

# Characterization of the Functional Domain of $\beta_2$ -Microglobulin from the Asian Seabass, *Lates calcarifer*

Hirzahida Mohd-Padil, Khairina Tajul-Arifin\*, Adura Mohd-Adnan

School of Biosciences and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

## Abstract

**Background:**  $\beta_2$ -Microglobulin ( $\beta_2$ M) is the light chain of major histocompatibility class I (MHC I) that binds non-covalently with the  $\alpha$  heavy chain. Both proteins attach to the antigen peptide, presenting a complex to the T cell to be destroyed via the immune mechanism.

**Methodology/Principal Findings:** In this study, a cDNA sequence encoding  $\beta_2$ M in the Asian seabass (*Lates calcarifer*) was identified and analyzed using *in silico* approaches to predict and characterize its functional domain. The  $\beta_2$ M cDNA contains an open reading frame (ORF) of 351 bases with a coding capacity of 116 amino acids. A large portion of the protein consists of the IG constant domain (IGc1), similar to  $\beta_2$ M sequences from other species studied thus far. Alignment of the IGc1 domains of  $\beta_2$ M from *L. calcarifer* and other species shows a high degree of overall conservation. Seven amino acids were found to be conserved across taxa whereas conservation between *L. calcarifer* and other fish species was restricted to 14 amino acids at identical conserved positions.

**Conclusion/Significance:** As the *L. calcarifer*  $\beta_2$ M protein analyzed in this study contains a functional domain similar to that of  $\beta_2$ M proteins in other species, it can be postulated that the  $\beta_2$ M proteins from *L. calcarifer* and other organisms are derived from a common ancestor and thus have a similar immune function. Interestingly, fish  $\beta_2$ M genes could also be classified according to the ecological habitat of the species, i.e. whether it is from a freshwater, marine or euryhaline environment.

**Citation:** Mohd-Padil H, Tajul-Arifin K, Mohd-Adnan A (2010) Characterization of the Functional Domain of  $\beta_2$ -Microglobulin from the Asian Seabass, *Lates calcarifer*. PLoS ONE 5(10): e13159. doi:10.1371/journal.pone.0013159

**Editor:** Vladimir Brusic, Dana-Farber Cancer Institute, United States of America

**Received:** June 10, 2010; **Accepted:** September 2, 2010; **Published:** October 6, 2010

**Copyright:** © 2010 Mohd-Padil et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Ministry of Science, Technology and Innovation (MOSTI), Malaysia under the Top-Down program (Project no. 05-02-02-T002). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: kay.tarifin@gmail.com

## Introduction

$\beta_2$ M is the light chain component of the class I major histocompatibility complex (MHC I) molecule. It consists of about 99 residues with a seven-stranded  $\beta$ -sandwich fold and a central disulfide bond [1] and belongs to the antibody constant domain-like family of proteins (immunoglobulin superfamily). At the cell surface, the MHC I complex is comprised of three extracellular domains from the  $\alpha$  chain ( $\alpha 1$ ,  $\alpha 2$  and  $\alpha 3$ ) plus the  $\beta_2$ M protein domain.  $\beta_2$ M ensures the proper folding and cell surface display of the MHC I molecule [2]. The classical MHC I molecule mainly functions as a component that binds antigenic peptides, presenting them to the T-cell receptor to trigger the cellular immune response [3].

It has been reported that there is a high degree of conservation of  $\beta_2$ M sequences of mammalian species as well as between mammalian and avian species [4,5]. The teleost  $\beta_2$ M sequences also exhibit high sequence similarity overall and conserved regions with warm-blooded vertebrates [6]. Previous phylogenetic analysis revealed the evolutionary diversion of the  $\beta_2$ M protein in warm-blooded vertebrates and fish [7,8]. Meanwhile, another earlier phylogenetic analysis indicated that the freshwater fish  $\beta_2$ M gene diverged from the common ancestor gene earlier than the seawater fish  $\beta_2$ M gene [6].

In an effort to improve our understanding of the molecular biology of *L. calcarifer*, several thousand expressed sequence tags (ESTs) have been derived from various cDNA libraries from several tissues [9]. Analyses of the EST data enabled us to identify novel gene sequences, including those with significant similarity to  $\beta_2$ M. Using the numerous  $\beta_2$ M sequences that are available in the public databases, we analyzed the protein sequence in an effort to better understand the immune system of *L. calcarifer*.

## Materials and Methods

The cDNA sequence of *L. calcarifer*  $\beta_2$ M obtained from the spleen EST library was translated into its potential open reading frame (ORF) using the ORF Finder algorithm (<http://www.ncbi.nlm.nih.gov/gorf/>). Domain analyses were carried out using several resources, including Simple Modular Architecture Research Tools (SMART) (<http://smart.embl-heidelberg.de/>) [10], Pfam 20.0 (<http://pfam.wustl.edu/>) [11] and Prosite 19.36 (<http://www.expasy.org/prosite/>) [12]. The profile of the IGc1 domain obtained from the Pfam domain database was used to search for other homologous proteins using the *hmmsearch* program in HMMER version 2.3.2 (<http://hmm.janelia.org/>) [13,14] in both the Swiss-Prot database Release 54 (<http://www.ncbi.nlm.nih.gov/>) and the fish genomes at Ensembl database Release 49

