



OPEN

## Cloning, computational analysis and expression profiling of steroid 5 alpha-reductase 1 (SRD5A1) gene during reproductive phases and ovatide stimulation in endangered catfish, *Clarias magur*

Deepak Agarwal<sup>1</sup>, Gulshan Kumar<sup>2</sup>, Mohd Ashraf Rather<sup>3</sup>✉ & Ishtiyaq Ahmad<sup>3</sup>

The cloning and characterization of the complete coding sequence of the *Clarias magur* SRD5A1 (CmSRD5A1) gene, which encodes an enzyme responsible for regulating steroid levels by converting testosterone into 5 $\alpha$ -dihydrotestosterone (DHT), have been successfully achieved. DHT plays a vital role in enabling the complete expression of testosterone's actions in neuroendocrine tissues. The ORF of the full-length cDNA sequence of SRD5A1 was 795 bp, translating into 265 amino acids, with a total length of 836 bp including UTRs. Like other vertebrates, the signal peptide analysis revealed that SRD5A1 is a non-secretory protein, and hydrophathy profiles indicated that it is hydrophobic in nature. The 3D structure of CmSRD5A1 sequence generated above was predicted using highly accurate AlphaFold 2 in Google Colab online platform. CmSRD5A1 contains seven transmembrane helices connected by six loops, with the N-termini located on the periplasmic side and C-termini on the cytosolic side. Structural superimposition with known bacterial and human SRD5As showed very high structural similarity. The electrostatic potential calculation and surface analysis of CmSRD5A1 revealed the presence of a large cavity with two openings one highly electropositive towards the cytosolic side and another relatively neutral towards the transmembrane region. The structural comparison revealed that the electropositive side of the cavity should bind to NADPH and the steroid hormone in the hydrophobic environment. Polar residues binding to NADPH are highly conserved and the same as known structures. The conserved residues involved in hydrogen bonding with the ketone group at C-3 in the steroids hence favoring  $\Delta 4$  double-bond reduction are identified as E66 and Y101. Our findings showed that SRD5A1 expression was lower during the spawning phase than the preparatory phase in female fish, while the administration of Ovotide (a GnRH analogue) resulted in up-regulation of expression after 6 h of injection in the ovary. In males, the lowest expression was observed during the preparatory phase and peaked at 16 h post- Ovotide injection in the testis. The expression of SRD5A1 in the brain of female fish was slightly higher during the Ovotide stimulation phase than the spawning phase. This study represents the first report on the cloning and characterization of the full-length cDNA of SRD5A1 in Indian catfish.

In neuroendocrine tissues, the metabolism of testosterone (T) involves two biologically significant pathways regulated by brain aromatase and 5 $\alpha$ -Reductase. Aromatase, which is a critical enzymatic complex, plays a key role in the last step of the synthetic steroidogenic pathway by converting aromatizable androgens into estrogens.

<sup>1</sup>Institute of Fisheries Post Graduate Studies, TNJFU, Kazhipattur, India. <sup>2</sup>College of Fisheries, Gumla, Phasia, India. <sup>3</sup>Division of Fish Genetics and Biotechnology, Faculty of Fisheries, SKUAST-Kashmir, Srinagar, India. ✉email: biotechashraf786@gmail.com

In several fish species, two aromatase genes have been identified: the *cyp19a1a* gene that codes for aromatase A, primarily expressed in the gonads, and the *cyp19a1b* gene that codes for aromatase B, mainly expressed in the brain<sup>4–5</sup>. On the other hand, 5 $\alpha$ -reductase, also known as 3-oxo-5 $\alpha$ -steroid 4-dehydrogenases, is a microsomal enzyme that reduces the C4–C5 double bond of various steroid substrates<sup>6,7</sup>. It converts testosterone into the more potent androgen 5 $\alpha$ -dihydrotestosterone (DHT)<sup>8</sup>. In rodents, monkeys, humans, and fishes (including zebrafish and medaka), two isozymes of 5 $\alpha$ -reductase encoded by distinct genes, namely 5 $\alpha$ R1 and 5 $\alpha$ R2, have been identified<sup>9–11</sup>. The PMET (Phospholipid methyltransferase) superfamily encompasses 5 $\alpha$ -reductase, a microsomal enzyme responsible for the conversion of testosterone into dihydrotestosterone which is a crucial reaction for androgen action, contributing to the formation of the male phenotype during embryogenesis and the growth of tissues like the prostate in humans<sup>12</sup>. Although aromatase in fish brain has been widely studied, limited reports exist for 5 $\alpha$ -reductase<sup>13,14</sup>. The steroidogenic metabolite DHT, produced by 5 $\alpha$ -reductase, has received little attention in non-mammalian species, as it is thought to be relatively biologically inactive in fishes. However, recent studies have demonstrated the potency of DHT as an androgen in fish<sup>15</sup>. Moreover, the 5 $\alpha$ -reductase has been suggested to be necessary for the full expression of testosterone actions in neuroendocrine tissues<sup>16</sup>. Despite its importance, the 5 $\alpha$ -reductase gene has not yet been characterized in fishes of the Clariidae family, and no earlier studies are available in any species of walking catfish.

*Clarias magur*, commonly known as the Indian walking catfish, is a seasonal breeder that reproduces during the rainy season from July to August<sup>17</sup>. In captivity, obtaining milt from male silurid species, including *Clarias magur*, is difficult and requires sacrificing the brooders<sup>18</sup>. In contrast, this fish breeds easily in its natural environment. The underlying physiological mechanisms of reproduction in this species remain largely unexplored. Despite a few reports on genes involved in reproductive physiology<sup>19–25</sup>, there is a lack of information on the molecular mechanisms involved in spawning. Therefore, the present study aims to characterize the coding sequences of 5 $\alpha$ -reductase in *Clarias magur* and investigate its expression profiling at different reproductive stages to explore its potential role in spawning. This investigation will provide new insights into the reproductive biology of *Clarias magur* and could contribute to the development of improved breeding strategies for this economically important fish species.

## Materials and methods

### Sample and tissue collection

The wet lab experiment was conducted at the Freshwater Fish Farm, ICAR-CIFE, Powarkeda, Madhya Pradesh, India. Four reproductive stages of *C. magur* were examined: (i) preparatory, (ii) mature, (iii) 6 h post Ovatide injection, and (iv) 16 h post Ovatide injection. A total of six male and six female live fish were sampled for each stage, with sampling performed based on the maturation month of the fish. For example, fish in the preparatory stage were sampled in February, while those in the mature stage were sampled in August<sup>17</sup>. All experimental protocols were approved by ICAR- Central Institute of Fisheries Education, Mumbai. To anesthetize the fish, clove oil @ 0.05 mL per 500 mL of water was used. To examine the reproductive genes profile, Ovatide hormone (Hemmo Pharmaceuticals Pvt. Ltd., Mumbai) was administered to male and female fishes at a rate of 0.5 ml kg<sup>-1</sup> and 1.0 ml kg<sup>-1</sup> per body weight<sup>26</sup>, respectively. Sampling was carried out after 6 h and 16 h, and both brain and gonadal tissues were collected in RNAlater (QIAGEN, USA) and stored at – 80 °C until further analysis. RNA was extracted from tissue pools comprising three specimens.

### Total RNA extraction and cDNA conversion

The total RNA from the tissue samples was extracted using Trizol reagent (Invitrogen, USA) following the manufacturer's guidelines. Genomic DNA contamination was removed from the extracted RNA (5  $\mu$ g) by treating it with DNase I (Thermo Scientific, USA). The purified RNA was evaluated for quality and quantity using gel electrophoresis on 1% Agarose and NanoDrop spectrophotometer (Thermo Scientific, USA). To synthesize cDNA, DNase I-treated total RNA (1  $\mu$ g) was reverse transcribed using Oligo dT primer and RevertAid reverse transcriptase First Strand cDNA Synthesis kit (Thermo Scientific, USA) according to the manufacturer's instructions<sup>19</sup>.

### Molecular cloning and sequence analysis of SRD5A1

The conserved regions of SRD5A1 genes were identified by multiple sequence alignment with orthologous sequences, and primers were designed against them through Primer3 plus online primer designing tool (Table 1) and synthesised from Eurofins Genomics India Pvt. Ltd. For amplification, the PCR was performed in a 96-well Takara PCR System with a total volume of 25  $\mu$ L. The obtained partial cDNA fragment of 382 bp was eluted from the gel, cloned into pTZ57R/T vector, sequenced and confirmed by BLASTn similarity search. For RACE PCR, gene-specific primers (Table 1) were designed from the identified partial sequence. The reaction was carried out for both 3' and 5' ends to obtain the full-length cDNA sequence of the gene comprising of the open reading frame (ORF), 3' and 5' untranslated (UTR) regions. The desired PCR product was purified, cloned, and sequenced, and the sequence was confirmed by BLASTn software available at <http://www.ncbi.nlm.nih.gov/blast>. The annealing temperature of each primer set was optimized by gradient PCR. The reaction mixture contained 2.5  $\mu$ L of 10 $\times$  Taq buffer, 0.5  $\mu$ L of 10 mM of each dNTP mix, 2.5  $\mu$ L of 25 mM MgCl<sub>2</sub>, 1  $\mu$ L each of sense and antisense primers (10 pmol), 1  $\mu$ L of cDNA (500 ng), 0.25  $\mu$ L of 5 U/ $\mu$ L Taq DNA polymerase, and 16.25  $\mu$ L of nuclease-free water<sup>19</sup>.

### ORF identification and Amino acid sequence prediction

The open reading frame (ORF) of the obtained full-length cDNA sequence was identified using the NCBI ORFfinder tool. Gene Runner software was used to deduce the amino acid sequence of the coding region and

Primer name	Sequence (5'-3')	Usage
CmSRD5A1F	TCGAGGAGGAAAGCCACAC	Partial cDNA amplification
CmSRD5A1R	GGCTCGACTGGAAGGACGA	Partial cDNA amplification
GSP1 3'	CTTCCCTAAGGGTTGGGTTTCACATCC	3' RACE for CmSRD5A1
GSP2 3'	CACAGTGTGCCTTTGCTTTCTTCACC	3' RACE for CmSRD5A1
GSP1 5'	CTCAAAGGCAGAAATAACGCAGGCAGC	5' RACE for CmSRD5A1
GSP2 5'	GCCAGAGCCATGAGAGTCATTAATAAC	5' RACE for CmSRD5A1
CmSRD5A1F_qRT	TGGAGGACATTCTGTTGGGGCTGT	Real time PCR
CmSRD5A1R_qRT	CACACGCTGCCAGAGCCATGAG	Real time PCR
B_actinF	GCCGAGAGGGAAATTGTCCGTG	Internal control
B_actinR	GCCAATGGTGATGACCTGTCCG	Internal control

**Table 1.** Primers used in the study.

conserved domains in the deduced amino acid sequence were identified using the NCBI Conserved Domain Database (CDD) search tool (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>)<sup>19</sup>.

