Molecular characterization of apolipoprotein A-I from the skin mucosa of *Cyprinus carpio*

Jolodar A.*

Received: October 2015 Accepted: May 2016

Abstract

Apolipoprotein A-I is the most abundant protein in *Cyprinus carpio* plasma that plays an important role in lipid transport and protection of the skin by means of its antimicrobial activity. A 527 bp cDNA fragment encoding C terminus part of apoA-I from the skin mucosa of common carp was isolated using RT-PCR. After GenBank database searching, a partial sequence containing a coding sequence (CDS) relating to this gene was found. Overlapping of the cDNA fragment with this CDS allowed us to obtain the full-length sequence including non-coding regions. This sequence has 1170bp including a polyA tail of 18 bp plus 45 and 354 bp at the 3'- and 5'-untranslatedregions, respectively. The complete sequence contained an open reading frame of 256 amino containing 5 amino acid propeptides with a predicted molecular mass of 29.967 kDa and theoretical pI of 6.13.The signal peptide of common carp apoA-I was predicted to have the most likely cleavage site between amino acid positions 17 and 18. Domain analysis of common carp apoA-I showed the conserved domain of Apolipoprotein A1/A4/E between amino acid resides 67 to 251. The similarity search indicated that common carp apoA-I matched apoA protein from the group of fish with 45-77% similarity, but showed relatively low levels of similarity to its mammalian counterparts (20-28%).It was shown that the secondary structure of *C. carpio* apoA-I consisted of α-helical predominantly amphipathic in nature and was characterized by the presence of thirteen conserved repeats.

Keywords: Apolipoprotein A-I, Common carp**,** *Cyprinus carpio*, Epidermal mucus, Full-length sequence

Department of Basic Sciences, Biochemistry and Molecular Biology Section, Faculty of Veterinary Medicine, Shahid Chamran University of Ahvaz, 61355-145, Ahvaz, Iran
*Corresponding author's Email:jolodara@scu.ac.ir

Introduction

Apolipoprotein A-I is the principal protein constituent of high-density lipoprotein (HDL), which is the most abundant plasma protein in the carp plasma. Besides the well-known role of apolipoprotein in lipid transport through the circulation system and its metabolism properties, much evidence has accumulated in the recent years that certain fish apolipoproteins have been shown to be potential immune modulators or antimicrobial proteins (Tada *et al.,* 1993). Preliminary studies detected the presence of apoA-I, in the skin and epidermal mucus of common carp *Cyprinus carpio* (Concha *et al.,* 2003). It was shown that apolipoproteins are expressed not only in the liver and intestine of *C. carpio* but also in the epidermis where they are apparently secreted to the mucus and display potent antimicrobial activity in the micro molar range against Gram positive and Gram negative bacteria, including some fish pathogens (Concha *et al.,* 2004). ApoA-Is from several teleost species, such as rainbow trout *Oncorhynchus mykiss*, shows potent antimicrobial activity *in vitro* against gram-positive and negative bacteria. The apoA-I responsible for this function is expressed in epidermal and mucosal barriers rather than in the liver, suggesting the tissue-specific role of apoA-I in fish (Villarroel *et al.,* 2007). These findings are very relevant, especially taking into consideration that the mucus layer is in direct contact with the aquatic environment and therefore constitutes the first and most extended defensive barrier against pathogen invasion.

Several reports have been published on the structures of apolipoproteins from mammals. However, little information about apolipoproteins has been obtained in lower vertebrates. Only a few fish apolipoprotein cDNAs encoding apoA-I have been reported, such as zebrafish *Danio rerio* (Babin *et al.,* 1997) and eel *Anguilla japonica* (Kondo *et al.,* 2001), rainbow trout *Oncorhynchus mykiss* ApoC-II (Shen *et al.,* 2000), rainbow trout ApoE (Durliat *et al.,* 2000), pufferfish *Takifugu rubripes* (Kondo *et al.,* 2005) and orange-spotted grouper *Epinephelus coioides* Apo-14 (Zhou *et al.,* 2005).

The aim of the present study was to characterize the full-length sequence from thee pidermal mucus of common carp apoA-I through RT-PCR and expressed sequence tag (EST) homology search.