**Table 1.**  $\beta$ <sub>2</sub>M sequences used in this analysis.

Group	Protein ID	Protein Name	Species Name	Common Name	Reference
Fish	Q8AXA0	Reg	<i>R. eglanteria</i>	Clearnose skate	[19]
	P55076	Bin	<i>L. intermedius</i>	Lake Tana barbels	[26]
	O42197	Ipu	<i>I. punctatus</i>	Channel catfish	[6]
	Q04475	Dre1	<i>D. rerio</i>	Zebrafish	[4]
	Q03422	Cca	<i>C. carpio</i>	Common carp	[27]
	Q9PRF8	Aba	<i>A. baerii</i>	Siberian sturgeon	[18]
	Q8AYH8	Pol	<i>P. olivaceus</i>	Japanese flounder	[8]
	Q90ZJ6	Ola	<i>O. latipes</i>	Japanese medaka	[28]
	NP_998291	Dre2	<i>D. rerio</i>	zebrafish	[29]
Amphibian	Q9IA97	Xla	<i>X. laevis</i>	African clawed frog	[7]
Monotremes	Q864T7	Oan	<i>O. anatinus</i>	Platypus	[30]
	Q864T6	Tac	<i>T. aculeatus</i>	Australian echidna	[30]
Avians	P21611	Gga	<i>Gallus gallus</i>	Chicken	[31]
	P21612	Mga	<i>M. gallopavo</i>	Turkey	[32]
Marsupials	Q9GKM2	Tvu	<i>T. vulpecula</i>	Brush-tail possum	[33]
	Q864T8	Mdo	<i>M. domestica</i>	Short-tailed opossum	[30]
Ruminants	Q6QAT4	Oar	<i>O. aries</i>	Sheep	Unpublished data
	P01888	Bta	<i>B. taurus</i>	Cattle	[34]
Equine	P30441	Eca	<i>E. caballus</i>	Horse	[35]
	Q861S3	Eas	<i>E. asinus</i>	Ass	[36]
Rodents	P01887	Mmu	<i>M. musculus</i>	Mouse	[37]
	P07151	Rno	<i>R. norvegicus</i>	Rat	[38]
	Q8CIQ3	Shi	<i>S. hispidus</i>	Hispid pocket mouse	Unpublished data
	Q9WV24	Cgr	<i>C. griseus</i>	Chinese hamster	[39]
	P01886	Cpo	<i>C. porcellus</i>	Domestic guinea pig	[40]
Primates	P55079	Soe	<i>S. oedipus</i>	Cotton-top tamarin	[41]
	Q9TSX4	Sfu	<i>S. fuscicollis</i>	Brown-headed tamarin	[41]
	O77517	Smn	<i>S. niger</i>	Black-handed tamarin	[41]
	O77518	Sim	<i>S. imperator</i>	Tamarin	[41]
	P63068	Sbb	<i>S. bicolor</i>	Pied bare-faced tamarin	[42]
	O77531	Pir	<i>P. irrorata</i>	Bald-faced saki	[41]
	Q6PZD3	Cae	<i>C.aethiops</i>	African green monkey	[43]
	O77529	Cho	<i>C. hoffmannsi</i>	Hoffmanns's titi	[41]
	O77532	Csa	<i>C. satanas</i>	Black-bearded saki	[41]
	Q71UN5	Cpe	<i>C. penicillata</i>	Black pencilled marmoset	[42]
	O77533	Cme	<i>C. melanocephalus</i>	Black headed Uacari	[42]
	O77519	Lch	<i>L. chrysopygus</i>	Golden-rumped lion tamarin	[42]
	O77535	Cpy	<i>C. pygmaea</i>	Pygmy marmoset	[42]
	O77530	Cto	<i>C. torquatus</i>	Yellow-handed titi	[42]
	O77526	Cpp	<i>C. personatus</i>	Masked titi	[42]
	O77528	Cpn	<i>C. p. nigrifrons</i>	Black titi	[42]
	O77521	Cem	<i>C. emiliae</i>	Emilia's marmoset	[42]
	O77536	Apa	<i>A. paniscus</i>	Black spider monkey	[42]
	O77534	Sbo	<i>S. boliviensis</i>	Bolivian squirrel monkey	[42]
	O77525	Lla	<i>L. lagotricha</i>	Brown woolly monkey	[42]
O77523	Ase	<i>A. seniculus</i>	Howler monkey	[42]	
O77520	Cgo	<i>C. goeldii</i>	Goeldi's marmoset	[42]	
P63063	Ale	<i>A. lemurinus</i>	Lemurine night monkey	[42]	
O77537	Aaz	<i>A. azarai</i>	Azara's night monkey	[42]	
Q6T672	Pan	<i>P. anubis</i>	Olive baboon	Unpublished data	

