Journal of Scientific & Industrial Research
Vol. 79, August 2020, pp. 705-708

Computational Modelling of *Tetraodon nigroviridis* Melanocortin-1 Receptor (MC1R) Protein and Identification of Natural Compounds as Putative Modulators

Saumya Pandey and Shrinivas Jahageerdar*

Fish Genetics and Biotechnology Division, ICAR-Central Institute of Fisheries Education, Mumbai– 400 061, India

Received:31 July 2018; revised 05 March 2020; accepted 23 May 2020

In ornamental fisheries, body ornamentation, including the body colour, is of great economic importance. Melanocortin-1 receptor (MC1R) gene is primarily responsible for pigmentation in the majority of the vertebrates, and its modulation will help to understand the pigmentation and external body coloration. The present study was aimed to predict the MC1R tertiary protein structure of *Tetraodon nigroviridis* through homology modelling and to identify putative modulators. The tertiary protein structure of the MC1R protein of *T. nigroviridis* was homology modelled and validated. BLAHtetrone with binding efficiency of -9.7 kcal/mol was identified as a putative modulator of the MC1R protein.

Keywords: BLAHtetrone, Docking, G-protein coupled receptor, Homology modeling, Virtual screening

Introduction

The price structure of ornamental fish is primarily determined by the body coloration, as the customer preference is for attractive coloration and ornamentation. Pigmentation in fishes is responsible for a broad spectrum of colors, which commands a better price for the fish.[1] In the vertebrate species, the melanogenesis pathway is primarily responsible for pigmentation. A polymorphic Melanocortin-1-receptor (MC1R) gene initiates the melanogenesis pathway.[2] MC1R protein is the smallest member of class A (rhodopsin-like) melanocortin (MC) receptor family of G-protein coupled receptors (GPCRs).[3] Eumelanin (dark coloration) and pheomelanin (red or yellow color) are produced in the active and inactive form of MC1R in the presence of α-MSH.[4] *Tetraodon nigroviridis* is a marine and brackish water aquarium fish with black spots on its green colored body. *T. nigroviridis* has been widely used as a model organism in genetic studies, including the color inheritance in fish species. The MC1R gene plays a vital role in the color formation; hence, the present study was envisaged to computationally characterize the MC1R protein of *T. nigroviridis* and identify natural compounds that can act as putative modulators of the MC1R protein.

*Author for Correspondence
E-mail: jahageerdar@cife.edu.in

Materials and Methods

Sequence Retrieval and Analysis

UniProt was used to retrieve the FASTA format of the MC1R protein sequence of *T. nigroviridis* with H3CV70(5) as its UniProtKB ID. Conserved domain regions of the MC1R protein sequence were identified with the help of the BLAST CD-search. The CLUSTALW was used for multiple alignments of MC1R protein sequences.

MC1R Protein Structure Analysis and Active Site Prediction

The physicochemical properties of the MC1R protein of *T. nigroviridis* were determined by the ProtParam tool. The TMHMM v.2.0 server determined the transmembrane information. The subcellular localization of the MC1R protein was predicted with the Hum-mPLoc v3.0 server.[6] For the secondary structure prediction, the SOPMA server[7] was used. The tertiary structure for the MC1R protein was predicted by employing homology modeling with the help of the SWISS-MODEL.[8] The Ramachandran plot generated by the PDBsum server was used to validate the predicted tertiary model. The COACH server[9] was used for predicting the active site residue.

Virtual Screening of MC1R Protein with Natural Compounds

For virtual screening, a total of 180 313 natural compounds library was prepared from the ZINC12 database. The AutoDock-Vina[10] was used for docking of the MC1R protein to the ligands. With a threshold

value of -7 kcal/mol, the putative ligands were selected based upon their lowest binding efficiencies and ADME properties, which were tested by the ADMETsar server.