### Evolutionary analysis

The amino acid sequence of SRD5A1 from various species including *C. magur* (Indian magur), *I. punctatus* (Channel catfish), *L. calcarifer* (Sea bass), *P. nattereri* (Piranha), *D. rerio* (Zebrafish), *C. carpio* (Common carp), *S. salar* (Atlantic salmon), *Homo sapiens* (human), *Mus musculus* (Mouse) and *Bos taurus* (Cattle) was analyzed for multiple sequence alignment using the CLC Genomics Workbench software (QIAGEN, USA). The resulting alignment was utilized to construct a Neighbor-joining tree using MEGA 11.0 software<sup>27</sup> with 1000 bootstrap replicates.

### In silico analysis and 3D structure prediction of CmSRD5A1

The deduced amino acid sequence of CmSRD5A1 underwent analysis using various tools. Physicochemical parameters were evaluated using the ProtParam tool at ExPASy<sup>28</sup>. The secondary structure was predicted through online tools including PSIPRED V.4.0 (<http://bioinf.cs.ucl.ac.uk/psipred/>) and SOPMA<sup>29</sup>. The protein's hydrophathy profile was examined using the ProtScale program<sup>28</sup>. Furthermore, signal peptides were predicted utilizing SignalP 4.0 software<sup>30</sup>. The AlphaFold 2 software<sup>31</sup> is available on the ColabFold platform (<https://colab.research.google.com>) and was utilized to predict the 3D structure of the CmSRD5A1 deduced sequence (accession no. BBE00816.1). AlphaFold is a highly accurate neural network-based tool for protein structure prediction, employing training procedures that incorporate evolutionary, physical, and geometric constraints of protein structures. In the 14th Critical Assessment of Protein Structure Prediction (CASP14) competition, AlphaFold demonstrated accuracy comparable to experimental structures in most cases and significantly outperformed other prediction methods<sup>32</sup>. ColabFold, hosted on Google's Jupyter Notebook platform<sup>33</sup>, provides a free and accessible environment for performing protein structure prediction by combining the fast homology search method MMseqs2 with AlphaFold2 or RoseTTAFold, thereby accelerating the prediction process<sup>34</sup>. The structure was predicted in no template mode. The top-ranked structure was relaxed using amber. Multiple sequence alignment (MSA) options were kept as msa\_mode: mmseqs2\_uniref\_env and pair\_mode: unpaired\_paired (default settings). Parameters of advanced and sample settings were kept as auto. Post-translational modifications (PTM) were not considered during structural prediction as no PTMs were reported for this enzyme in available literatures.

### Model evaluation and Protein–protein interactions of CmSRD5A1

The quality assessment of the predicted 3D structure of CmSRD5A1 involved the analysis of the Ramachandran plot and AlphaFold quality estimates. The Ramachandran plot illustrates the structure's geometry by representing torsion angles on a residue-wise basis. To evaluate the quality, the Procheck tool<sup>35</sup> available on SAVESv6.0 (<https://saves.mbi.ucla.edu/>) was utilized. This tool provides information on the total number and percentage of residues distributed in allowed and disallowed regions of the plot.

AlphaFold, on the other hand, assesses errors in predicted structures using two approaches: one that verifies if a residue is correctly positioned within its local environment, and another that determines if domains are accurately positioned relative to each other<sup>36</sup>. The predicted local distance difference test (pLDDT) score, ranging from 0 to 100, serves as a per-residue confidence score. A pLDDT score above 90 indicates a high level of confidence, while a score below 50 suggests low confidence. This score estimates whether the predicted residue's distances to neighbouring C-alpha atoms (within 15 Angstroms) align with the distances observed in the true structure. The study also found that a pLDDT < 50 is a reliable indicator of disorder, implying that such a region is likely to be unstructured in physiological conditions or only structured when part of a complex<sup>36</sup>. Another metric used in the study is the Predicted Aligned Error (PAE), which measures the distance error between pairs of residues<sup>36</sup>. PAE provides AlphaFold's estimation of the positional error at residue x when aligning the predicted and true structures based on residue y. These values typically range from 0 to 35 Å and are represented as a heatmap image, where the vertical and horizontal axes correspond to residue numbers and the color at each pixel indicates the PAE value for the corresponding pair of residues. When the relative position of two domains

is confidently predicted, the PAE values for residue pairs involving one residue from each domain tend to be low (less than 5 Å).

The protein–protein interactions of the *CmSRD5A1* protein with other proteins involved in the steroidogenesis pathway were determined using the online software STRING<sup>37</sup>, which predicted the interactions (Fig. 1).

### Structural comparison, ligand interaction and visualization

The enzymes 3D structures were analyzed using UCSF-Chimera software<sup>38</sup>. To superimpose *CmSRD5A1* onto human SRD5A2 (HsSRD5A2; PDB ID: 7bw1) and SRD5A of *Proteobacteria* bacterium (PbSRD5; PDB ID: 7c83), the Matchmaker tool of Chimera software was utilized, with the latter serving as the reference structure. The alignment algorithm employed was Needleman-Wunsch with the BLOSUM-62 matrix. Atom pairs exceeding 2.0 Å were pruned iteratively until none remained. Following the superimposition, a structure-based sequence alignment was generated, setting a residue-residue cut-off of 5 Å. NADPH and NADP-dihydrofinasteride adduct molecules were then added to the *CmSRD5A1* structure. Ligand binding and catalytic residues were identified in the *CmSRD5A1* structure based on the structure-based sequence alignment. Additionally, various images depicting the structure and ligand interactions were generated.

### Calculation of electrostatic potential

The 3D structure of *CmSRD5A1* was initially prepared using the PDB2PQR tool available on the APBS-PDB-2PQR server (<https://server.poissonboltzmann.org/>). The resulting PDB2PQR output was then utilized to calculate the electrostatic potential through the APBSA tool in auto-mode. APBS (Adaptive Poisson-Boltzmann Solver) is an electrostatics solver that addresses the equations of continuum electrostatics for large biomolecular assemblies. The electrostatic potential obtained from the PDB2PQR server was visualized using the server's

```

1 ← M E D I L L G L F S S E E E E L Y M L
1 AGCATGAGGACATTCTGTTGGGGCTGTTTTCTCTGAAGAGGAAGAACTGTACATGTT
20 D S L S Y L M T L M A L A A C V A L L F
61 GGACAGTCTTTTCGTATTTAATGACTCTCATGGCTCTGGCAGCGTGTGTTGCTCTGCTGTT
40 L D V P Y G R Y T S R K Y G V Q V H V K
121 TCTGGACGTCCCTTATGGCAGGTACACTTCCCCTAAATACGGGGTTCAGGTTTCATGTTAA
60 V A W F I Q E L P A L F L P L S L I L W
181 AGTGGCGTGGTTTATTCAAGAGCTGCCTGCGTTATTTCTGCCTTTGAGTTTAATATTATG
80 A S P A K I S N L P N Q V L L G M Y F C
241 GGCTTCGCCCCGGAAGATATCAAACCTTCCCTAACCCAGGTGCTTCTTGGCATGTACTTCTG
100 H Y F Q R S L I Y P F L I R G G K P T P
301 TCATTATTTCCAAAGGTCTCTCATATATCCCTTTCTAATTCGAGGAGGAAAGCCACACC
120 F A S F A L A F V F C I Y N G Y L Q V R
361 CTTTGCATCCTTTGCTCTGGCATTGTTTTTGCATTTATAATGGATATCTGCAAGTGAG
140 Y L S Q Y A D F P K G W V S H P C F I T
421 ATACCTGAGCCAGTATGCAGACTTCCCTAAGGGTTGGGTTTACATCCCTGCTTCATCAC
160 G S C L W F V G W V I N L H S D H I L R
481 AGGCTCCTGTCTATGGTTTGTGGATGGGTGATTAACCTTCATTCTGACCACATTCTGAG
180 N L R N P G E T G Y K I P R G G V F E Y
541 GAACCTGCGCAATCCAGGAGAAACAGGCTACAAAATACCAAGAGGTGGAGTATTCGAGTA
200 V S G A N Y L G E I V E W T G F A L A A
601 TGTGTCTGGTGCAAATTATCTGGGTGAGATAGTGAATGGACTGGATTTGCTCTGGCTGC
220 H S I H S V A F A F F T F V V L S S R A
661 TCACTCCATCCACAGTGTTCCTTTGCTTTCTTACCTTCGTGCTCCTTTCCAGTTCGAGC
240 I S H H K W Y F E K F E D Y P K S R R A
721 CATCTCCCACCACAAGTGGTATTTTGAGAAATTCGAGGATTATCCGAAATCAAGGAGAGC
260 L I P F V F *
781 TCTCATTCCGTTTGTGTTCTGATGAAGTAACTCAGTATGCGTTTCTCATACAATGAA

```

**Figure 1.** Nucleotide and the deduced amino acid sequence of full-length cDNA of *CmSRD5A1*. The total number of nucleotides 836 bp. Arrows indicate 5' UTR and 3' UTR and boxes indicate the start and stop codons.



provided tool. To generate colored images of the CmSRD5A1 surface, UCSF-Chimera was employed with three color combinations:  $-10 \text{ kcal}/(\text{mol}^{-\circ})$  represented in red,  $0 \text{ kcal}/(\text{mol}^{-\circ})$  in white, and  $+10 \text{ kcal}/(\text{mol}^{-\circ})$  in blue.

### CmSRD5A1 mRNA expression analysis

The basal mRNA expression analysis of SRD5A1 in the selected tissues of *C. magur* was conducted using the LightCycler 450 Real-time PCR detection system (Roche, USA). Each reaction consisted of a 10  $\mu\text{L}$  reaction mix, which included 5  $\mu\text{L}$  of  $2\times$  Maxima SYBR Green qPCR master mix (Thermo Scientific, USA), 0.5  $\mu\text{L}$  of each gene-specific primer (0.3 pM) (Table 1), 2  $\mu\text{L}$  of cDNA (20 ng), and 2  $\mu\text{L}$  of nuclease-free water. The PCR amplification utilized the default thermal profile, commencing with an initial denaturation step at 95  $^{\circ}\text{C}$  for 10 min, followed by 40 cycles of denaturation at 95  $^{\circ}\text{C}$  for 20 s, annealing at 65  $^{\circ}\text{C}$  for 20 s, and extension at 72  $^{\circ}\text{C}$  for 30 s. To ensure accuracy, all reactions were run in triplicate and the experiment was repeated twice. The relative expression of the target gene was determined using the comparative Ct method ( $2^{-\Delta\Delta\text{Ct}}$ )<sup>39</sup> with  $\beta$ -actin as the housekeeping gene for gene normalization<sup>19</sup>.