**Table 1.** Cont.

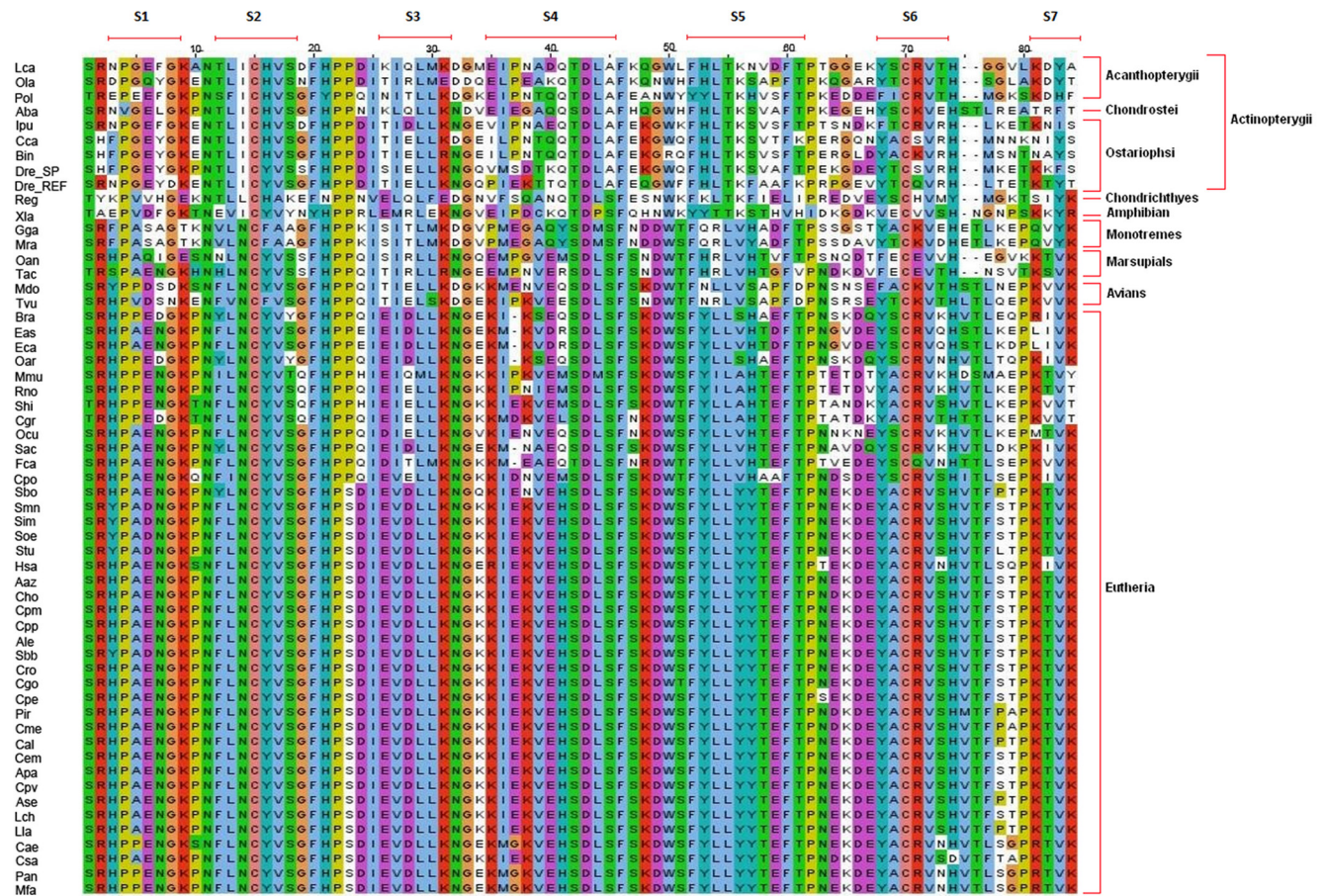
Group	Protein ID	Protein Name	Species Name	Common Name	Reference
	Q8SPW0	Mfa	<i>M. fascicularis</i>	Crab-eating macaque	Unpublished data
	P61769	Has	<i>H. sapiens</i>	Human	[44]
Others	Q07717	Scs	<i>S. scrofa</i>	Pig	[45]
	P01885	Ocu	<i>O. cuniculus</i>	Rabbit	[46]
	Q5MGS7	Fca	<i>F. catus</i>	Domestic cat	Unpublished data

The 55 protein sequences used in this study were retrieved from the Swiss-Prot and Refseq databases. Protein ID shows the accession number of the sequence in the database while Protein Name is the protein name designated in this study. Equine, Rodents, Ruminants, Primates and Others are largely grouped as Eutheria. doi:10.1371/journal.pone.0013159.t001

http://www.ensembl.org). The IGC1 domain sequences of the homologous proteins thus identified were extracted for subsequent analyses.

The sequence alignment for the IGC1 domains was built using the *hmalign* program in the HMMER package against the profile of the IGC1 domain obtained from Pfam to enable the pattern of  $\beta_2$ M protein change across the taxa to be examined. PHYLIP (http://evolution.genetics.washington.edu/phylip) [15] was then used to perform phylogenetic analyses. A neighbor-joining tree was built using

the *protdist* and *neighbor* programs with the Jones-Taylor-Thornton substitution model. The robustness of the trees was evaluated by bootstrap analysis of 1000 random iterations using *seqboot*, while *consense* was used to generate the consensus tree. All programs used to construct the phylogenetic trees are contained in PHYLIP packages [15]. Subsequently, MEGA4 (http://megasoftware.net/) [16] was utilized to view the resultant phylogenetic trees. The *L. calcarifer*  $\beta_2$ M sequence analyzed in this study has been deposited in the GenBank database under the accession number FJ200516.