Materials and methods

Fish specimens and tissue sampling

Common carp with an average body weigth of 1200 g were obtained from the Shahid Maleki Fish Culture Ponds located in the Khouzestan Province of Iran and maintained in an indoor aqvarium tank with running river water. Fish were adapted at 20±2°C for at least three weeks before they were killed. For tissue sampling from the skin of common carp, the mucus was previously removed by blotting the epithelia or mucosa with a tissue paper

and then cells were collected by scraping the epithelial layer with a sterile glass microscope slide and processed immediately or rapidly frozen in liquid nitrogen, and then stored at-80°C until they were used for RNA extraction.

RNA isolation

Total RNAs from the skin mucosa of common carp were prepared using the RNX plus solution (CinnaGen, Iran) according to the manufacturer's instructions. The purified total RNA was quantified by absorbance at 260 nm and used immediately or stored precipitated in ethanol at -80°C until use.

RT-PCR

Briefly, 12 µL (2 μg each) of skin total RNA was incubated with 0.5 μg of Oligo(dT)18 primer at 70°C, for 10 min followed by a brief centrifugation. The reaction was chilled on ice for a few minutes and then 1 µL Rnasin(CinnaGen, Iran), 1 µL dNTP mixture (120 µM of each nucleotide), 2.5 µL of 5 X enzyme buffer and 1 µL (200 U) of Moloney Murine Leukemia Virus (M-MulV) reverse transcriptase (CinnaGen, Iran) were added. The reaction was incubated at 42°C for 1h followed by a brief centrifugation and then inactivation of the enzyme by heating at100°C for 10 min. Reverse transcriptase was omitted in the tubes corresponding to the negative controls.Primer set was generated based on a coding sequence (CDS) (KF268349)and an EST clone (CA967348) for common carp apoA-I.RT-PCR reactions were carried out for cDNA template using standard reaction conditions. The reaction mixture (50 µL) contained 5 µL of the reverse transcription reaction, 0·2 µM of each primer, 250 µM of each dNTP and 1U of *Taq* DNA polymerase in a standard PCR buffer. The thermocycler was programmed as follows: Initial denaturation (94°C, 3 min), followed by 25 cycles (95°C for 40 s, annealing at 58°C for 50s, and extension at 72°C for 50s) and a final extension step at 72°C for 5 min. The amplification product was then electrophoresed on 1% (w/v) agarose gel and visualized by staining with ethidium bromide. DNA fragments were then extracted from the gel using the QIAquick Gel Extraction Kit (QiaGen, Iran) according to the manufacturer's instructions.

DNA sequencing and sequence analysis

The DNA was sequenced from both ends using a dideoxy termination method in an Applied Biosystems 373 DNA sequencer. The sequence was determined by using overlapping fragments. Primer sets were generated using Primer3 program (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). Each cDNA sequence was translated into the amino acid sequence using the Translate tool software available at the Expasy website (http://ca.expasy.org/tools/pi_tool.html). The predicted molecular mass and

theoretical pI value were estimated using ProtParam (http://www.expasy.org/tools/protparam.html). The putative signal peptides were identified using the SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalIP/). The secondary structure was predicted by using the PSIPRED Protein Structure Prediction Software which was from the websit (http://bioinf.cs.ucl.ac.uk/psipred/). A motif search was conducted using the Motif Search Software (http://pfam.sanger.ac.uk/search/sequence). The sequence alignments were performed using ClustalW1.8 program (Thomopson *et al*., 1994) and edited with the BOXSHADE software (http://www.ch.embnet.org/software/BOX_form.html). The computer program software was used to calculate the frequency codon usage for each amino acid http://www.bioinformatics.org/SMS/.

Results

Isolation and sequence analysis of C. carpio apoA-I

The results obtained by RT-PCR analysis confirmed the presence of apolipoprotein A-I in common carp skin since the expected 527 bp cDNA fragment was amplified from the skin. No amplification product was obtained in the control samples where reverse transcriptase was omitted (Fig. 1). When the amino acid sequences derived from the amplified product were searched against an amino acid sequence data base, the greatest homology was found for several fish apoA-Is. After database searching, a 774-bp CDS clone(KF268349)was found revealing an open reading frame (ORF) contiguous with the isolated cDNA fragment. The CDS clone has been deposited as direct submission in July 2014. This seqence includes only a coding region for protein without any characterization. The complete nucleotide sequence (including non-coding regions) was assembled from these two overlapping cDNA clones. The full-length cDNA sequence of *C. carpio* apoA-I containing both coding and non-coding region was shown in Fig. 2.