### Ethics statement

The guidelines of the CPCSEA (Committee for the Purpose of Control and Supervision of Experiments on Animals), Ministry of Environment and Forests (Animal Welfare Division), Government of India on care and use of animals in scientific research were followed to care and rear the animals used in the present study. The study was approved by the Board of studies and authorities of ICAR-Central Institute of Fisheries Education (ICAR- CIFE), Mumbai-61.

### Statistical analysis

The one-way analysis of variance (ANOVA) of statistical package SPSS 17.0 (USA) was used to test the statistical significance of differences in mRNA transcript levels.  $P < 0.05$  was considered statistically significant. The results were expressed as mean  $\pm$  standard error (bars).

### ARRIVE guidelines

The present study has been carried out in accordance with ARRIVE guidelines.

## Results

### Molecular cloning of CmSRD5A1 gene

A 382 bp partial cDNA fragment of CmSRD5A1 was successfully cloned from the brain of *C. magur*. Employing the 5'/3' RACE-PCR technique, a full-length cDNA fragment of 836 bp including UTRs (ORF 795 bp) was obtained for CmSRD5A1 (Fig. 2), which encoded a putative protein consisting of 265 amino acids. Figure 3 presents the ORF of the CmSRD5A1 protein, along with its signature domains. The complete CDS sequence derived from this study has been submitted to NCBI GenBank (Accession number: LC383722 {Protein Id: BBE00816}). Comparative analysis of the amino acid sequence of CmSRD5A1 obtained in this study with similar proteins documented in NCBI indicated a substantial level of similarity (Fig. 4). The amino acid sequence exhibited the highest similarity with *Ictalurus punctatus* (83%), followed by *Pygocentrus nattereri* and *Cyprinus carpio* (75%).

### In silico analysis of CmSRD5A1 amino acid sequences

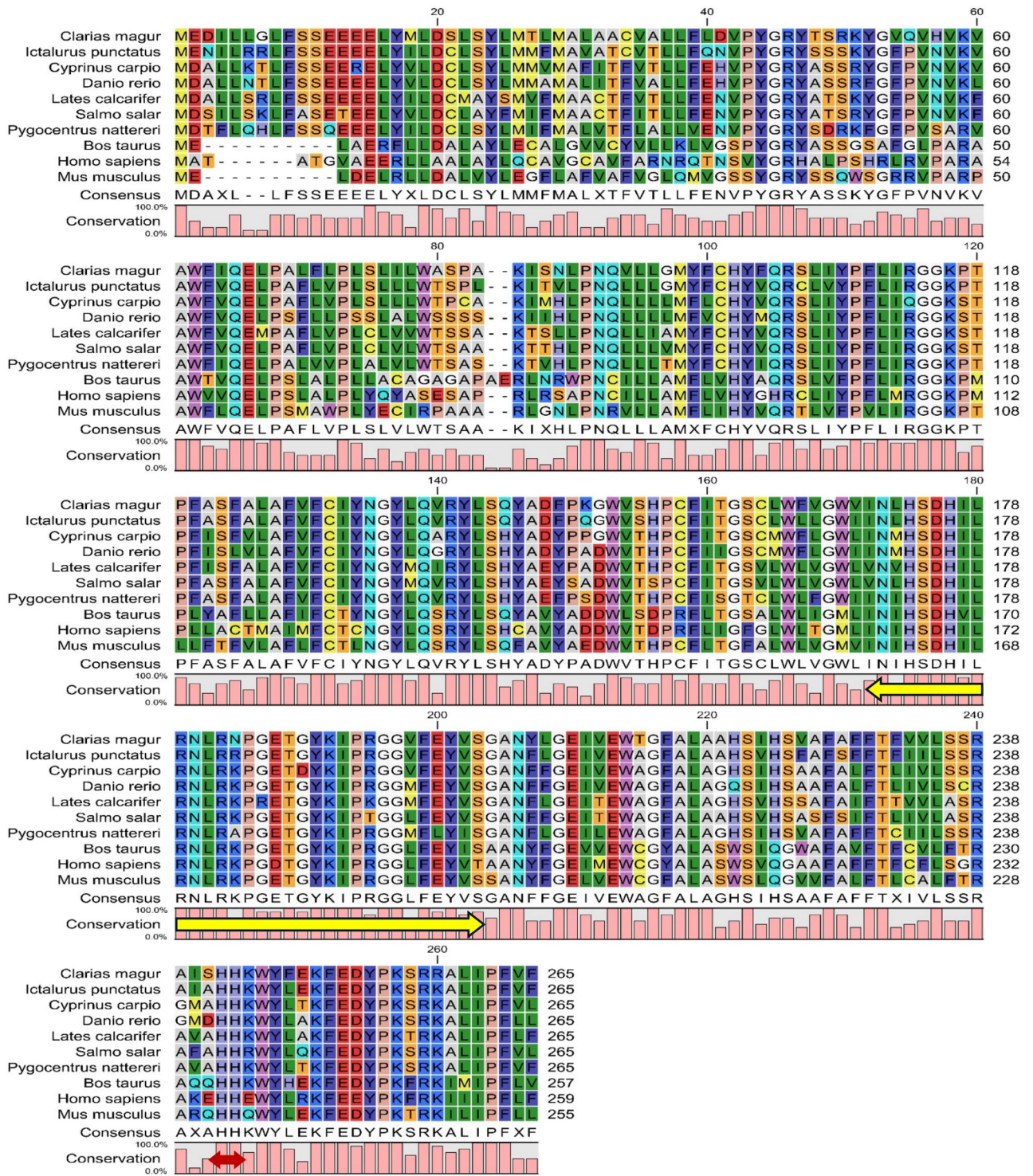
The multiple sequence alignment of the deduced amino acid sequence of CmSRD5A1 with sequences from other species revealed well-conserved regions across the sequences (Fig. 3). Furthermore, the CDD search analysis identified two transmembrane regions and a highly conserved steroid\_dh domain spanning amino acids 116–265 in the CmSRD5A1 protein, exhibiting a remarkable E-value of  $2.06e^{-63}$  across all species included in the multiple alignment analysis (Fig. 5). The Neighbour-joining tree constructed based on the SRD5A1 protein distinguished closely related species into separate clades. The catfish species formed one distinct cluster, while the cyprinids formed another separate group (Fig. 6). Bovine species clustered together, representing a distinct group, while mammals and humans formed a sister group. Notably, the CmSRD5A1 protein displayed close relatedness to the channel catfish and Piranha, followed by the cyprinids. This finding aligns with the classification and evolutionary status of the species.

### Secondary structure analysis of CmSRD5A1 protein

The CmSRD5A1 protein was shown to have an apparent molecular weight (MW) of 30.428 Kilodaltons with a pI of 8.21. The secondary structure of CmSRD5A1 predicted by the PSPRED server showed 10  $\alpha$ -helices, 11 coils and no  $\beta$ -sheets in the protein. No signal peptide could be identified in CmSRD5A1 by SignalP 4.1 server. The hydropathy analysis of the CmSRD5A1 peptide chains revealed the hydrophobic nature of the protein. CmSRD5A1 protein exhibited 16 serine (Ser), 3 threonines (Thr), and 5 tyrosine (Tyr) phosphorylation sites that are uniformly distributed throughout the polypeptide chain as predicted by NetPhos 3.1 Server.

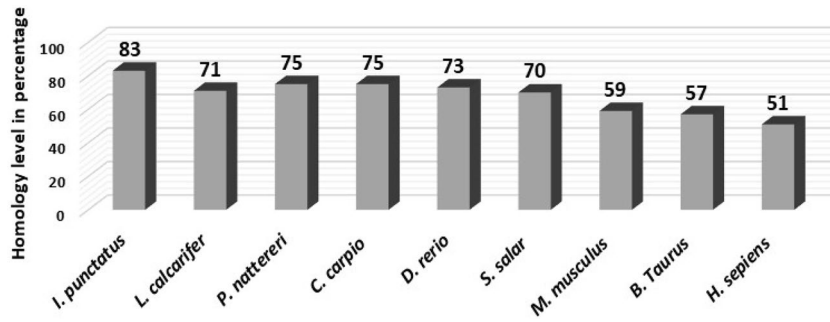
### 3D structure prediction, superimposition, quality evaluation and electrostatic potential

The 3D structure of the CmSRD5A1 sequence generated above was predicted using highly accurate AlphaFold 2 in Google Colab online platform. There were five structures generated. The rank\_1 among all five was selected for further analysis. CmSRD5A1 contains seven transmembrane segments (TMs), with the N-termini located on the periplasmic side of the plasma membrane and the C-termini on the cytosolic side (Fig. 7A,B). The 7-TM helices are connected by six loops (L1–6). The surface representation of CmSRD5A1 revealed a large cavity inside the 7-TM domain at the cytosolic side formed by all 7-TMs and L1, L3, and L5. The predicted structure

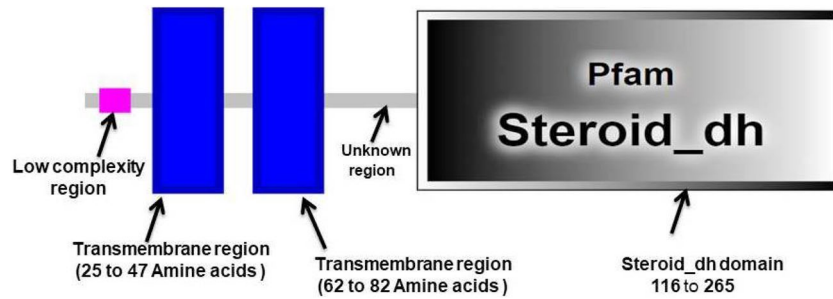


**Figure 2.** Conserved amino acid sequences of the SRD5A1 protein across different vertebrates. Amino acid sequences encoding the SRD5A1 protein were aligned using the Clustal W algorithm of the CLC Genomics Workbench. The bars below indicate the conserved motifs in different vertebrates. Brown arrow represents the histidine residues. Yellow arrow represents the highly conserved NADPH-binding domain. The amino acid sequences of *I. punctatus* (XP\_017336166.1), *C. carpio* (XP\_042632703.1), *D. rerio* (XP\_009292518.1), *L. calcarifer* (XP\_018556926.1), *S. salar* (XP\_014033435.1), *P. nattereri* (XP\_017566885.1), *B. Taurus* (NP\_001092607.1), *H. sapiens* (NP\_001038.1) and *M. musculus* (NP\_780492.2) were retrieved from NCBI database.