**Figure 1. Multiple sequence alignment of IGC1 domains.** The alignment consists of IGC1 domain sequences from organisms of various taxa such as Eutheria, Marsupials, Monotremes, Avians, Chondrichthyes fish and Actinopterygii fishes. Lca is the *L. calcarifer* protein sequence and the conserved residues are marked by (\*). S1, S2, S3, S4, S5, S6 and S7 indicate the regions of seven  $\beta$  strands in the IGC1 domain. Numbers at the top indicate amino acid positions. Information on the sequences is given in Table 1. doi:10.1371/journal.pone.0013159.g001

**Results**

**Analysis of the *L. calcarifer*  $\beta_2$ M Sequence**

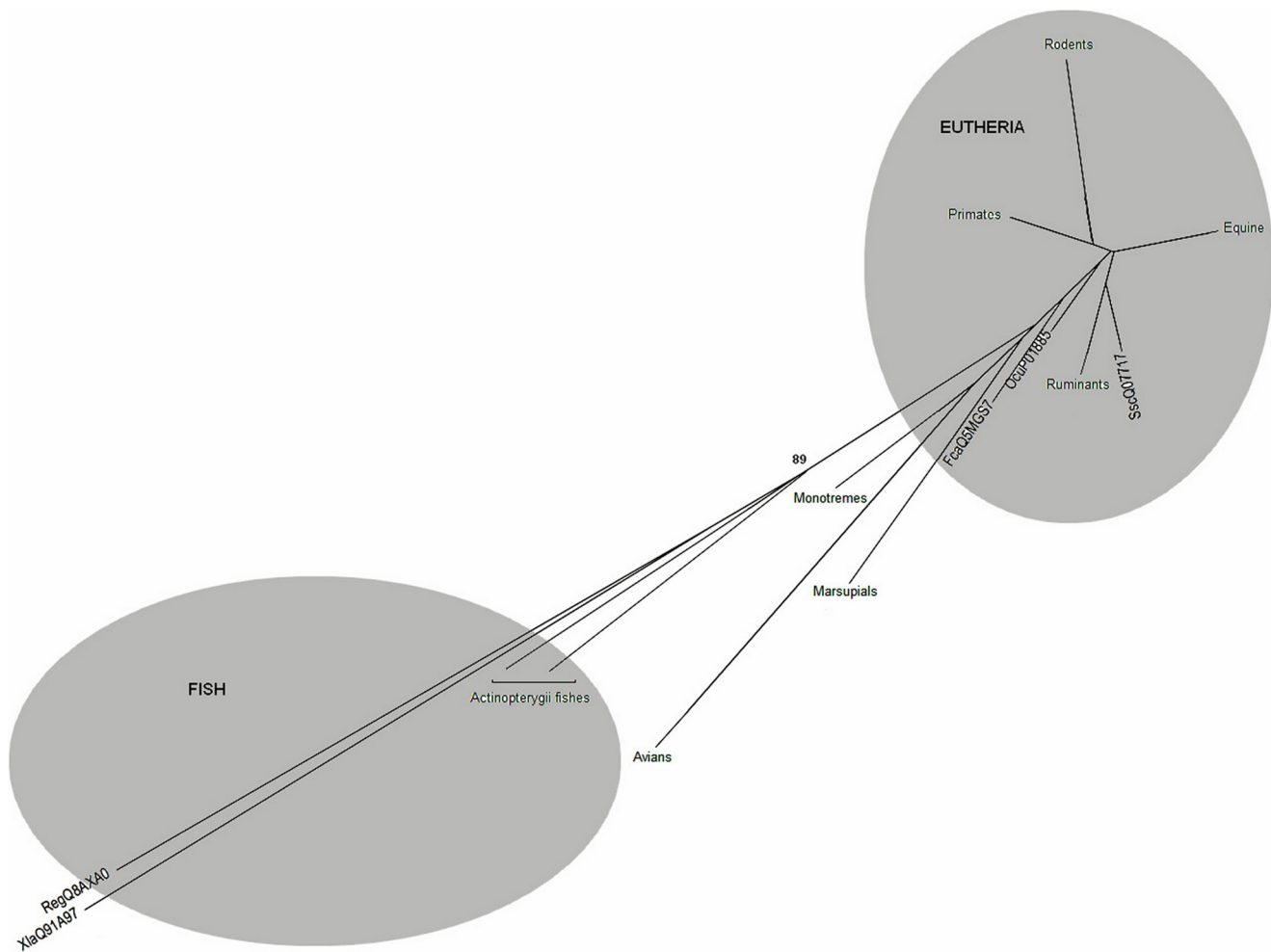
Analyses of the cDNA sequence of *L. calcarifer*  $\beta_2$ M (clone LSE48F06) from the spleen EST library indicated the most probable ORF codes for a polypeptide of 116 amino acids in length. A domain search revealed that a large portion of the protein sequence matched the immunoglobulin C-type (IGc1) domain in the SMART, Pfam and Prosite databases. Almost half of the amino acid residues of  $\beta_2$ M form two large  $\beta$  structures, which are linked by a central disulfide bond. Its conformation thus strongly resembles the overall tertiary structure of the IgC1 domain [17]. In addition, analysis against the Prosite database showed the presence of an immunoglobulin and major histocompatibility complex protein signature, YSCRVTH, located at residues 97–103 in the *L. calcarifer*  $\beta_2$ M sequence.