The region surrounding the first initiator sequence of *C. carpio* apoA-I matches the consensus eukaryotic initiation sequence, predicted to be the first ATG at position +1 (Fig. 2). The prediction of this initiation codon was made because it was the first ATG in the open reading frame followed by a hydrophobic leader sequence. There was Arg, a positive charge residue at position 2 after the first predicted Met in addition to a purine (Adenine) three bases before and one base after this initiator. By employing these features that are prerequisite for an initiation codon, it is reasonable to propose that this ATG is the actual initiation codon. This conservative ANNATGA feature of *C. carpio* mRNA at the translational start site is in agreement with the Kozak (Kozak, 1981; Kozak, 1986) initiation consensus.

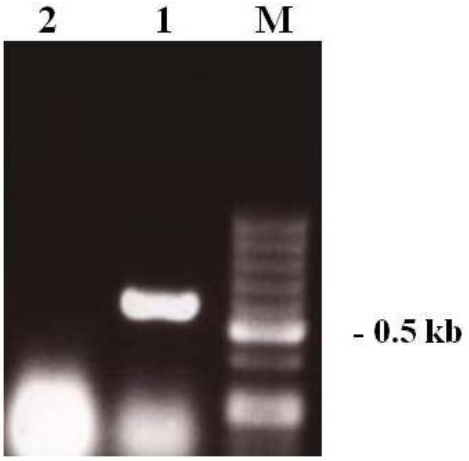


Figure 1: Agarose gel electrophoresis of RT-PCR products from apoA-I gene isolated from the skin mucosa of *Cyprinus carpio*. M: DNA size marker. Lane 1: RT-PCR amplification products. Lane2: negative control (water). Each lane was loaded with 8 µL of the total reaction.

```
                                                                  +1
-45  cggccgggagtctcttcccagaccagctacatcaacagatccatgATGAGGTTCGTAGTT
                                                          M  R  F  V  V
16   CTCGCCCTCGCTGTTTTGCTGGCAGGCTGCCAGGCCCGGTTCCTGCAGGACGAGCCGCCG
      L  A  L  A  V  L  L  A  G  C  Q  A  R  F  L  Q  D  E  P  P
77   TCGCAGGTGGAGCACGTGAAGTCTGCGCTCCAGCTTTACGCTGATCAGCTGAAGCAGGCG
      S  Q  V  E  H  V  K  S  A  L  Q  L  Y  A  D  Q  L  K  Q  A
137  GCGCACAAGAGCCTCACTCACCTCGACGACACAGAGTTCGCAGACTACAAGGAATTCCTG
      A  H  K  S  L  T  H  L  D  D  T  E  F  A  D  Y  K  E  F  L
197  GGCCAGTCTGTGGACAACCTCCACGGCTACTTTCAGAACGGCTTCCAAGCCATCACCCCA
      G  Q  S  V  D  N  L  H  G  Y  F  Q  N  G  F  Q  A  I  T  P
257  ATTGGTGACCAGGTGCTGGAGGCCACTAAAGACACACGCGAGAAGCTGGTCAAGGACGTG
      I  G  D  Q  V  L  E  A  T  K  D  T  R  E  K  L  V  K  D  V
317  GAAGAGCTCCGCAAGAAGATCGAGCCCATGCGCGCGGAGCTCAGGCAGGTGCTGGAGAAG
      E  E  L  R  K  K  I  E  P  M  R  A  E  L  R  Q  V  L  E  K
377  CACTTACAGGAGTACAGAGACGAGCTGAAGCCTTTCGTCGAGGAGTACCTGACCAAACAT
      H  L  Q  E  Y  R  D  E  L  K  P  F  V  E  E  Y  L  T  K  H
437  CAAAAGTTCCTGGAGGAGATGAGGATCAAGCTGGAGCCTGTGGTCAAGAGCTTGAGAGAG
      Q  K  F  L  E  E  M  R  I  K  L  E  P  V  V  K  S  L  R  E
497  AAGTTTGGACCCAACTGGGAGGAGACCAAGTCCAAGCTGATGCCCATCTTGGAGGCTGTG
      K  F  G  P  N  W  E  E  T  K  S  K  L  M  P  I  L  E  A  V
557  CGCGAGAAGGTGGCGGAGCATCTCCAGGACCTGAAGAAACTGCTGGAGCCCTACATGCAG
      R  E  K  V  A  E  H  L  Q  D  L  K  K  L  L  E  P  Y  M  Q
617  GATTATAGGGAGCAGATGGAGAAGGGAGCCCAGGAGTTCCGCCAGAGCGTGAAATCTGGA
      D  Y  R  E  Q  M  E  K  G  A  Q  E  F  R  Q  S  V  K  S  G
677  GAACTGAGGAAAAAGATGAACGAGTTGGGCGAGGAGGTGAAGCCTCACTTCGAGGCTATT
      E  L  R  K  K  M  N  E  L  G  E  E  V  K  P  H  F  E  A  I
737  TTCGCAGCCGTCCAAAAGGCCATTTACAAGCCGTAAaccgacccttttttacaccatctc
      F  A  A  V  Q  K  A  I  Y  K  P  .
797  cgcctcctttcttccttgccaccaggggccgagtaccaacacccgaaacatcgtcatccag
857  aaggcttttctcagagcttaactgcctttcactgaagcacacacacactaacacatgctc
917  acaacactcacatgcacttttgacactcacattcactctaatggccaaacacagctaaga
977  attaaagaacactgaaacctgttctttcaaaaatgcctagcaaaattccttttgttaagct
1037 gctttttgttaaaaactgattttcaatgttcagtgaaataaataatacaaataaaaactt
1097 tcatactatttcaaaaaaaaaaaaaaaaa
```