Results and Discussion

Sequence Analysis

The nucleotide sequence length of the MC1R gene of *T. nigroviridis* was 978 bp, which encodes an amino acid sequence length of 325. Conserved domain analysis affirmed that the MC1R protein belonged to the GPCR family, with three domains lying between the amino acids 62 to 302. The first one was the serpentine type 7TM GPCR Srsx domain, which was found at the N-terminal and is a chemoreceptor. A seven-transmembrane receptor from the Class A-Rhodopsin family was identified as the second domain. The third domain was a G-protein coupled chemokine receptor, which was found at the C-terminal end. Chemoreceptors in the first domain suggest that the MC1R protein is capable of intaking chemical compounds. Moderate to high homology ranging from 52% to 84% was found between MC1R protein of *T. nigroviridis* with human, Guppy, mouse, pig, cattle, and dog MC1R, depicting that MC1R protein is conserved across the species. Earlier studies have also suggested that MC1R is conserved across the species.[11]

MC1R Protein Structure Analysis and Active Site Prediction

The molecular weight of H3CV70 was predicted to be 37043.99 da and the theoretical pI 6.61, indicating a negatively charged, acidic protein. The Hum-mPLoc v3.0 and TMHMM server showed that the MC1R protein is localized in the plasma membrane and comprises seven trans-membranes. The transmembrane sequence positions are provided in Table 1. The secondary structure analysis of the MC1R protein revealed that the protein is composed of 44.00% Alpha helix (Table 2). The presence of a high percentage of α-helix implies that the MC1R protein of *T. nigroviridis* is highly robust and stable to mutations due to a higher number of inter-residue contacts in helices.[12]

Beta1 adrenergic receptor, a GPCR protein[13] of turkey with a sequence identity of 31.44%, was used as a modeling template for predicting the tertiary structure of MC1R protein of *T. nigroviridis*. The predicted tertiary structure is shown in Fig. 1. The global model quality estimate (GMQE) of 0.57 and

Table 1 — Expression of MC1R protein amino acid residues in different regions

Transmembrane region information				
Query accession number	Server	Position	Start amino acid	End amino acid
H3CV70	TMHMM2.0	Outside	1	50
H3CV70	TMHMM2.0	Tmhelix	51	70
H3CV70	TMHMM2.0	inside	71	82
H3CV70	TMHMM2.0	Outside	83	105
H3CV70	TMHMM2.0	Tmhelix	106	124
H3CV70	TMHMM2.0	inside	125	147
H3CV70	TMHMM2.0	Outside	148	167
H3CV70	TMHMM2.0	Tmhelix	168	190
H3CV70	TMHMM2.0	inside	191	199
H3CV70	TMHMM2.0	Outside	200	222
H3CV70	TMHMM2.0	Tmhelix	223	242
H3CV70	TMHMM2.0	inside	243	265
H3CV70	TMHMM2.0	Outside	266	279
H3CV70	TMHMM2.0	Tmhelix	280	302
H3CV70	TMHMM2.0	inside	303	325

Table 2 — Secondary structure prediction of *T. nigroviridis* using SOPMA server

Protein structure unit	No. of amino acids	Structural unit (%)
Alpha helix	143	44.00
3_{10} helix	0	0.00
Pi helix	0	0.00
Beta bridge	0	0.00
Extended strand	89	27.38
Beta-turn	23	7.08
Bend region	0	0.00
Random coil	70	21.54
Ambiguous state	0	0.00
Other states	0	0.00

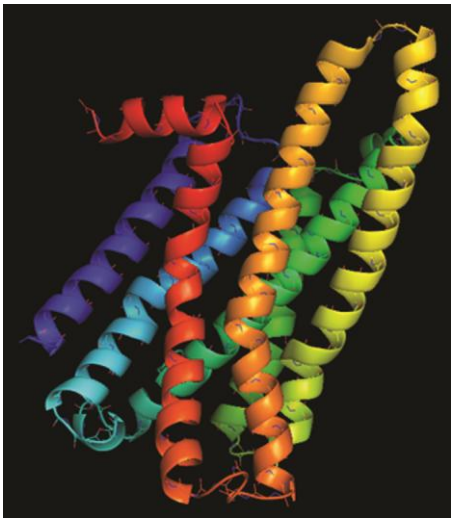

Fig. 1 — Predicted tertiary model of MC1R protein of *T. nigroviridis* by using SWISS-MODEL, based on homology modelling, viewed using PyMol

qualitative model energy (QMEAN) of 3.44 was present between the MC1R protein and the template sequence. The Ramachandran plot analysis revealed that of the 272 residues, 234 (91.1%) amino acid residues were present in the most favoured core region. The rest 23 (8.9%) residues were present in the additionally allowed regions. The presence of a high percentage of residues in the favoured region indicates that the predicted structure of the MC1R protein of *T. nigroviridis* is reliable and was used for further analysis.

