cytochrome P450S monooxygenases. In most of the published results [16, 51], have worked on DNA barcoding. The Secondary and 3D structure of cytochrome c oxidase was totally ignored. The 3-Dimensional structure of cytochrome c oxidase (COI) in Indian Snow trout hitherto unreported. This prompted us to investigate the COI of mitochondrial gene in *Schizothorax* species. The homology modelling will give more insight on DNA bar coding. The insilico approach helps the researchers by giving them a hand-in idea, so that they can happily and easily work on the

structure predication of the target protein COI. This type of study is helpful to conclude the 3-D structure and functions of COI and also it is useful for DNA bar coding studies.

CONCLUSION

Cytochrome c oxidase I (COI) isolated from *S. richardsonii* was found to having the functional residues involved in Proton pumping mechanism of bovine COI fully conserved. By explanation of mechanism at 3D level, from the modeled structure of COI, it can be concluded that COI isolated from *S. richardsonii* is structurally similar to COI, a major subunit of Cytochrome Oxidase isolated protein from *Bovine* heart and contains same proton channel residues in helices III and IV, also found in major bacterial and animal enzymes. Further, residues essential for Proton pumping mechanism observed from sequence alignment are found to be well conserved in structure, thus, validating the built model.

ACKNOWLEDGEMENT

The authors are grateful to the fishermen and highly acknowledged for their support in fish collection.

REFERENCES

1. Malmstrom, B.G., 1990. Cytochrome c oxidase as a redox-linked proton pump. Chemical Reviews, 90: 1247-1260.
2. Babcock, G.T. and M. Wikstrom, 1992. Oxygen activation and the conservation of energy in cell respiration. Nature, 356: 301-309.
3. Kadenbach, B. and P. Merle, 1981. The function of multiple subunits of cytochrome c oxidase from higher eukaryotes. FEBS Lett, 135: 1-11.
4. Kadenbach, B., U. Ungibaver, J. Jaraush, U. Buge and L. Kuhn-Neutwig, 1983. The complexity of respiratory complexes. Trends Biochemical Sciences, 8: 398-400.
5. Capaldi, R.A., 1990. Structure and function of cytochrome c oxidase. Annual Review of Biochemistry, 59: 569-596.
6. Warburg, O. and E. Negelein, 1929. The absorption spectrum of the respiratory enzyme. Biochemistry, 214: 64-100.
7. Bajorath, J., R. Stenkamp and A. Aruffo, 1994. Knowledge-based model building of proteins: Concepts and examples. Protein Science, 2: 1798-1810.
8. Rahman, A., N. Chaturvedi, S. Sinha, N.P. Pandey, K.D. Gupta, S. Sundaram and A. Tripathi, 2013. Computational protein structure modeling and analysis of UV-B stress protein in *Synechocystis* PCC 6803. Bioinformation, 9(12): 639-644.
9. Vishwanath, W., 2010. Schizothorax richardsonii. In: IUCN 2013. IUCN Red List of Threatened Species Version 2013.1. (www.iucnredlist.org).
10. Talwar, P.K. and G.A. Jhingran, 1991. Inland fisheries of India and adjacent countries. Vol I & II, Oxford and IBH Publishing Company Pvt. Limited, New Delhi.
11. Lee, J.S., M. Miya, Y.S. Lee, C.G. Kim, E.H. Park and N.M. Aokiy, 2001. The complete DNA sequence of the mitochondrial genome of the self fertilizing fish *Rivulus marmoratus* (Cyprinnodontifome Rivuilidae) and the first description of the duplication of control region in fish. Gene, 280: 1-7.
12. Kim, I.C., S.O. Jung, Y.M. Lee, C.J. Lee, J.K. Park and J.S. Lee, 2005. The Complete mitochondrial genome of the ray fish *Raja porosa* (Chondrichtyes; Rajiidae). DNA Seq, 16: 187-194.
13. Avise, C., 2001. The History and Formation of Species. Harward University Press. Cambridge. Phylogeography.
14. Kartavtsev, K.P., T.J. Park, K.A. Vinnikov, S.N. Ivankov Sharina and J.S. Lee, 2007. Cytochrome b cyt-b gene sequence analysis in six flatfish species (Teleostei; Pleuronectide) with Phylogenetic and Taxonomic insights. Mar. Sci., 152: 757-773.
15. Wallace, D.C., 1992. Disease of the mitochondrial DNA. Annu. Rev. Biochem., 61: 1175-1212.
16. Hebert, P.D.N., A. Cywinska, S.L. Ball and J.R. Dewarrd, 2003. Biological identifications through DNA Barcodes. Proceeding of the Royal Society of London: Biological Sciences, 270: 313-321.
17. Sambrook, J. and D. Russell, 2001. Molecular Cloning: a Laboratory Manual, 3rd edn. Cold Spring Harbor Press, New York.
18. Thompson, J.D., D.G. Higgins and T.J. Gibson, 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22(22): 4673-6680.
19. Tamura, K., J. Dudley, M. Nei and S. Kumar, 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution, 24: 1596-1599.