A total of 81 IGc1 domains contained in  $\beta_2$ M sequences were obtained by protein search against the Swiss-Prot database (version 14 updated 23 October 2007) and the known proteins of five fish species (medaka, stickleback, zebrafish, pufferfish and

spotted green pufferfish) in the Ensembl database (version 49 updated March 2008). Of these 81 domain sequences, only 56 were used to build the multiple sequence alignment (MSA) (see Table 1); the remaining sequences were excluded as they were considered identical, truncated or replicated between the two databases used. As the IGc1 domain is the functional domain in all  $\beta_2$ M sequences analyzed in this study, subsequent analyses focused on this domain as a representation of the  $\beta_2$ M protein.

**Sequence Alignment**

Alignment of the IGc1 domains of  $\beta_2$ M from *L. calcarifer* and other species showed a high degree of overall conservation across taxa (see Figure 1). The IGc1 domain of *L. calcarifer* starts with a Ser residue, which is conserved among the taxa, with the exception of domains from the Japanese flounder (Q8AYH8), Chinese hamster (Q9WV24), hispid pocket mouse (Q8CIQ3) and Australian echidna (Q864T6), in which the starting residue is a Thr. The fish sequences, with the exception of Siberian sturgeon (Q9PRF8), are two residues shorter (lacking residue-75 and residue-76) than the human (P61769) sequence. The deletions in



**Figure 2. NJ phylogenetic tree of  $\beta_2$ M protein sequences representing whole organisms.** The phylogenetic tree shown is the collapsed tree of 55 sets of sequence data. This tree shows that  $\beta_2$ M sequences are clustered together according to their taxons.  $\beta_2$ M sequences from Eutheria are clustered together and consist of sequences from Primates, Equine, Rodents, Ruminants, SscQ07717, OcuP01885 and FcaQ5MG57. Marsupials, Monotremes and Avians are the intermediate taxons between Eutheria and Fish. Amphibian Xla protein Q91A97 is clustered together with Actinopterygii fishes while the outgroup in this tree is a cartilaginous fish Reg Q8AXA0. The divergence of fish and mammalian  $\beta_2$ M received a high bootstrap value (89) to support the reliability of this phylogenetic tree. doi:10.1371/journal.pone.0013159.g002

the teleost sequences are located in the loop between the anti-parallel beta-strands S6 and S7 [18]. Seven positions or residues, Asn<sub>11</sub>, Cys<sub>15</sub>, Pro<sub>22</sub>, Asp<sub>44</sub>, Phc<sub>47</sub>, Cys<sub>70</sub> and Val<sub>72</sub> (numbering refers to the position in the MSA) are found to be completely conserved. These residues or regions are believed to be in the active or binding site of the  $\beta_2$ M protein. A previous study reported that the two cysteine residues (Cys<sub>15</sub> and Cys<sub>70</sub>), which covalently link to form a disulfide bridge (thus connecting the two  $\beta$  sheets), are important elements in protein folding that contribute to stabilization of the MHC class I molecule [1].

### Phylogenetic Analysis

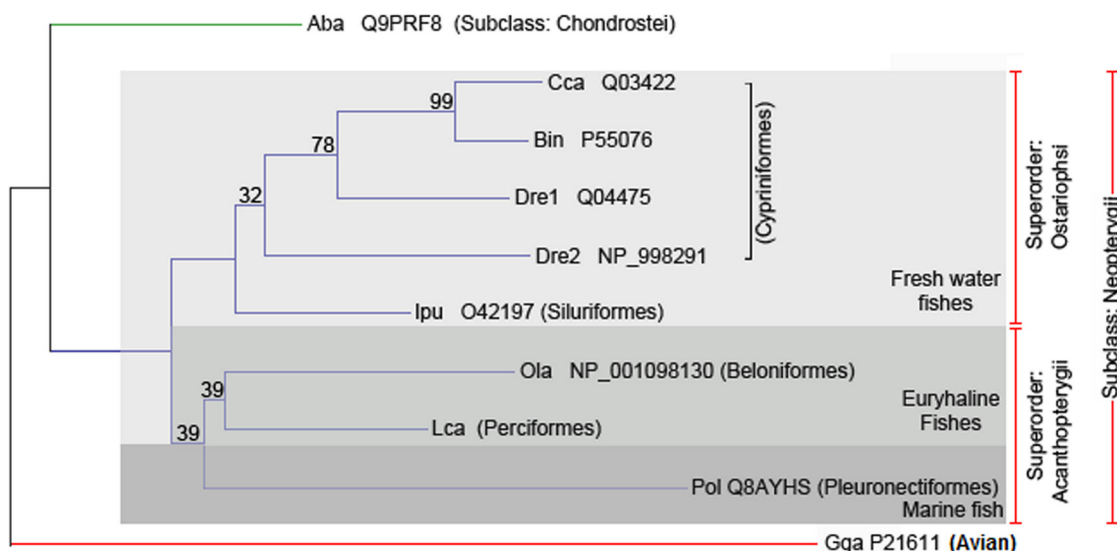
The phylogenetic tree built is an unrooted tree (see Figure 2) with two distinct clades, the mammalian  $\beta_2$ M and fish  $\beta_2$ M. The avian, marsupial and monotreme sequences form an intermediate group between the fish and the mammalian clades. All fish sequences are from the Actinopterygii class with the exception of the clearnose skate (Q8AXA0) [19], which is a cartilaginous fish (class Chondrichthyes). In order to clarify the relationships among the  $\beta_2$ M molecules of the Actinopterygii fish (Siberian sturgeon, channel catfish (O42197), Japanese medaka (Q90ZJ6), Japanese flounder, zebrafish (Q04475, NP\_998291), common carp (Q03422) and Lake Tana barbel (P55076)), an NJ tree consisting of fish sequences only was constructed using the chicken (P21611)  $\beta_2$ M sequence as an outgroup (see Figure 3). The  $\beta_2$ M from *L. calcarifer*, which is a euryhaline and catadromous species, clustered together with Japanese medaka  $\beta_2$ M, forming a separate clade from the freshwater fishes (common carp, Lake Tana barbel, zebrafish and channel catfish). The phylogenetic trees also reveal a molecular signal of the ecological distinction between marine and freshwater fish. The Siberian sturgeon  $\beta_2$ M sequence is the most basal lineage, which is placed outside the main cluster of teleosts in the tree (see Figure 3).