The confidence score (C-score) for the MC1R protein of *T. nigroviridis* was found to be 0.34 with the 2ycxA_BS01_P32 template cluster. The amino acid residues of TM-5 and TM-6, i.e., I-127, D-128, C-132, L-199, T-206, F-209, W-258, F-262, L-265, L-288, I-292, were predicted as the active site residues using the COACH server. These are the part of the conserved regions of the MC1R protein of *T. nigroviridis* and humans and are suitable sites for ligand binding activity.[14]

Virtual Screening and Molecular Docking

From a list of 180 313 natural compounds, three compounds 1H-Indole-2-methanol (ZINC20763420), BLAHtetrone (ZINC05220992) and 2-naphthyl BLAHone (ZINC04237106), had the lowest binding efficiency and acceptable ADME property and were selected for usage in fishes (Table 3). With the binding efficiency of -9.70 kcal/mol, the natural compound BLAHtetrone (ZINC05220992), extracted from *Camu camu* berry, was found to be a better modulator of the MC1R protein of *T. nigroviridis* than others. Figs 2 and 3 depicted the 2D and 3D docked sites for BLAHtetrone to the protein. Since the identified compounds are bound in the active sites of the MC1R, they may have the ability to act as the MC1R modulators, which further needs to be confirmed with lab validation and appropriate biochemical experiments.

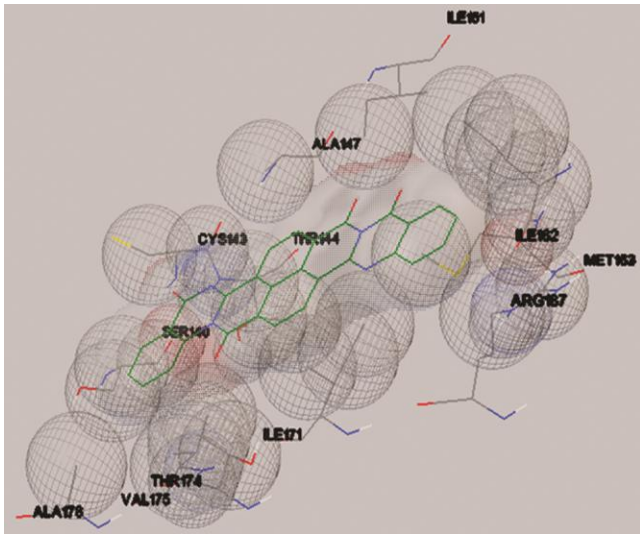

Fig. 2 — 2D ligand (BLAHtetrone) – MC1R protein interaction map viewed using AutoDock

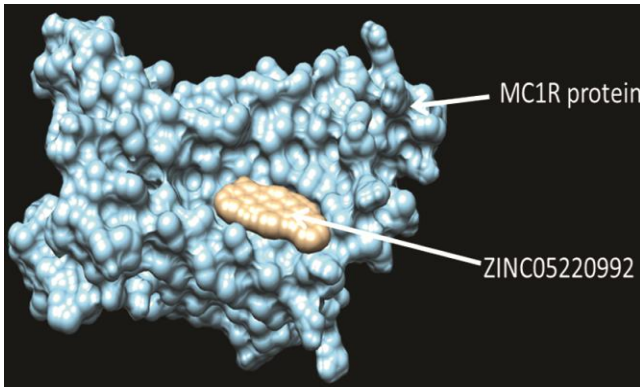

Fig. 3 — 3D ligand (BLAHtetrone) – MC1R protein interaction map viewed using AutoDock

Table 3 — Binding efficiencies of the natural compounds to the MC1R protein of *T. nigroviridis* along with the ADMET properties of top 3 ZINC compounds

S.No.	Ligands	Binding efficiency (kcal/mol)	Hydrogen bond donors	Hydrogen bond acceptors	Molecular mass (Dalton)	Log P	Fish toxicity	Chemical
1	1H-Indole-2-methanol (ZINC20763420)	-9.90	2	8	497.60	3.28	low	
2	BLAHtetrone (ZINC05220992)	-9.70	0	8	468.43	4.39	low	
3	2-naphthylBLAHone (ZINC04237106)	-8.30	2	3	317.41	3.20	low	

Conclusions

MC1R protein is the primary protein that governs the melanogenesis pathway. The study of *T. nigroviridis* MC1R protein provides a useful insight into its protein structure. The identified putative modulators may help to modify the body pigmentation of *T. nigroviridis*. There is a high similarity among the MC1R sequences of various species. It is suggested that the *T. nigroviridis* MC1R protein can be used as a model to understand the protein structure of the majority of the other fish species. Further, it can also be used to explore the possibilities of developing novel MC1R modulators for commercially important aquarium fishes.