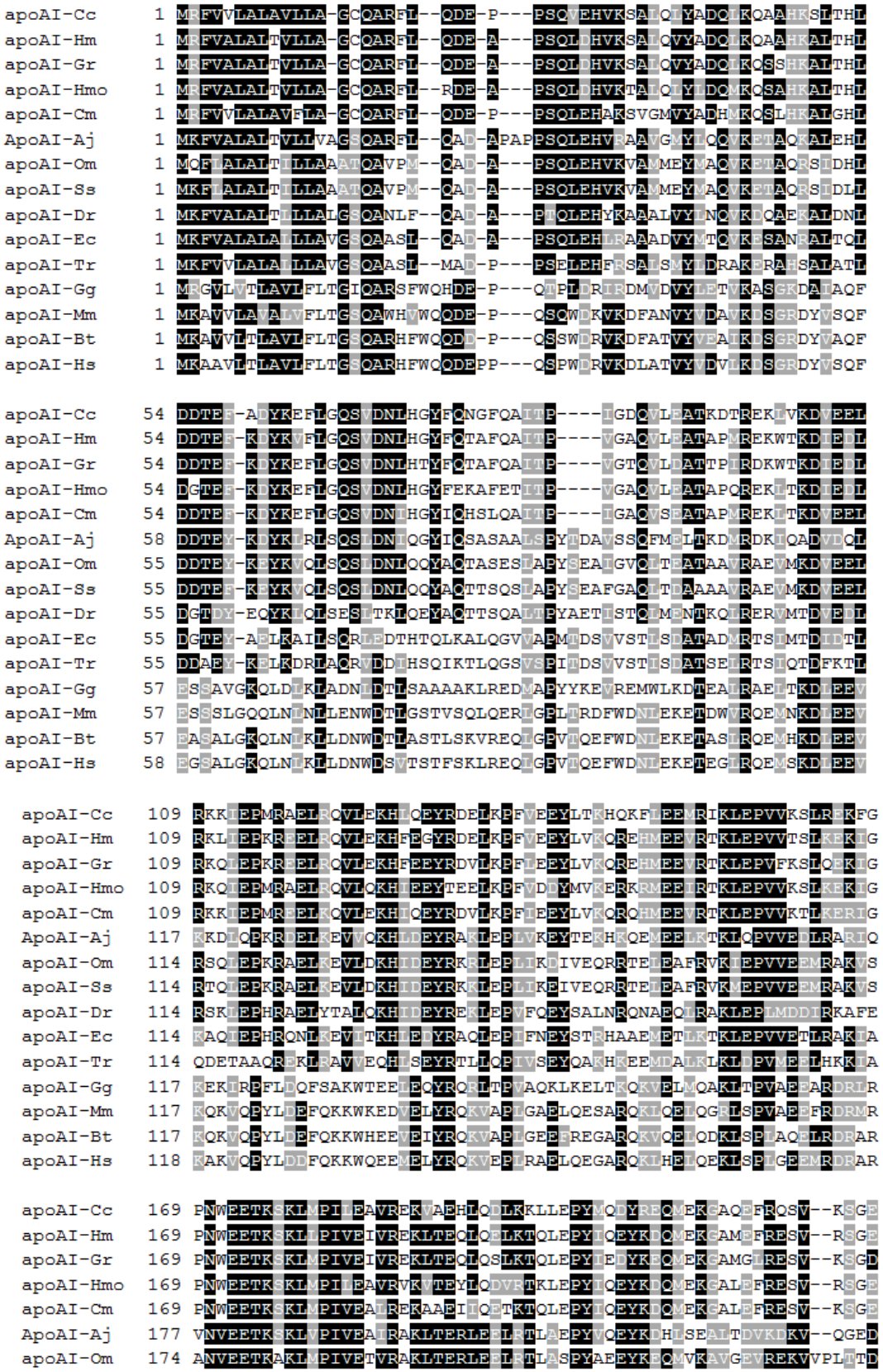
Figure 2: The complete nucleotide sequence of *Cyprinus carpio* apoA-I and its predicted primary structure. The amino acid sequence of the coding region is shown in one letter code below the nucleotide sequence. Amino acid number +1 is assigned to the first residue of the pre-enzyme. The predicted signal peptide is shown in bold letters. The position of the propeptides are underlined. The 5´- and 3´-untranslated regions show in lower cases. The presumed polyadenylation signals (aataaa) in the 3´-untranslated region are underlined and bold.