20. Altschul, S.F., W. Gish, W. Miller, E.W. Myres and J. Lipman, 1990. Basic local alignment search tool. Journal of Molecular Biology, 215: 403-410.
21. Berman, H., K. Henrick and H. Nakamura, 2003. Announcing the worldwide Protein Data Bank. Nature Structural & Molecular Biology, 10: 980.
22. Eswar, N., M.A. Marti-Renom, B. Webb, M.S. Madhusudhan, D. Eramian, M. Shen, U. Pieper and A. Sali, 2006. Comparative Protein Structure Modeling with MODELLER. Current Protocols in Bioinformatics, John Wiley & Sons Inc, Supplement, 15: 5.6.1-5.6.30.
23. Marti-Renom, M.A., A. Stuart, A. Fiser, R. Sánchez, F. Melo and A. Sali, 2000. Comparative protein structure modeling of genes and genomes. Annual Review of Biophysics & Bimolecular Structure, 29: 291-325.
24. Fiser, A., R.K. Do and A. Sali, 2000. Modeling of loops in protein structures. Protein Science, 9: 1753-1773.
25. Sali, A. and T.L. Blundell, 1993. Comparative protein modeling by satisfaction of spatial restraints. Journal of Molecular Biology, 234: 779-815.
26. Sippl, M.J., 1995. Knowledge-based potentials for proteins. Current Opinion Structure Biology, 5: 229-235.
27. Banci, L., I. Bertini, F. Cantini, N. DellaMalva, T. Herrmann, A. Rosato and K. Wuthrich, 2006. Structural genomics of proteins involved in copper homeostasis. Journal of Biological Chemistry, 281: 141-147.
28. Llorca, O., M. Betti, J.M. Gonzlez, A. Valencia, A.J. Mrquez and J.M. Valpuesta, 2006. The three-dimensional structure of an eukaryotic glutamine synthetase: functional implications of its oligomeric structure. Journal of Structural Biology, 156: 469-479.
29. Telium, K., J.C. Hoch, V. Goffin, S. Kinet, J.A. Martial and B.B. Kragelund, 2005. Solution structure of human prolactin. Journal of Molecular Biology, 351: 810-823.
30. Panteri, D., A. Paiardini and F. Keller, 2006. A 3D model of Reelin subrepeat regions predicts Reelin binding to carbohydrates. Brain Research, 1116: 222-230.
31. Mansfeld, J., S. Gebauer, K. Dathe and R. Ulbrich-Hofmann, 2006. Secretory phospholipase A_2 from *Arabidopsis thaliana*: Insights into the three-dimensional structure and the amino acids involved in catalysis. Biochemistry, 45: 5687-5694.
32. Schnuchel, A., R. Wiltscheck, M.H. Czisch, M. Wllllmsky, P. Graumann, M.A. Marahiel and T.A. Holak, 1993. Universal nucleic acid-binding domain revealed by crystal structure of the *B. subtilis* major cold-shock protein. Nature, 364: 169-171.
33. Beissenhirtz, M.K., F.W. Scheller, M.S. Viezzoli and F. Lisdat, 2006. Engineered superoxide dismutase monomers for superoxide biosensor applications. Analytical Chemistry, 78: 928-935.
34. Glaser, F., 2003. Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information. Bioinformatics, 19: 163-164.
35. Alasdair, T., R. Laurie and M.R. Jackson, 2005. Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. Bioinformatics, 21(9): 1908-1916.
36. Tsukihara, T., 2003. The low-spin heme of cytochrome c oxidase as the driving element of the proton-pumping process. Proceedings of the National Academy of Sciences, 100(26): 15304-15309.
37. Tsukihara, T., H. Aoyama, E. Yamashita, T. Tomizaki, H. Yamaguchi, K. Shinzawa-Itoh, R. Nakashima, R. Yaono and S. Yoshikawa, 1996. The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 Å. Science, 5265: 1136-1144.
38. Iwata, S., C. Ostermeier, B. Ludwig and H. Michel, 1995. Structure at 2.8 A resolution of cytochrome c oxidase from *Paracoccus denitrificans*. Nature, 376: 660-669.
39. Thomas, J.W., A. Puustinen, J.O. Alben, R.B. Gennis and M. Wikstrbm, 1993. Substitution of Asparagine for Aspartate 135 in Subunit I of the Cytochrome bo Ubiquinol Oxidase of *Escherichia coli* Eliminates Proton Pumping Activity. Biochemistry, 32: 10923Boore, J.L., 1999. Animal mitochondrial genomes. Nucleic Acids Research, 27: 1767-1780.
40. Garcia-Horsman, J.A., A. Puustinen, R.B. Gennis and M. Wikström, 1995. Proton transfer in cytochrome bo3 ubiquinol oxidase of *Eschericia coli*. Second site mutations in subunit I that restore proton pumping in the mutant Asp135-Asn. Biochemistry, 34: 4428-4433.
41. Yoshikawa, S., K. Shinzawa-Itoh, R. Nakashima, R. Yaono, E. Yamashita, N. Inoue, M. Yao, M.J. Fei, C.P. Libeu and T. Mizushima, 1998. Redox-coupled crystal structural changes in bovine heart cytochrome c oxidase. Science, 280: 1723-1729.
42. Souliman, T., G. Buse, G.P. Bourenkov, H.D. Bartunik, R. Huber and M.E. Than, 2000. Primary structure of a novel subunit in ba3-cytochrome oxidase from *Thermus thermophilus*. EMBO Journal, 19: 1766-1776.