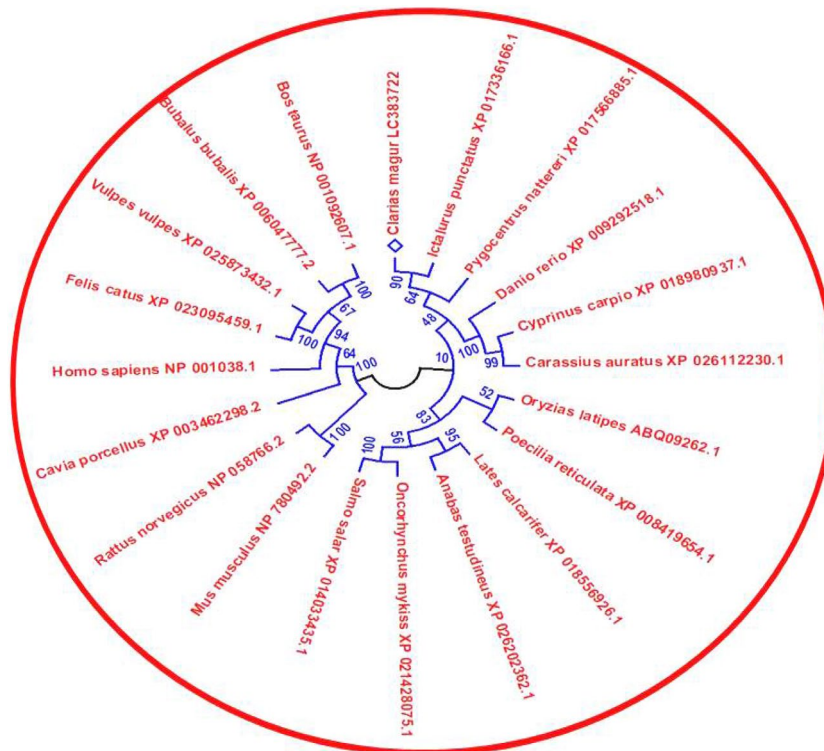




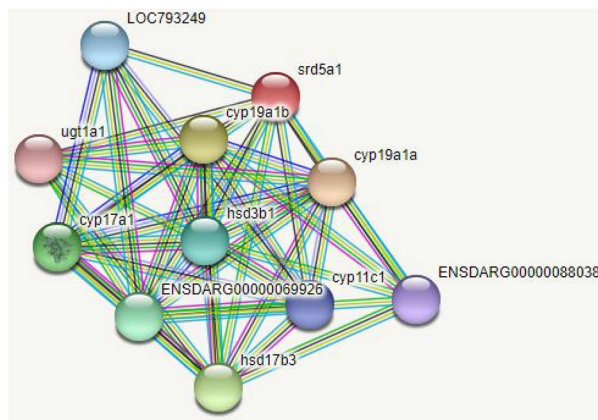
**Figure 3.** Homology level of amino acid sequences of different species.



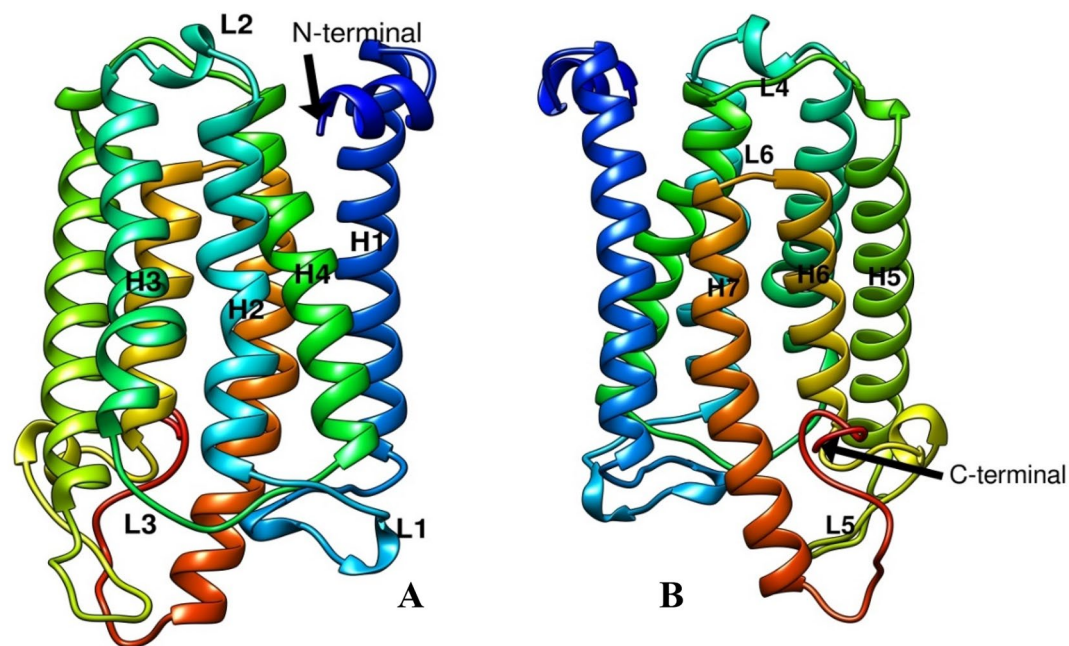
**Figure 4.** Conserved domains of CmSRD5A1 protein.



**Figure 5.** Phylogenetic analysis of CmSRD5A1.



**Figure 6.** Protein–Protein interaction of CmSRD5A1 with different proteins.



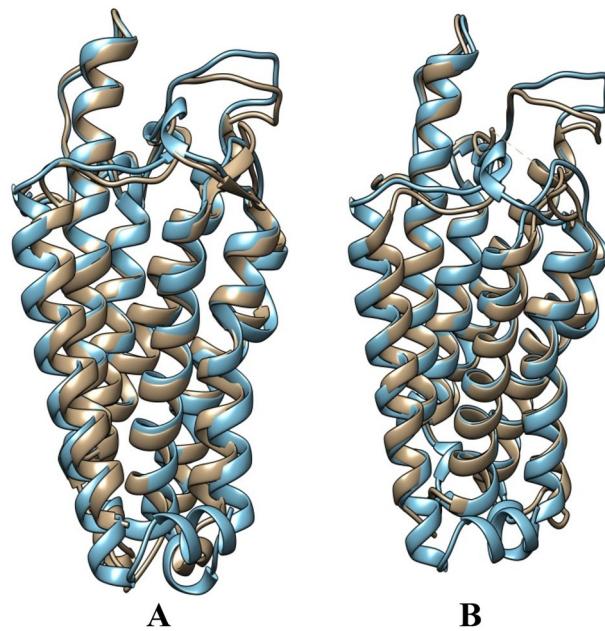
**Figure 7.** Ribbon diagram showing the tertiary structure of CmSRD5A1 protein in which alpha helices and loops are labelled as H1–H7 and L1–6 respectively and are shown in different colours. A. Shows one face of the structure with N-terminal and H1–H4 and L1–L3. B. Structure A was rotated 180° to show remaining helices H5–H7 and loops L4–L6.

was superimposed to SRD5A of Proteobacteria bacterium (PDB ID. 7c83; Fig. 8A) and human SRD5A2 (PDB ID. 7bw1; Fig. 8B) to understand structural similarity. The observed RMSD for the human structure was 1.363 Å for 245 aligned residues and for the bacterial structure was 1.506 Å for 251 aligned residues. This shows a high degree of structural conservation of fish enzymes.

The quality of the predicted structure was analysed by examining using the Ramachandran plot, pLDDT and PAE scores. Ramachandran plot analysis revealed that 94.8% of the total residues are in the most favoured region and none of them in the disallowed region (Fig. 9). IDDT and PAE plots (Fig. 10A,B) generated by AlphaFold revealed a very good quality structure. The N-terminal region around the residues 40–50 was of slightly lower quality with IDDT score of 60. Other regions are of high quality as can be seen in the IDDT and PAE plots.

In the predicted structure it was observed that there is a cavity towards the cytoplasmic side of the enzyme. The cavity has two openings one towards the cytosolic side and the other towards the transmembrane side between the TM1 and TM4. The substrate-binding cavity shows two relatively separate tunnel-like pockets for two ligands. It is reported that this enzyme binds to NADPH and steroid hormones. The ligands NADPH and NADP-dihydrofinasteride (Fig. 11A) were transferred after superimposition from bacterial and human enzymes to understand the ligand interaction. It was observed that NADPH insert into the binding pocket inside the





**Figure 8.** Structural superimposition of CmSRD5A1 with bacterial and human orthologs. **(A)** CmSRD5A1 is superimposed to SRD5A of *Proteobacteria bacterium* (PDB ID: 7c83). **(B)** CmSRD5A1 is superimposed to human SRD5A2 (PDB ID: 7bw1) to understand structural similarity. CmSRD5A1 is coloured in cyan, and the known structures are coloured in gold.