As shown in Fig. 2, the full-length cDNA of *C. carpio*apoA-Icontained1170 bp with a single open reading frame of 768 bp flanked by 45 and 354bp at 5'- and 3'-untraslated regions, respectively. Two polyadenylylation consensus signals (AATAAA) were found 41 and 54 bp upstream from the first residue of the poly(A) tail. The cDNA possessed two copies of the putative polyadenylation signals suggesting the possibility of processing their mRNA species with differential lengths. The coding region encodes a polypeptide of 256 amino acids with a calculated molecular mass of 29.967 kDa and theoretical pI of 6.13. The predicted molecular masses of apoA-I polypeptide agrees closely with the apoAI-1 and apoAI-2 of *Hemibarbus mylodon* (Kim *et al.*, 2009) in the range of 29.97 to 30.62 kDa, but the theoretical pI value of *C. carpio* apoA-I (6.13) was higher than its counterparts (5.78-5.25).

Signal peptide analysis of C. carpio apoA-I

An analysis using the signal peptide software indicated that *C. carpio* apoA-I posses a single peptide domain structure. The signal sequence is involved in sequestration of the protein to the rough endoplasmic reticulum as a prelude to extracellular secretion. The N-terminal extension in the predicted amino acid sequence of *C. carpio* apoA-I has several features which are characteristic of signal peptides found in the precursor of most secreted proteins. Computer analysis for the hydropathy profile of the deduced amino acid sequence at the N-terminus region indicated that one stretch of about 17 hydrophobic amino acids residue directly follows the initiation codon (Fig. 2). Peptide signal cleavage site can be investigated by application of " (-3, -1)-rule"(Von Heijne, 1986)· Based on this rule, in addition to high overall hydrophobicity, the presence of an amino acid with a short side chain in position -1, and a lack of an aromatic, charged or polar residue at position -3 are necessary. Moreover, it has been suggested that proline must not be present in position -3 through +1. In addition to these, the proteolytic cleavage site must be located near the point where the hydrophobic index drops dramatically and enters the hydrophilic range. Therefore, the signal peptide of *C. carpio* apoA-I was predicted to have the most likely cleavage site between amino acid positions 17 and 18 with a high probability score (0.991) and the Arg at position 18 was assumed to represent the start of the mature protein.