43. Ostermeier, C., A. Harrenga, U. Ermler and H. Michel, 1997. Structure at 2.7 Å resolution of the *Paracoccus denitrificans* two-subunit cytochrome c oxidase complexed with an antibody F_V?fragment. Proceeding of the National Academy of Sciences USA, 94: 0547-10553.
44. Svensson-Ek, M., J. Abramson, G. Larsson, S. Tornroth, P. Brzezinski and S. Iwata, 2002. The X-ray crystal structures of wild-type and EQ (I-286) mutant cytochrome c oxidases from *Rhodobacter sphaeroides*. Journal of Molecular Biology, 321: 329-339.
45. Wiederstein, M. and M.J. Sippl, 2005. Protein sequence randomization: efficient estimation of protein stability using knowledge-based potentials. Journal of Molecular Biology, 345: 1199-1212.
46. Sippl, M.J., 1990. Calculation of Conformational Ensembles from Potentials of Mean Force: an approach to the knowledge based prediction of local structures in globular proteins. Journal Molecular Biology, 213: 859-883.
47. Sippl, M.J., 1993. Recognition of Errors in Three-Dimensional Structures of Proteins. Proteins, 17: 355-362.
48. Campbell, S.J., N.D. Gold, R.M. Jackson and D.R. Westhead, 2003. Ligand binding: functional site location, similarity and docking. Current Opinion in Structure Biology, 13: 389-395.
49. Chandra, S., A. Barat, M. Singh, K.B. Singh and R. Matura, 2012. DNA Bar-Coding of Indian Coldwater Fishes of Genus *Schizothorax* (Family: Cyprinidae) from Western Himalaya. World Journal of Fish and Marine Sciences, 4(4): 430-435.
50. Rupasinghe, S. and M.A. Schuler, 2006. Homology modeling of plant P450s. Phytochemistry Reviews, 5: 473-505.
51. Kranthi, S., K.R. Kranthi, A.A. Bharose, S.N. Syed, C.S. Dhawad and E.K. Patil, 2006. Cytochrome oxidase I sequence of *Helicoverpa* (Noctuidae: Lepidoptera) species in India-Its utility as a molecular tool. Journal of Biotechnology, 5: 195-199.