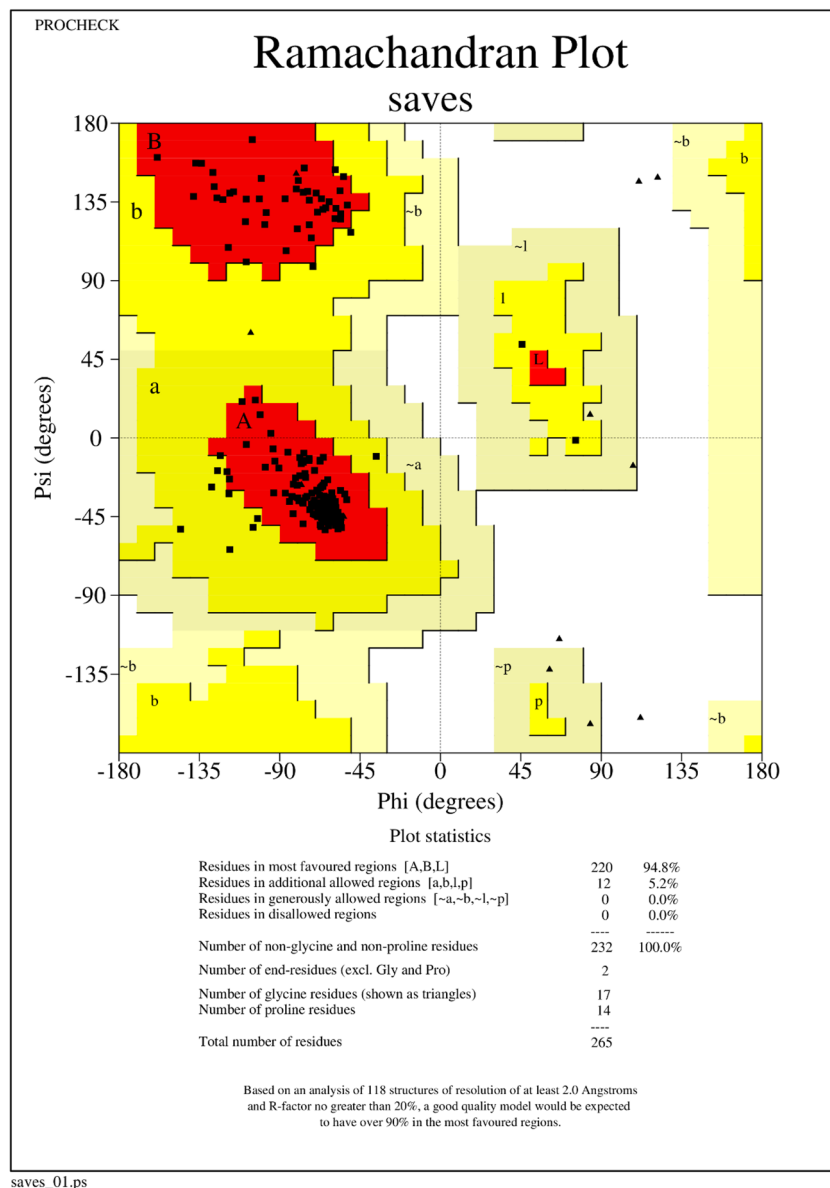
7-TM bundle. The binding pocket for NADPH is enclosed on the cytosolic side by the cytosolic loops, almost completely shielding it from the cytosol. The nicotinamide-ribose moiety is buried inside 7-TM away from the opening. The interaction of steroid moiety was also revealed by comparing it with the human structure having the complex NADP-dihydrofinasteride (Fig. 11A). The binding mode of steroid is in such a way that its  $\Delta^4$  double-bond comes in close proximity to the nicotinamide ring. The electrostatic potential analysis revealed that the cytosolic side of the tunnel is highly electropositive but the transmembrane side opening is neutral (Fig. 11B). NADPH is enclosed towards the cytosolic side in highly electropositive side of the tunnel. The steroid hormone binds towards the transmembrane opening in a relatively hydrophobic environment.

#### Ligand interaction and catalytic mechanism

Based on the structure based sequence alignments the amino acid residues involved in ligand interaction were revealed. The conserved residues forming hydrogen bonds with NADPH were identified as N171, D175, and R182 on TM5, N204, Y205, and E211 on TM6, and T231, R238, and H242 on TM7 (Fig. 12A,B). The residues on intracellular loop 1 (Y44 and R46) and loop 3 (Y189) are also conserved and can be contributed to NADPH binding through direct hydrogen bonds. Conserved residues partly interacted with NADPH through water mediated hydrogen bonds are R46 on intracellular loop 1, N204 on TM6 and R238 on TM7. Polar residues binding to the ketone group in the steroids were identified as E66 and Y101 (Fig. 12C,D). The hydrophobic residues binding to steroids are not so conserved. The region of the substrate binding pocket involved in binding with steroids contains a conserved signature motif identified as E66 and Y101 important for catalysis. This motif forms a triangular hydrogen bond that coordinates the ketone group at C-3 in the steroids. This coordination brings the steroid into proximity to the nicotinamide of NADPH, leading to the hydride ( $H^-$ ) transfer and  $\Delta^4$  double-bond reduction. Additionally, Q65 was also identified which can bind and stabilize E66 as reported in human and bacterial enzymes.

#### Expression analysis of CmSRD5A1 gene

The mRNA expression pattern of SRD5A1 gene in *C. magur* adult male and female fishes was determined through qRT-PCR in the brain and gonads at four different maturation stages namely preparatory, spawning (before injection), 6 h post-Ovatide injection and 16 h post-Ovatide injection (post-spawning). In female fish, the CmSRD5A1 gene was expressed in both the tissues of all stages but significantly higher in the ovary compared to the brain. The expression of CmSRD5A1 in the ovary was decreased at spawning stage compared to the preparatory stage. The highest mRNA expression of CmSRD5A1 in the ovary was found after the 6 h of Ovatide induction (Fig. 13). In testis, significant peak mRNA expression of CmSRD5A1 was observed at 16 h post-Ovatide injection stage compared to the preparatory stage (Fig. 10). Whereas in case of the brain the lowest mRNA expression of CmSRD5A1 in female was obtained at spawning stage which was at peak in preparatory stage (Fig. 11). In the brain of male fish, the mRNA expression of CmSRD5A1 was suddenly dropped at 16 h-PI stage with significantly threefold change (Fig. 14).

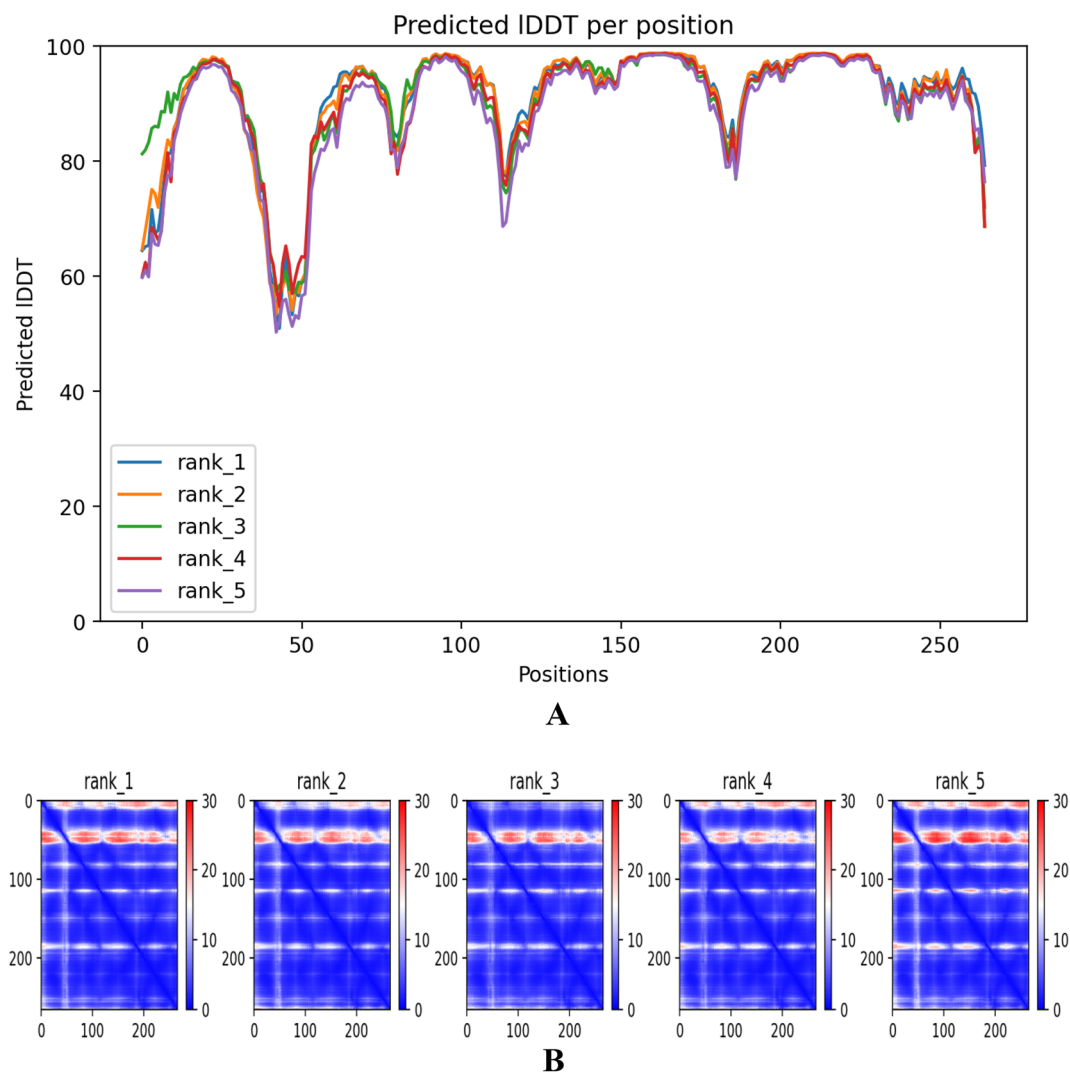


**Figure 9.** Ramachandran plot CmsRD5A1 showing the dihedral angles Psi and Phi of amino acid residues. The residues which lie in the most favoured regions are shown as black dots in red area and the residues which lie in additional allowed regions are in red dots.

## Discussion

The full-length cDNA sequence of the SRD5A1 gene (795 bp ORF) was cloned from *C. magur* using the 5'/3' RACE-PCR strategy. Previous studies have demonstrated the presence of 5 $\alpha$ -reductase activity in select fish species such as goldfish, toadfish, lamprey, rainbow trout, and zebrafish<sup>16,40</sup>. Comparative analysis of the CmsRD5A1 amino acid sequence with sequences from other fish and mammal species revealed significant regions of similarity, including a putative NADPH-binding domain. Within this domain, a highly conserved region consisting of 31 amino acids (residues 170–200) was tentatively identified (Fig. 3). Further sequence analysis indicated that CmsRD5A1 contains two conserved histidine residues (H242, H243) in its C-terminal region, which are known to be essential for substrate binding and catalytic activity<sup>41</sup> (Fig. 3). In a previous report by<sup>42</sup>, a tetrapeptide segment involved in substrate binding was identified in the chimeric SRD5A1 enzyme constructed from human and rat. In our study, a similar tetrapeptide segment (-A-A-C-V-) for steroid binding was also found in CmsRD5A1 (Fig. 3). However, the sensitivity of this segment to inhibitors of steroid 5 $\alpha$ -reductase remains unknown. These findings are consistent with previous reports, confirming the high conservation of these motifs<sup>11,43</sup>.

The phylogenetic analysis of CmsRD5A1, with other species, revealed high similarity with *I. punctatus* followed by other fishes like *L. carcerifer*, *P. nattereri* and *D. rerio*. The *C. magur* SRD5A1 showed 51% similarity with its ortholog from the human.