The pro-segment in *C. carpio* apoA-I appeared to be only 4 amino acids (between positions 18-21) long in accordance with other fish apoA-Is (Fig. 3).


apoAI-Cc
apoAI-Hm
apoAI-Gr
apoAI-Hmo
apoAI-Cm
ApoAI-Aj
apoAI-Om
apoAI-Ss
apoAI-Dr
apoAI-Ec
apoAI-Tr
apoAI-Gg
apoAI-Mm
apoAI-Bt
apoAI-Hs

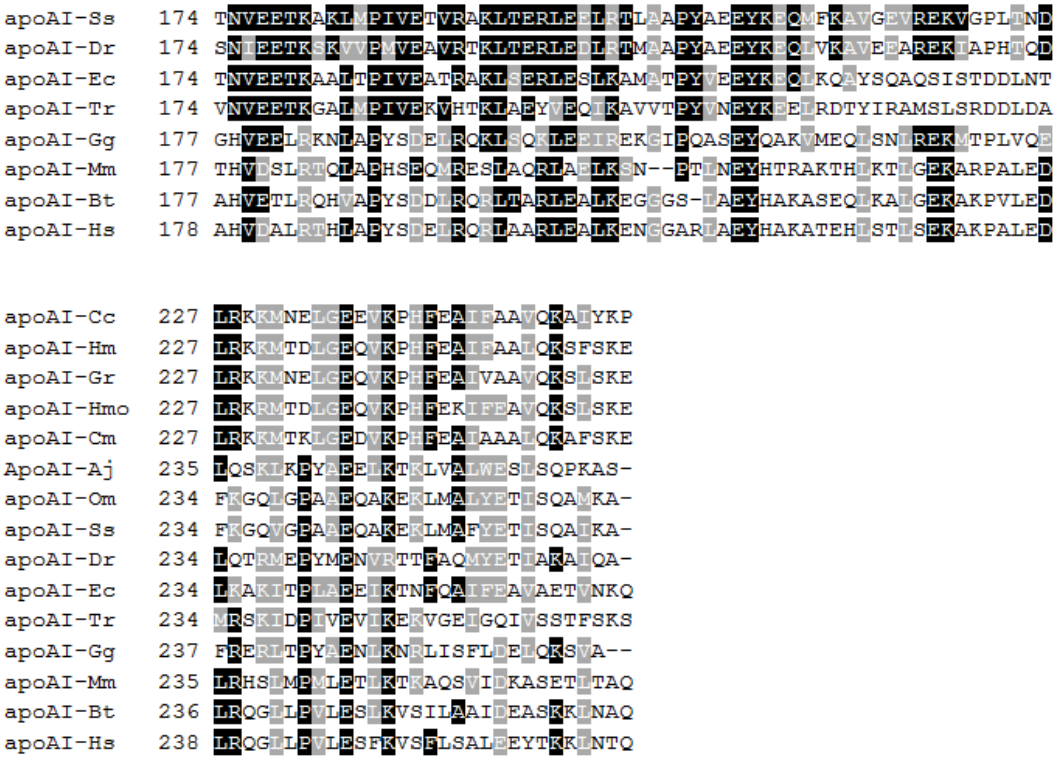


Figure 3: Multiple sequence alignment of *Cyprinus carpio* apoA-I. The *C. carpio* sequence is aligned with apoA-I-1 from Korean doty barbel *Hemibarbus mylodon* (Hm; ACI15889), apoA-I from silver carp *Hypophthalmichthys molitrix* (Hmo; ADF97611), mud carp *Cirrhinus molitorella* (Cm; ACY82518), Gobiocypris rarus (Gr; ABY47600), 28kDa-2 apolipoprotein from Japanese eel *Anguilla japonica* (Aj; BAB40960), apoA-I-2 from rainbow trout *Oncorhynchus mykiss* (Om; NP_001117720), apoA-I from orange-spotted grouper *Epinephelus coioides* (Ec; ACM48181), ApoA-I from Atlantic Salmon *Salmo* salar (Ss; NP_001134612), apoA-I from Fugu rubripes *Takifugu rubripes* (Tr; NP_001072100), apoA-I from zebrafish *Danio rerio* (Dr; NP_571203), apoA-I precursor from chicken *Gallus gallus* (Gg; AAA48592), ApoA-I from house mouse *Mus musculus* (Mm; CAA45560), apoA-I precursor from cattle *Bos taurus* (Bt; AAA30381) and ApoA-I from Human *Homo sapiens* (Hs; AAH05380). Shading indicates identity (black) or conservative substitutions (grey) relative to *Cyprinus carpio*. Gaps inserted to optimize alignments are indicated by dashes.