### Discussion

Our analysis of the novel *L. calcarifer*  $\beta_2$ M gene recovered in this study indicates that, overall, fish  $\beta_2$ M sequences have high sequence similarity and share many conserved features with published sequences from mammals and birds. Using  $\beta_2$ M as both a

phylogenetic marker and a source of information, we confirmed previous studies indicating that the  $\beta_2$ M proteins of mammals and fish represent clearly distinct evolutionary paths, with fish  $\beta_2$ M genes more closely related to avian sequences than to those of mammals [7,8]. The divergence between fish and mammals is partially a consequence of several unreversed changes in the ORF. For example, at site-14 (S2) in all mammalian  $\beta_2$ M sequences, the residue is Arg, whereas in Actinopterygii fish  $\beta_2$ M sequences it is Ile. The Ser residue in mammalian sequences is substituted by Ala at site-45 (S4) in all Actinopterygii fish, whereas Lys at position 55 (S5) in mammalian  $\beta_2$ M is consistently replaced by Thr in all Actinopterygii fish. Although these changes of amino acids are located within the mature protein region, the residues at those sites are not involved in any important stabilizing interactions of the protein [20].

Our analyses of the Actinopterygii fish  $\beta_2$ M sequences employed sequences representing two subclasses: Neopterygii (teleosts) and Chondrostei. Within the Neopterygii, two superorders are evident: Ostariophsi and Acanthopterygii. The molecular results showed phylogenetic relationships that support those established based on fish morphology [21]. Ostariophsi includes two different orders (Cypriniformes and Siluriformes), whereas Acanthopterygii includes three different fish orders (Beloniformes, Perciformes and Pleuronectiformes). With the exception of Siberian sturgeon, our results also showed that fishes included in Ostariophsi are mainly freshwater fishes, whereas those of Acanthopterygii are marine or euryhaline fishes (see Figure 3). The results reconfirm a previous paper that indicated an evolutionary divergence had occurred between freshwater and marine or euryhaline fish  $\beta_2$ M sequences [8]. At the molecular level among the Actinopterygii fishes studied, Asian seabass, Japanese medaka and Japanese flounder sequences share synapomorphic amino acids at three sites: Thr<sub>73</sub> (S6) Gly<sub>78</sub> (loop between S6 and S7) and Asp<sub>82</sub> (S7). Siberian sturgeon (Chondrost) resolves as the most basal lineage in agreement with other studies [21,18], confirming this fish is the most primitive member of the subclass Actinopterygii. Indeed, a two-codon (residues 75 and 76 in the alignment) deletion is synapomorphic in all teleost  $\beta_2$ M sequences in contrast to Siberian sturgeon  $\beta_2$ M.



**Figure 3. NJ phylogenetic tree of IGc1 domains present in  $\beta_2$ M protein sequences from fish.** The phylogenetic tree shows that fish  $\beta_2$ M proteins are clustered according to the fish's ecological habitat, which may be fresh water, euryhaline or marine. The chicken sequence was used as the outgroup. Information on the sequences is given in Table 1. doi:10.1371/journal.pone.0013159.g003

Given the preponderance of analytical results and qualitative comparative results, we suggest that the *L. calcarifer*  $\beta_2$ M gene recovered here is likely to function similarly to previously characterized  $\beta_2$ M genes, and that the protein it encodes acts as a light chain that binds non-covalently to the heavy chain of the MHC class I molecule. The two proteins would then create a complex with the antigen peptide and present the antigen to T cells to be destroyed by the immune mechanism [22].  $\beta_2$ M is also involved in stabilizing the MHC class I molecules [23,24]. Since  $\beta_2$ M is clearly most closely related to the IGc1-type domains of MHC class I and II, its gene must have been linked to that of the MHC at some point of evolution [7]. Furthermore, the similarity between the structures of  $\beta_2$ M and the IGc1-type domains of MHC I and II suggests that they share a common ancestor encoded in MHC genes [25]. In this study, we have identified the  $\beta_2$ M gene in *L. calcarifer* and confirm its phylogenetic placement within a group of related fish species. We believe that, as more fish  $\beta_2$ M sequences become available, reanalysis of the data may be

able to better resolve the evolutionary history of the seeming ecological divergence detected among fish sequences and that of the *L. calcarifer*  $\beta_2$ M gene from the rest of the  $\beta_2$ M gene tree. Again, the overall utility of our approach in the detection, recovery and delineation of genes within *L. calcarifer* is emphasized by its success in our study of  $\beta_2$ M.