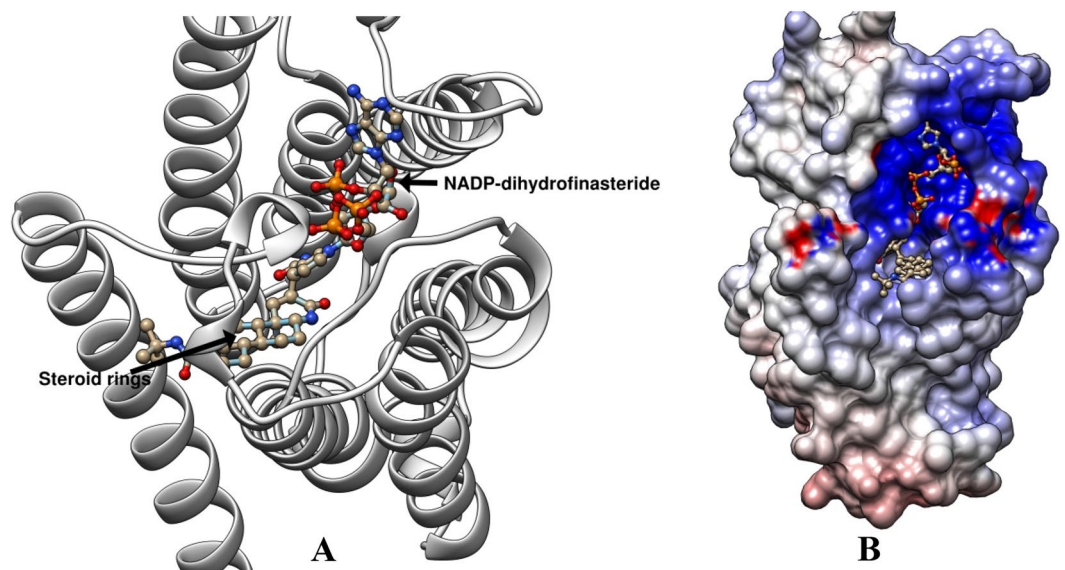
**Figure 10.** Structural quality parameters given by AlphaFold 2. **(A)** Predicted local distance difference test (pLDDT) plot for five generated structures. **(B)** Predicted aligned error (PAE) plot for predicted CmSRD5A1 structures. Rank\_1 structure was used for further analysis.

The hydrophobicity analyzed using the Kyte-Doolittle algorithm predicted that CmSRD5A1 protein is more hydrophobic at N-terminus region due to the presence of a highly hydrophobic region of nonpolar amino acids and is believed to function as a transmembrane signal anchor. Similar kind of hydrophobicity profile has been observed in its ortholog proteins<sup>44</sup> supporting that CmSRD5A1 protein is hydrophobic. Since the SRD5A1 is a membrane-associated protein with a hydrophobic transmembrane domain, no signal peptide could be identified.

SRD5As belong to a large group of eukaryotic membrane embedded steroid reductases catalyzing the irreversible reduction of the  $\Delta 4,5$  bond in  $\Delta 4-3$ -ketosteroids using reduced NADPH as the hydride donor cofactor<sup>45,46</sup>. The membrane-embedded SRD5A family in humans includes five members, SRD5A1–3 and glycoprotein synaptic 2 (GSPN2) and GSPN2<sup>46</sup>. There is another group of steroid reductase from the Aldo-Keto Reductase (AKR) superfamily which is soluble and involved in numerous metabolic transformations<sup>47</sup>. There are numerous structures available for soluble steroid reductases but the 3D structure and mechanism of action of eukaryotic membrane-embedded steroid reductases are less explored. To date, there are two structures of SRD5A enzymes namely PbSRD5A from Proteobacteria bacterium<sup>45</sup> and human SRD5A2<sup>46</sup> are reported. Although these steroid/sterol reductases share very little sequence similarity, they all use NADPH as the cofactor to reduce specific carbon-carbon double bonds in their steroid substrates. It is not surprising that SRD5As exhibit similar fold and the NADPH cofactor-binding sites are highly conserved in the steroid reductase families. Biochemical and structural studies carried out by<sup>45,46</sup> on human and bacterial enzymes have led to the understanding of the structural and catalytic mechanism of this enzyme.

Both the reported structures have very high degree of structural similarity. They adopt a monomeric 7-TM structural topology with a largely enclosed binding cavity inside the transmembrane domain. Structure predicted in the current study from the catfish *Clarias magur* which is another vertebrate-like mammal has also shown the same structural topology. Structural superimposition of current structure with above mentioned human and



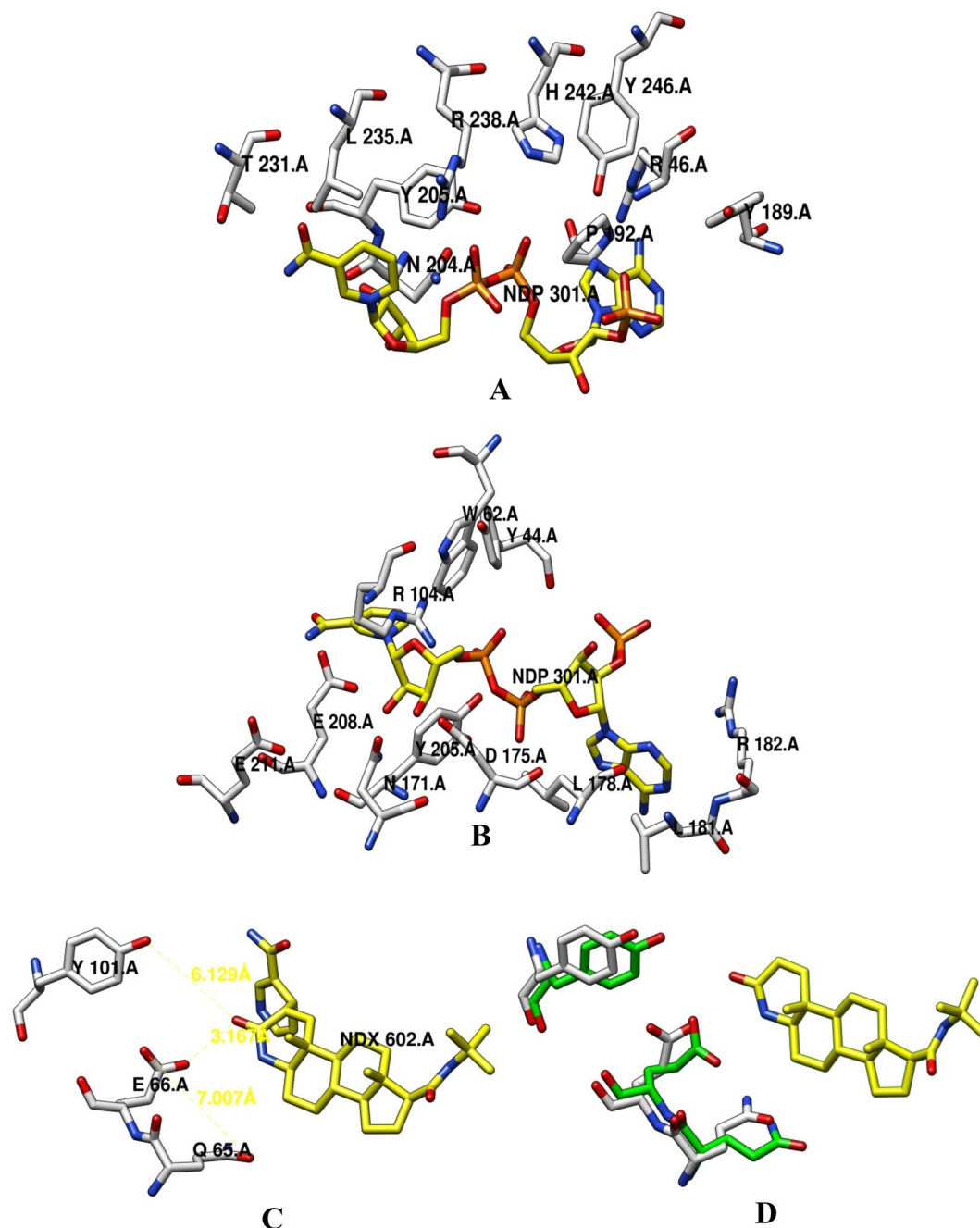


**Figure 11.** CmSRD5A1 structure having inserted with NADP-dihydrofinasteride based on the structural superimposition with human SRD5A2 (PDB ID: 7bw1). The steroid rings group is also marked.

bacterial SRD5A given an RMSD value  $< 2.0 \text{ \AA}$  for more than 80% of the residues aligned showing the very high degree of structural conservation. Analysis of surface features revealed a tunnel towards the cytoplasmic side having two openings one towards the lipid bilayer and the other towards the cytoplasm region. The cytoplasmic loop partially covers the cytoplasmic opening. The electrostatic potential calculation revealed that the cytoplasmic region of the cavity is highly electropositive but the membrane side opening is hydrophobic. This type of asymmetry is noted in both the bacterial and human SRD5A structures. The electropositive side of the tunnel is involved in the binding of NADPH and the hydrophobic part binds with steroids. The ligands were transferred to CmSRD5A1 from known SRD5A structures.

Ligand binding and catalytic residues were identified based on structural superimposition based sequence alignment in the predicted CmSRD5A1. The amino acid comparison with bacterial enzyme revealed almost conserved residues for NADPH binding. On the other hand, residues involved in steroid binding are not well conserved, but the locality was hydrophobic. Structural, computational and mutagenesis studies reveal the molecular mechanisms of catalysis and finasteride inhibition involving residues E57 and Y91 in human SRD5A<sup>46</sup>. The corresponding catalytic identified as E66 and Y101 are important for catalysis. This motif forms a network of hydrogen bonds that coordinate the ketone group at C-3 in the steroids. This coordination brings the steroid into proximity to the nicotinamide of NADPH, leading to the hydride ( $\text{H}^-$ ) transfer and  $\Delta^4$  double-bond reduction.