In the *C. carpio* protein, cleavage may occur following the conserved Gln residue to give a mature 235-residue apoA-I protein with an N-terminal aspartate. It should be noted that Mammalian and chicken apoA-Is have been reported to contain a 6 amino acid pro-segment. It means; the comparison with derived protein sequences of common *C. carpio* apoA-I with other mammals and chicken apoA-Is indicates that in those animals two additional codons have led to the increasing of 2 amino acids (Trp and Gln) between residues 20 and 21 (based on the *C. carpio* apoA-I numbering).Posttranslational processing has been shown to occur in mammalian proteins (Cheung and Chang, 1983) following the conserved Gln-Gln

neutral dipeptide (instead of one Gln in *C. carpio* protein) to give a mature protein.

Patterns of codon usage in C. carpio apoA-I

It is interesting to note that bias of codon usage in gene sequences used by an organism is non-random and differences in codon usage pattern not only exist between species but also may be found within a species. The results of the codon usage in apoAI presented in Table 1 shows thatthe codon usage for the various amino acids in *C. carpio* apoA-Iis extremely skewed. This is especially evident when some of the more frequently represented amino acids are examined. For example, there are 31 residues of Leu in the sequence.

Table 1: Codon usage table for *Cyprinus carpio* apoA-I protein.

Amino acids	Codon Usage	Total #	*C. carpio* apoAI %	Amino acids	Codon Usage	Total #	*C. carpio* apoAI %
Ala	GCG	5	26	Asp	GAT	2	17
	GCA	3	16		GAC	10	83
	GCT	4	21	Glu	GAG	30	91
	GCC	7	37		GAA	3	9
Cys	TGT	0	0	Phe	TTT	2	17
	TGC	1	100		TTC	10	83
Gly	GGG	0	0	His	CAT	2	25
	GGA	3	33		CAC	6	75
	GGT	1	11	Lys	AAG	23	82
	GGC	5	56		AAA	5	18
Ile	ATA	0	0	Leu	TTG	4	13
	ATT	3	43		TTA	1	3
	ATC	4	57		CTG	16	52
Met	ATG	7	100		CTA	0	0
Pro	CCG	3	27		CTT	1	3
	CCA	1	9		CTC	9	29
	CCT	3	27	Asn	AAT	0	0
	CCC	4	36		AAC	4	100
Gln	CAG	16	84	Arg	AGG	5	38
	CAA	3	16		AGA	2	15
Ser	AGT	0	0		CGG	1	8
	AGC	3	38		CGA	0	0
	TCG	1	13		CGT	0	0
	TCA	0	0		CGC	5	38
	TCT	3	38	Thr	ACG	0	0
	TCC	1	13		ACA	2	29
Val	GTG	11	61		ACT	2	29
	GTA	1	6		ACC	3	43
	GTT	2	11	Trp	TGG	1	100
	GTC	4	22	End	TGA	0	0
Tyr	TAT	1	13		TAG	0	0
	TAC	7	88		TAA	1	100

Sixteen of them (52%) are encoded by CTG whereas none of them are encoded by CTA. Such extreme bias in the preference for specific codons is also evident in some other amino acids including Phe, Gln, Lys, Asp and Glu. Morover, 2 amino acids with more than one codon (Asp and Asn) are found to be coded by only one specific triplet.

Sequence alignment of C. carpio apoA-I.

Analignment of the deduced amino acid sequences of *C. carpio* apoA-Iwith the group of fish and mammalian apoA-Is is shown in Fig.3. The *C. carpio* sequence is aligned apoA-I-1 from Korean doty barbel *Hemibarbus mylodon* (ACI15889) (Kim*et al.,* 2009), apoA-I from silver carp*Hypophthalmichthys molitrix* (ADF97611), mud carp*Cirrhinus molitorella* (ACY82518), and*Gobiocypris rarus* (ABY47600), 28kDa-2 apolipoprotein from Japanese eel *Anguilla japonica* (BAB40960), ApoA-I-2 precursor from rainbow smelt *Osmerus mordax* (ACO09807), apoA-I-2 from rainbow trout*Oncorhynchus mykiss* (NP_001117720), apoA-I from orange-spotted grouper *Epinephelus coioides* (ACM48181), ApoA-I from Atlantic Salmon *Salmosalar* (NP_001134612), apoA-I from Fugu rubripes *Takifugu rubripes* (NP_001072100), apoA-I from zebrafish *Danio rerio* (NP_571203), apoA-I precursor from chicken *Gallus gallus* (AAA48592), ApoA-I from house mouse *Mus musculus* (CAA45560), apoA-I precursor from cattle *Bos taurus* (AAA30381) and ApoA-I from Human *Homo sapiens* (AAH05380). The differences in amino acid sequence were scattered throughout the sequence. The amino acid sequence similarity with apolipoprotein AIb1 (AII80532) and apolipoprotein AIb2 (AII80537) from *C. carpio* were 98% and 97%, respectively. The deduced amino acid sequence is most similar to fish apoA-I with 76% identity with Korean doty barbel, *Hemibarbus mylodon*, 74% identity with *Gobiocypris rarus*,46% identity with Japanese eel, *Anguilla japonica*, 45% identity with zebrafish, *Danio rerio* and it is similar to mammalian apoA-I with 25% identity with human, *homosapiens* and 23% identity with mouse, *Mus musculus*.Gaps were introduced for maximum alignment of these sequences. The *C. Carpio* apoA-I was aligned along with the corresponding mammalian sequences since these apolipoproteins have been revealed to arise from the common ancestor gene (Luo and Li, 1986).