## Acknowledgments

The authors thank Azrol Ridzuan Abd Aziz for providing the  $\beta_2$ M nucleotide sequence, Dr. Michael R. J. Fortsner for helpful discussion and revision of the manuscript and Mohd. Zulqurnain Abd. Razak for help with the figures.

## Author Contributions

Conceived and designed the experiments: KTA AMA. Performed the experiments: HMP. Analyzed the data: HMP. Wrote the paper: HMP KTA AMA. Interpreted data: KTA AMA.

## References

- Bjorkman PJ, Saper MA, Samraoui B, Bennett WS, Strominger JL, et al. (1987) Structure of the human class I histocompatibility antigen, HLA-A2. *Nature* 329: 506–512.
- Coico R, Sunshine G, Benjamin E (2003) *Immunology A Short Course*. New Jersey: John Wiley & Son, Inc. 361 p.
- Collins RW (2004) Human MHC class I chain related (MIC) genes: their biological function and relevance to disease and transplantation. *Eur J Immunogenet* 31: 105–114.
- Ono H, Figueroa F, O'huigin C, Klein J (1993) Cloning of the  $\beta_2$ -microglobulin gene in the zebrafish. *Immunogenetics* 38: 1–10.
- Hui FH, Tian YY, Ruo QY, Feng SG, Chun X (2006) cDNA cloning and genomic structure of grass carp (*Ctenopharyngodon idellus*) B2-microglobulin gene. *Fish Shellfish Immunol* 20: 118–123.
- Criscitello MF, Benedetto R, Antao A, Wilson MR, Chinchar VG, et al. (1998)  $\beta_2$ -microglobulin of *Ictalurid* catfishes. *Immunogenetics* 48: 339–343.
- Stewart R, Ohta Y, Minter RR, Gibbons T, Horton TL, et al. (2005) Cloning and characterization of *Xenopus*  $\beta_2$ -microglobulin. *Dev Comp Immunol* 29: 723–732.
- Choi W, Lee EY, Choi TJ (2006) Cloning and sequence analysis of the  $\beta_2$ -microglobulin transcript from flounder, *Paralichthys olivaceus*. *Mol Immunol* 43: 1565–1572.
- Tan SL, Mohd-Adnan A, Mohd-Yusof NY, Forstner MRJ, Wan K-L (2008) Identification and analysis of a prepro-chicken gonadotropin releasing hormone II (preprocGnRH-II) precursor in the Asian seabass, *Lates calcarifer*, based on an EST-based assessment of its brain transcriptome. *Gene* 411: 77–86.
- Schultz J, Milpetz F, Bork P, Ponting CP (1998) SMART, a simple modular architecture research tool: identification of signaling domain. *Proc Natl Acad Sci U S A* 95: 5857–5864.
- Robert DF, Jaina M, Benjamin SB, Sam GJ, Volker H, et al. (2006) Pfam: clans, web tools and services. *Nucleic Acids Res* 34: 247–251.
- Hulo N, Bairoch A, Bulliard V, Cerutti L, De Castro E, et al. (2006) The Prosite database. *Nucleic Acids Res* 34: 227–230.
- Eddy SR (1998) Profile hidden Markov models. *Bioinformatics* 14: 755–763.
- Durbin R, Eddy S, Krogh A, Mitchison G (1998) *Biological sequence analysis: probabilistic models of proteins and nucleic acids*. United Kingdom: Cambridge University Press. 95 p.
- Felsenstein J (1996) Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. *Methods Enzymol* 266: 418–427.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24(8): 1596–1599.
- Becker JW, Reece GN (1985) Three-dimensional structure of  $\beta_2$ -microglobulin. *Proc Natl Acad Sci U S A* 82: 4225–4229.
- Lundqvist ML, Appelkvist P, Hermsen T, Pilstrom L, Stet RJM (1999) Characterization of beta-2-microglobulin in a primitive fish, the Siberian sturgeon (*Acipenser baeri*). *Immunogenetics* 50(1-2): 79–83.
- Cannon JP, Haire RN, Litman GW (2002) Identification of diversified genes that contain immunoglobulin-like variable regions in a protochordate. *Nat Immunol* 3(12): 1200–1207.
- Benyamini H, Gunasekaran K, Wolfson H, Nussinov R (2003)  $\beta_2$ -Microglobulin amyloidosis: insights from conservation analysis and fibril modeling by protein docking techniques. *J Mol Biol* 330: 159–174.
- Hurley IA, Mueller RL, Dunn KA, Schmidt EJ, Friedman M, et al. (2007) A new time-scale for ray-finned fish evolution. *Proc Biol Sci* 274: 489–498.
- Hansen TH, Lee DR (1997) Mechanism of class I assembly with  $\beta_2$ -microglobulin and loading with peptide. *Adv Immunol* 64: 105–137.
- Van KL, Ashton RPG, Ploegh HL, Tonegawa S (1992) TAP1 mutant mice are deficient in antigen presentation, surface class molecules, and CD4+8+ cells. *Cell* 71: 1205–1214.