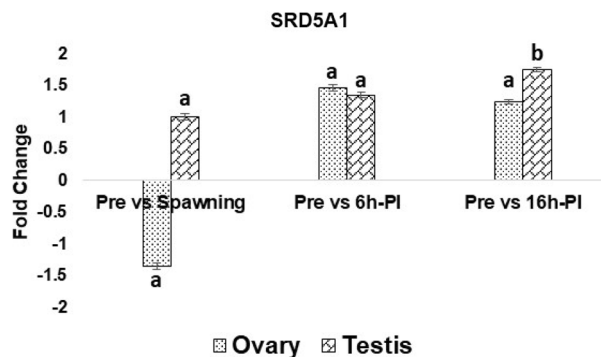
Two maturation stages and two after 6 h and 16 h Ovotide injected stages of *C. magur* adult male and female fishes were taken for CmSRD5A1 mRNA expression study by qRT-PCR. The differential mRNA expression obtained for the ovary was not statistically significant. In the ovary, the CmSRD5A1 gene showed lower mRNA expression at the spawning stage compared to the preparatory stage while expression peaked at 6 h post-Ovotide injection stage. This mRNA expression profile in the ovary suggested that more CmSRD5A1 at preparatory stage leads to the conversion of more testosterone into DHT. It has been reported that DHT stimulates the secondary sexual characteristics and masculinization of the female genotype to a higher than testosterone<sup>48</sup>. However, the actual physiological role of testosterone and DHT have not yet been fully described in female fish but considered that the higher production of CmSRD5A1 might alter the reproductive hormonal secretion in the female through the accumulation of DHT. On the other hand, DHT can increase the  $17\beta$ -estradiol output from the teleost ovary<sup>40</sup> which is the principal estrogen secreted in all vertebrates including fish and is essential for the development and maintenance of female reproductive tissues like the ovary<sup>49,50</sup>. Therefore the lower expression of CmSRD5A1 at the mature stage in the ovary as well as in the brain of female fish resulted in less accumulation of DHT which might reduce the production of  $17\beta$ -estradiol. In agreement with the earlier studies<sup>51,52</sup> the mRNA expression of CmSRD5A1 in the male brain was observed and found higher than in the testis. In the case of the male brain, the mRNA expression of CmSRD5A1 was found to be down-regulated significantly ( $P < 0.05$ ) after 16 h of Ovotide injection. The lower expression of CmSRD5A1 in the male brain at 16 h post-Ovotide injection stage resulted in less availability of DHT which might be the possible reason for the dysfunction of the testis in captivity. As reviewed<sup>15</sup>, different several androgen receptors have been characterized in fishes, but none is specific for KT indicating that some other androgens like DHT might be involved in the mediation of the androgenic response in different target tissues. Few studies have reported an androgen receptor with high binding affinity to DHT in many teleost species and suggested DHT is often one of the most potent inducers of androgen receptor transcriptional activity in vitro<sup>53,54</sup>. If DHT is the more potent androgen and preferred ligand for androgen receptor (AR) transactivation than T and KT, the overexpression of DHT through the SRD5A1 activation might help in spermatogenesis.



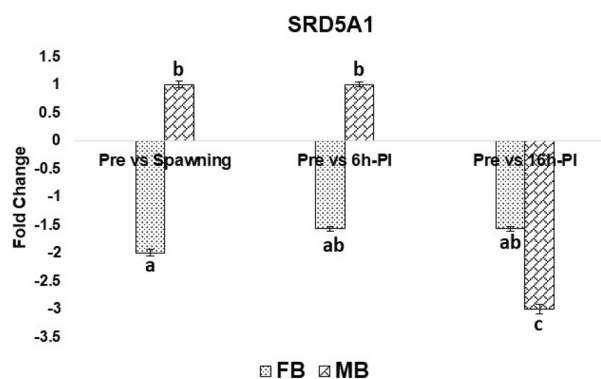
**Figure 12.** Ligand interaction and catalytic residues based on structural superimposition. (A,B) Interaction of NADP with amino acid residues predicted based on superimposition and the structural based sequence alignment with *Proteobacteria bacterium* SRD5A (PDB ID. 7c83). (C) Catalytic residues involved in H-bonding with ketone group at C-3 of steroid. (D) Catalytic residues superimposed with human SRD5A2 ((PDB ID. 7bw1).

## Conclusion

In summary, the present study reports the full-length coding sequence of the SRD5A1 gene from *C. magur*. The deduced amino acid sequence was analysed using various bioinformatics tools to predict the structural and functional features. The conserved catalytic residues involved in  $\Delta 4$  double-bond reduction are identified as E66 and Y101. The differential mRNA expression analysis of CmSRD5A1 revealed the highest expression in the ovary at 6 h post-Ovatide injection stage and in testis at 16 h post-Ovatide injection stage compared to the preparatory stage. In both sex gonads, the lowest expression of CmSRD5A1 was observed at the spawning stage. Whereas in the case of the brain, the lowest mRNA expression of CmSRD5A1 was obtained at spawning stage in females and males at 16 h post-Ovatide injection stage. As per our knowledge, this is the first report of cloning and structural analysis of CmSRD5A1 full-length cDNA in Indian catfish. In general, this study is a basis for



**Figure 13.** qRT-PCR analysis of SRD5A1 mRNA expression in ovary and testis at different maturation stages of *C. magur*. The relative expression of SRD5A1 normalized with  $\beta$ -actin was calculated using the comparative Ct method. Data for real-time PCR were expressed as mean  $\pm$  S.E.M (n = 3). The superscripts (a,b,c,d) above the bar represents the statistical significance at  $P < 0.05$ .



**Figure 14.** qRT-PCR analysis of SRD5A1 mRNA expression in brain of both the sexes of *C. magur* at different maturation stages. The relative expression of SRD5A1 normalized with  $\beta$ -actin was calculated using the comparative Ct method. Data for real-time PCR were expressed as mean  $\pm$  S.E.M (n = 3). The superscripts (a,b,c,d) above the bar represents the statistical significance at  $P < 0.05$ .

further exploration of structural and functional characteristics of the SRD5A1 gene in *C. magur*, including its protein structure and involvement in signal transduction. The expression pattern of the transcript in gonads and during different reproductive phases provides insights into their involvement in reproductive cycles. Its actual role during the reproductive phases needs to be explored in further research.

### Data availability

Data will be made on special request to the corresponding author. However, the complete CDS sequence derived from this study has been submitted to NCBI GenBank (Accession number: LC383722 {Protein Id: BBE00816}) which is available there.

Received: 25 July 2023; Accepted: 7 November 2023

Published online: 09 November 2023

### References

- Chiang, E. F. L., Yan, Y. L., Guiguen, Y., Postlethwait, J. & Chung, B. C. Two Cyp19 (P450 aromatase) genes on duplicated zebrafish chromosomes are expressed in ovary or brain. *Mol. Biol. Evol.* **18**(4), 542–550. <https://doi.org/10.1093/oxfordjournals.molbev.a003833> (2001).
- Dalla Valle, L., Lunardi, L., Colombo, L. & Belvedere, P. European sea bass (*Dicentrarchus labrax* L.) cytochrome P450arom: cDNA cloning, expression and genomic organization. *J. Steroid Biochem. Mol. Biol.* **80**(1), 25–34. [https://doi.org/10.1016/S0960-0760\(01\)00170-4](https://doi.org/10.1016/S0960-0760(01)00170-4) (2002).
- Greytak, S. R., Champlin, D. & Callard, G. V. Isolation and characterization of two cytochrome P450 aromatase forms in killifish (*Fundulus heteroclitus*): Differential expression in fish from polluted and unpolluted environments. *Aquat. Toxicol.* **71**(4), 371–389. <https://doi.org/10.1016/j.aquatox.2004.12.007> (2005).
- Kwon, J. Y., McAndrew, B. J. & Penman, D. J. Cloning of brain aromatase gene and expression of brain and ovarian aromatase genes during sexual differentiation in genetic male and female Nile tilapia *Oreochromis niloticus*. *Mol. Reprod. Dev.* **59**(4), 359–370. <https://doi.org/10.1002/mrd.1042> (2001).