Acknowledgment

The authors are grateful to the Director, ICAR-Central Institute of Fisheries Education, Mumbai, India, for providing the facility to carry this work. The authors are thankful to Ms Archana N Panche, Assistant Professor, MGM's Institute of Biosciences & Technology Aurangabad, Maharashtra, for the technical guidance. The first author acknowledges the financial support provided by the ICAR, New Delhi, in the form of a junior research fellowship (JRF).

References

1 Gupta S K, Jha A K, Pal A K & Venkateshwarlu G, Use of natural carotenoids for pigmentation in fishes, *Nat Prod Rad*, **6** (2007) 46–49

2 Dores R M, Londraville R L, Prokop J, Davis P, Dewey N & Lesinski N, Molecular evolution of GPCRs: melanocortin/ melanocortin receptors, *J Mol Endocrinol*, **52(3)** (2014), T29–T42.

3 Horrell E M W, Boulanger M C & D'Orazio J A, Melanocortin 1 receptor: structure, function & regulation, *Front Genet*, **7 (**2016) 95.

4 Virador V M, Muller J, Wu X, Abdel-Malek Z A, Yu Z X, Ferrans V J, Kobayashi N, Wakamatsu K, Ito S, Hammer J A & Hearing, V J, Influence of α-melanocyte-stimulating hormone & ultraviolet radiation on the transfer of melanosomes to keratinocytes, *The FASEB Journal*, **16(1)** (2002) 105–107.

5 Jaillon O, Aury J M, Brunet F, Petit J L, Stange-Thomann N, Mauceli E, Bouneau L, Fischer C, Ozouf-Costaz C, Bernot A & Nicaud S, Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype Nature, **431** (2004) 946–957.

6 Zhou H, Yang Y & Shen H-B,Hum-mPLoc 3.0: Prediction enhancement of human protein subcellular localization through modeling the hidden correlations of gene ontology & functional domain features, *Bioinformatics*, **33(6)** (2016) 843–853.

7 Geourjon C & Deleage G, SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments *Bioinformatics*, **11(6)** (1995) 681–684.

8 Biasini M, Stefan B, Andrew W, Konstantin A, Gabriel S, Tobias S, Florian K, Tiziano G C, Martino B, Lorenza B & Torsten S, SWISS-MODEL: modelling protein tertiary & quaternary structure using evolutionary information, *Nucleic Acids Res*, **42(W1)** (2014) 252–258.

9 Yang J, Roy A & Zhang Y, BioLiP: a semi-manually curated database for biologically relevant ligand-protein interactions, *Nucleic Acids Res*, **41(D1)** (2013) D1096–D1103.

10 Morris G M, Huey R, Lindstrom W, Sanner M F, Belew R K, Goodsell D S & Olson A J, Autodock4 & AutoDockTools4: automated docking with selective receptor flexibility, *J Computational Chemistry*, **16** (2009) 2785–2791.

11 Switonski M, Mankowska M & Salamon S, Family of melanocortin receptor (MCR) genes in mammals—mutations, polymorphisms & phenotypic effects, *J Appl Genet*, **54** (2013) 461–472.

12 Abrusán G & Marsh J A, Alpha helices are more robust to mutations than beta strands, *PLoS Comput Biol* **12(12)** (2016) e1005242.

13 Moukhametzianov R, Warne T, Edwards P C, Serrano-Vega M J, Leslie A G, Tate C G and Schertler G F, Two distinct conformations of helix 6 observed in antagonist-bound structures of a β1-adrenergic receptor, *Proc Natl Acad Sci U S A*, **108** (2011) 8228–8232.

14 Wolf Horrell, E M, Boulanger, M C & D'Orazio J A, Melanocortin 1 receptor: structure, function, & regulation, *Front Genet*, **7** (2016) 95.