- Zijlstra M, Bix M, Simister N, Loring J, Raulat D, et al. (1990)  $\beta_2$ -Microglobulin deficient mice lack CD4+8+ cytolytic T cells. *Nature* 344: 742–746.
- Shum BP, Azumi K, Zhang S, Kehrer SR, Raison RL, et al. (1996) Unexpected  $\beta_2$ -microglobulin sequence diversity in individual rainbow trout. *Proc Natl Acad Sci U S A* 93: 2779–2784.
- Dixon B, Nagelkerke LA, Sibbing FA, Egberts E, Stet RJ (1996) Evolution of MHC class II beta chain-encoding genes in the Lake Tana barbel species flock (*Barbus intermedius* complex). *Immunogenetics* 44(6): 419–431.
- Dixon B, Van ESHM, Rodrigues PNS, Egberts E, Stet RJM (1993) Fish major histocompatibility complex genes: an expansion. *Dev Comp Immunol* 19: 109–133.
- Naruse K, Fukamachi S, Mitani H, Kondo M, Matsuoka T, et al. (2000) A detailed linkage map of medaka, *Oryzias latipes*: comparative genomics and genome evolution. *Genetics* 154(4): 1773–1784.
- Vihtelic TS, Fadool JM, Gao J, Thornton KA, Hyde DR, et al. (2005) Expressed sequence tag analysis of zebrafish eye tissues for NEIBank. *Mol Vis* 11: 1083–1100.
- Miska KB, Hellman L, Miller RD (2003) Characterization of  $\beta_2$ -microglobulin coding sequence from three non-placental mammals: the duckbill platypus, the short-beaked echidna, and the grey short-tailed opossum. *Dev Comp Immunol* 27(3): 247–256.
- Riegert P, Andersen R, Bumstead N, Doehring C, Dominguez-Steglich M, et al. (1996) The chicken beta 2-microglobulin gene is located on a non-major histocompatibility complex microchromosome: a small, G+C-rich gene with X and Y boxes in the promoter. *Proc Natl Acad Sci U S A* 93(3): 1243–1248.
- Welinder KG, Jespersen HM, Walther-Rasmussen J, Skjoldt K (1991) Amino acid sequences and structures of chicken and turkey  $\beta_2$ -microglobulin. *Mol Immunol* 28(1-2): 177–182.
- Western AH, Eckery DC, Demmer J, Juengel JL, McNatty KP, et al. (2003) Expression of the FcRn receptor (and) gene homologues in the intestine of suckling brushtail possum (*Trichosurus vulpecula*) pouch young. *Mol Immunol* 39(20): 707–717.
- Groves ML, Greenberg R (1982) Complete amino acid sequence of bovine  $\beta_2$ -microglobulin. *J Biol Chem* 257(5): 2619–2626.
- Ellis SA, Martin AJ (1993) Nucleotide sequence of horse  $\beta_2$ -microglobulin cDNA. *Immunogenetics* 38(5): 383.
- Tallmadge RL, Lear TL, Johnson AK, Guerin G, Millon LV, et al. (2003) Characterization of the  $\beta_2$ -microglobulin gene of the horse. *Immunogenetics* 54(10): 725–733.
- Daniel F, Morello D, Le Bail O, Chambon P, Cayre Y, et al. (1983) Structure and expression of the mouse beta 2-microglobulin gene isolated from somatic and non-expressing teratocarcinoma cells. *EMBO J* 2(7): 1061–1065.
- Cole T, Dickson PW, Esnard F, Averill S, Risbridger GP, et al. (1989) The cDNA structure and expression analysis of the genes for the cysteine proteinase inhibitor cystatin C and for  $\beta_2$ -microglobulin in rat brain. *Eur J Biochem* 186(1-2): 35–42.
- Gastinel LN, Valente G, Bjorkman PJ (1992) cDNA sequence of CHO-K1 hamster beta2-microglobulin. *Proc Natl Acad Sci U S A* 89(2): 638–642.
- Wolfe PB, Cebrá JJ (1980) The primary structure of guinea pig  $\beta_2$ -microglobulin. *Mol Immunol* 17(12): 1493–1505.
- Canavez FC, Ladasky JJ, Muniz JA, Seuanez HN, Parham P, et al. (1998)  $\beta_2$ -Microglobulin in neotropical primates (Platyrrhini). *Immunogenetics* 48(2): 133–140.

42. Canavez FC, Moreira MA, Simon F, Parham P, Seuanez HN (1999) Phylogenetic relationships of the Callitrichinae (Platyrrhini, primates) based on  $\beta$ 2-microglobulin DNA sequence. *Am J Primatol* 48(3): 225–236.
43. York IA, Grant EP, Dahl AM, Rock KL (2005) A mutant cell with a novel defect in MHC class I quality control. *J Immunol* 174(11): 6839–6846.
44. Gussow D, Rein R, Ginjaar I, Hochstenbach F, Seeman G, et al. (1987) The human  $\beta$ 2-microglobulin gene, primary structure and definition of the transcriptional unit. *J Immunol* 139: 3132–3138.
45. Milland J, Loveland BE, McKenzie IF (1993) Isolation of a clone for pig  $\beta$ 2-microglobulin cDNA. *Immunogenetics* 38(6): 464.
46. Gates FT, Coligan JE, Kindt TJ (1979) Complete amino acid sequence of rabbit  $\beta$ 2-microglobulin. *Biochemistry* 18(11): 2267–2272.