5. Tanaka, M. *et al.* Cloning and sequence analysis of the cDNA encoding P-450 aromatase (P450arom) from a rainbow trout (*Oncorhynchus mykiss*) ovary; relationship between the amount of P450arom mRNA and the production of oestradiol-17 $\beta$  in the ovary. *J. Mol. Endocrinol.* **8**(1), 53–61. <https://doi.org/10.1677/jme.0.0080053> (1992).
6. Do Rego, J. L. *et al.* Neurosteroid biosynthesis: enzymatic pathways and neuroendocrine regulation by neurotransmitters and neuropeptides. *Front. Neuroendocrinol.* **30**(3), 259–301. <https://doi.org/10.1016/j.yfrne.2009.05.006> (2009).
7. Russell, D. W. & Wilson, J. D. Steroid 5 $\alpha$ -reductase: Two genes/two enzymes. *Annu. Rev. Biochem.* **63**(1), 25–61. <https://doi.org/10.1146/annurev.bi.63.070194.000325> (1994).
8. Langlois, V. S., Zhang, D., Cooke, G. M. & Trudeau, V. L. Evolution of steroid-5 $\alpha$ -reductases and comparison of their function with 5 $\beta$ -reductase. *Gen. Comp. Endocrinol.* **166**(3), 489–497. <https://doi.org/10.1016/j.ygcen.2009.08.004> (2010).
9. Andersson, S., Berman, D. M., Jenkins, E. P. & Russell, D. W. Deletion of steroid 5 $\alpha$ -reductase 2 gene in male pseudo hermaphroditism. *Nature* **354**(6349), 159–161. <https://doi.org/10.1038/354159a0> (1991).
10. Berman, D. M. & Russell, D. W. Cell-type-specific expression of rat steroid 5 $\alpha$ -reductase isozymes. *Proc. Natl. Acad. Sci.* **90**(20), 9359–9363. <https://doi.org/10.1073/pnas.90.20.9359> (1993).
11. Levy, M. A. *et al.* Cloning, expression and functional characterization of type 1 and type 2 steroid 5 $\alpha$ -reductases from cynomolgus monkey: Comparisons with human and rat isoenzymes. *J. Steroid Biochem. Mol. Biol.* **52**(4), 307–319. [https://doi.org/10.1016/0960-0760\(94\)00183-m](https://doi.org/10.1016/0960-0760(94)00183-m) (1995).
12. Wilson, J. D. The endocrine control of sexual differentiation. *Harvey lectures* **79**, 145–172 (1983).
13. Margiotta-Casaluci, L., Hannah, R. E. & Sumpter, J. P. Mode of action of human pharmaceuticals in fish: The effects of the 5 $\alpha$ -reductase inhibitor, dutasteride, on reproduction as a case study. *Aquat. Toxicol.* **128**, 113–123. <https://doi.org/10.1016/j.aquatox.2012.12.003> (2013).
14. Mathieu, M. *et al.* Immunohistochemical localization of 3 $\beta$ -hydroxysteroid dehydrogenase and 5 $\alpha$ -reductase in the brain of the African lungfish *Protopterus annectens*. *J. Comp. Neurol.* **438**(2), 123–135. <https://doi.org/10.1002/cne.1304> (2001).
15. Margiotta-Casaluci, L. & Sumpter, J. P. 5 $\alpha$ -Dihydrotestosterone is a potent androgen in the fathead minnow (*Pimephales promelas*). *Gen. Comp. Endocrinol.* **171**(3), 309–318. <https://doi.org/10.1016/j.ygcen.2011.02.012> (2011).
16. Diotel, L. *et al.* The brain of teleost fish, a source, and a target of sexual steroids. *Front. Neurosci.* **5**, 137. <https://doi.org/10.3389/fnins.2011.00137> (2011).
17. Joy, K. P., Singh, M. S., Senthilkumaran, B. & Goos, H. T. Pituitary-gonadal relationship in the catfish *Clarias batrachus* (L): A study correlating gonadotrophin-II and sex steroid dynamics. *Zool. Sci.* **17**(3), 395–404. <https://doi.org/10.2108/jzs.17.395> (2000).
18. Agarwal, D. *et al.* Transcriptome analysis of *Clarias magur* brain and gonads suggests neuro-endocrine inhibition of milt release from captive GnRH-induced males. *Genomics* **112**(6), 4041–4052. <https://doi.org/10.1016/j.ygeno.2020.07.012> (2020).
19. Agarwal, D. *et al.* Molecular characterization and expression profiling of 17 $\beta$ -hydroxysteroid dehydrogenase 2 and spermatogenesis associated protein 2 genes in endangered catfish, *Clarias magur* (Hamilton, 1822). *Anim. Biotechnol.* **31**(2), 93–106. <https://doi.org/10.1080/10495398.2018.1545663> (2018).
20. Ahmad, I. *et al.* Molecular characterization of Inhibin-A: Structure and expression analysis in *Clarias batrachus*. *Gen. Comp. Endocrinol.* **261**, 104–114. <https://doi.org/10.1016/j.ygcen.2018.02.007> (2018).
21. Chakrabarty *et al.* Endosulfan and flutamide, alone and in combination, target ovarian growth in juvenile catfish, *Clarias batrachus*. *omp. Biochem. Physiol. C: Toxicol. Pharmacol.* **155**(3), 491–497. <https://doi.org/10.1016/j.cbpc.2011.12.007> (2012).
22. Chutia *et al.* Gene Regulation on Spermiation of Catfish in *Recent updates in molecular Endocrinology and Reproductive Physiology of Fish: An Imperative step in Aquaculture* (ed. Sundaray *et al.*) 161–169. [https://doi.org/10.1007/978-981-15-8369-8\\_11](https://doi.org/10.1007/978-981-15-8369-8_11) (2021).
23. Mamta, S. K. *et al.* Cloning and expression analysis of tyrosine hydroxylase and changes in catecholamine levels in brain during ontogeny and after sex steroid analogues exposure in the catfish *Clarias batrachus*. *Gen. Comp. Endocrinol.* **197**, 18–25. <https://doi.org/10.1016/j.ygcen.2013.11.022> (2014).
24. Mir, I. N., Srivastava, P. P., Bhat, I. A., Dar, J. Y. & Agarwal, D. Expression profiling of lipolytic and long-chain fatty acid biosynthesis genes in catfish larvae fed with graded lipid levels. *Front. Mar. Sci.* **791**, 1. <https://doi.org/10.3389/fmars.2020.561402> (2020).
25. Rajakumar, A. & Senthilkumaran, B. Molecular cloning and expression analysis of 17 $\beta$ -hydroxysteroid dehydrogenase 1 and 12 during gonadal development, recrudescence and after in vivo hCG induction in catfish, *Clarias batrachus*. *Steroids* **92**, 81–89. <https://doi.org/10.1016/j.steroids.2014.09.009> (2014).
26. Sharma, K., Yadava, N. K. & Jindal, M. Effect of different doses of ovatein on the breeding performance of *Clarias batrachus* (Linn.). *Livest. Res. Rural. Dev.* **22**(4) (2010).
27. Tamura, K., Stecher, G. & Kumar, S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**(7), 3022–3027. <https://doi.org/10.1093/molbev/msab120> (2021).
28. Gasteiger, E. *et al.* Protein identification and analysis tools on the ExpASY server in *The proteomics protocols handbook*. 571–607. (Humana press). <https://doi.org/10.1385/1-59259-890-0-571> (2005).
29. Geourjon, C. & Deleage, G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Bioinformatics* **11**(6), 681–684. <https://doi.org/10.1093/bioinformatics/11.6.681> (1995).
30. Petersen, T. N., Brunak, S., Von Heijne, G. & Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**(10), 785–786. <https://doi.org/10.1038/nmeth.1701> (2011).
31. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2> (2021).
32. Kryshtafovich, A., Schwede, T., Topf, M., Fidelis, K. & Moutl, J. Critical assessment of methods of protein structure prediction (CASP): Round XIV. *Proteins* **89**, 1607–1617. <https://doi.org/10.1002/prot.26237> (2021).
33. Mirdita, M. *et al.* ColabFold: Making protein folding accessible to all. *Nat. Methods* **19**, 679–682. <https://doi.org/10.1038/s41592-022-01488-1> (2022).
34. Kumar, G., Singh, A. K. & Agarwal, D. Structural and functional characterization of RNA dependent RNA polymerase of *Macrobriachium rosenbergii* nodavirus (MnRdRp). *J. Biomol. Struct. Dyn.* **1**, 1–13. <https://doi.org/10.1080/07391102.2023.2175384> (2023).
35. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**(4), 477–486. <https://doi.org/10.1007/BF00228148> (1996).
36. Tumasuvunakool, K. *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* **596**(7873), 590–596. <https://doi.org/10.1038/s41586-021-03828-1> (2021).
37. Szklarczyk, D. *et al.* STRING v10: Protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**(D1), D447–D452. <https://doi.org/10.1093/nar/gku1003> (2015).
38. Petersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. <https://doi.org/10.1002/jcc.20084> (2004).
39. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2 $^{-\Delta\Delta CT}$  method. *Methods* **25**(4), 402–408. <https://doi.org/10.1006/meth.2001.1262> (2001).
40. Martyniuk, C. J., Bissegger, S. & Langlois, V. S. Reprint of “Current perspectives on the androgen 5 $\alpha$ -dihydrotestosterone (DHT) and 5 $\alpha$ -reductases in teleost fishes and amphibians”. *Gen. Comp. Endocrinol.* **203**, 10–20. <https://doi.org/10.1016/j.ygcen.2014.06.011> (2014).

41. Bhattacharyya, A. K. & Collins, D. C. Site-directed mutagenesis studies of rat steroid 5 $\alpha$ -reductase (isozyme-1): Mutation of residues in the cofactor binding and C-terminal regions. *J. Steroid Biochem. Mol. Biol.* **77**(2–3), 177–182. [https://doi.org/10.1016/S0960-0760\(01\)00023-1](https://doi.org/10.1016/S0960-0760(01)00023-1) (2001).
42. Thigpen, A. E. & Russell, D. W. Four-amino acid segment in steroid 5 alpha-reductase 1 confers sensitivity to finasteride, a competitive inhibitor. *J. Biol. Chem.* **267**(12), 8577–8583. [https://doi.org/10.1016/S0021-9258\(18\)42482-9](https://doi.org/10.1016/S0021-9258(18)42482-9) (1992).
43. Ramos, L., Chávez, B. & Vilchis, F. Cloning and differential expression of steroid 5 $\alpha$ -reductase type 1 (Srd5a1) and type 2 (Srd5a2) from the Harderian glands of hamsters. *Gen. Comp. Endocrinol.* **166**(2), 388–395. <https://doi.org/10.1016/j.ygcen.2009.12.010> (2010).
44. Andersson, S., Bishop, R. W. & Russell, D. W. Expression cloning and regulation of steroid 5 alpha-reductase, an enzyme essential for male sexual differentiation. *J. Biol. Chem.* **264**(27), 16249–16255. [https://doi.org/10.1016/S0021-9258\(18\)71614-1](https://doi.org/10.1016/S0021-9258(18)71614-1) (1989).
45. Han, Y. *et al.* Crystal structure of steroid reductase SRD5A reveals conserved steroid reduction mechanism. *Nat. Commun.* **12**(1), 449. <https://doi.org/10.1038/s41467-020-20675-2> (2021).
46. Xiao, Q. *et al.* Structure of human steroid 5 $\alpha$ -reductase 2 with the anti-androgen drug finasteride. *Nat. Commun.* **11**(1), 5430. <https://doi.org/10.21203/rs.3.rs-40159/v1> (2020).
47. Barski, O. A., Tipparaju, S. M. & Bhatnagar, A. The aldo-keto reductase superfamily and its role in drug metabolism and detoxification. *Drug Metab. Rev.* **40**(4), 553–624. <https://doi.org/10.1080/03602530802431439> (2008).
48. Borg, B. Androgens in teleost fishes. *Comp. Biochem. Physiol. Part C: Toxicol. Pharmacol.* **109**(3), 219–245. [https://doi.org/10.1016/0742-8413\(94\)00063-G](https://doi.org/10.1016/0742-8413(94)00063-G) (1994).
49. Carvalho, C. V. A. D., Passini, G., Costa, W. D. M., Vieira, B. N. & Cerqueira, V. R. Effect of estradiol-17 $\beta$  on the sex ratio, growth and survival of juvenile common snook (*Centropomus undecimalis*). *Acta Scientiarum. Anim. Sci.* **36**(3), 239–245. <https://doi.org/10.4025/actascianimsci.v36i3.22839> (2014).
50. Casey, M. L., MacDonald, P. C. & Andersson, S. 17 beta-Hydroxysteroid dehydrogenase type 2: Chromosomal assignment and progesterone regulation of gene expression in human endometrium. *J. Clin. Investig.* **94**(5), 2135–2141. <https://doi.org/10.1172/JCI117569> (1994).
51. Pasmanik, M. & Callard, G. V. Aromatase and 5 $\alpha$ -reductase in the teleost brain, spinal cord, and pituitary gland. *Gen. Comp. Endocrinol.* **60**(2), 244–251. [https://doi.org/10.1016/0016-6480\(85\)90320-x](https://doi.org/10.1016/0016-6480(85)90320-x) (1985).
52. Pasmanik, M. & Callard, G. V. Changes in brain aromatase and 5 $\alpha$ -reductase activities correlate significantly with seasonal reproductive cycles in goldfish (*Carassius auratus*). *Endocrinology* **122**(4), 1349–1356. <https://doi.org/10.1210/endo-122-4-1349> (1988).
53. Fitzpatrick, M. S., Gale, W. L. & Schreck, C. B. Binding characteristics of an androgen receptor in the ovaries of coho salmon *Oncorhynchus kisutch*. *Gen. Comp. Endocrinol.* **95**(3), 399–408. <https://doi.org/10.1006/gcen.1994.1139> (1994).
54. Olsson, P. E. *et al.* Molecular cloning and characterization of a nuclear androgen receptor activated by 11-ketotestosterone. *Reprod. Biol. Endocrinol.* **3**(1), 1–17. <https://doi.org/10.1186/1477-7827-3-37> (2005).

## Acknowledgements

We appreciate our colleagues for their valuable suggestions and support. The authors acknowledge ICAR-Central Institute of Fisheries Education, Mumbai, and Tamil Nadu Dr. J. Jayalalithaa Fisheries University, OMR Campus, Chennai, India for providing all the facilities to conduct the present research.

## Author contributions

D.A. generated the research idea performed Wet lab and Molecular analysis and wrote the manuscript. G.K. performed the computational analysis, M.A.R. and I.A. formulated work plan and corrected the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to M.A.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023