Sequence characteristics of C. carpio apoA-I

Domain analysis of *C. carpio* apoA-Ishowed e-value of 3.1e-35 with conserved domain of Apolipoprotein A1/A4/E domain (pfam01442)between amino acid resides 67 to 251. Multiple sequence alignment shown in Fig. 3illustrates the most conserved amino acid sequence scattered among the

entire sequence. Although apo A-I is modestly conserved among species, both at the level of nucleotide and amino acid sequence, the presence of thirteen conserved repeats predicted to form amphipathic α-helical secondary structures are quite conserved (Luo and Li, 1986; Li *et al.*, 1988). Using the PSIPRED Protein Structure Prediction Software we obtained a prediction of 89.4% and 10.6% of *a*-helical and random coils, respectively for the full-length coding sequence of *C. carpio* apoA-I (Fig. 4).

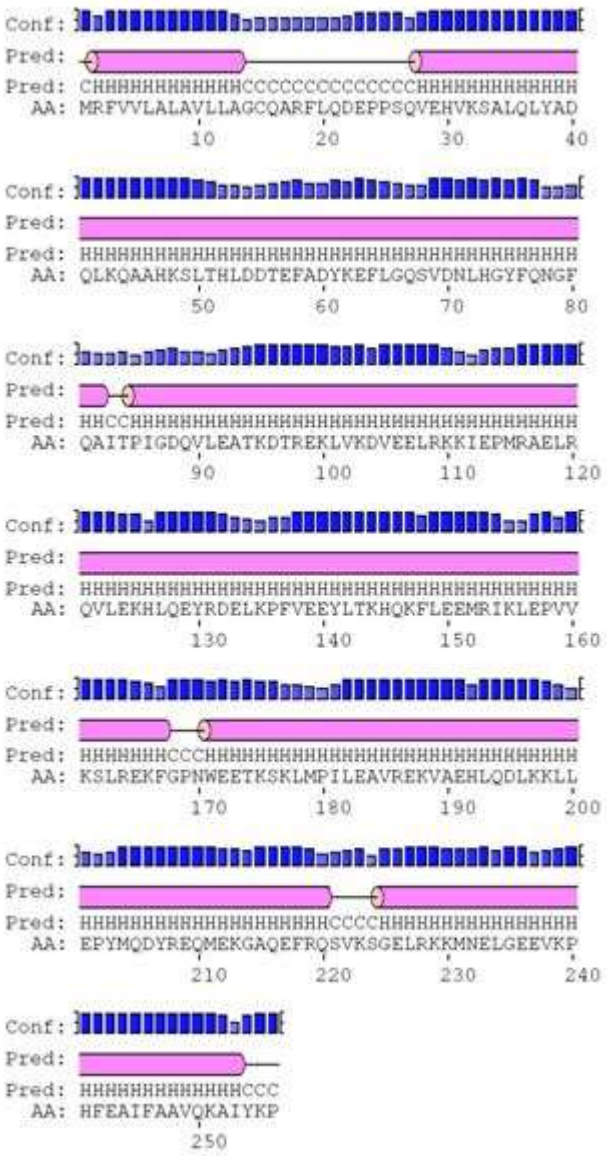


Figure 4: Schematic representation for the secondary structure of the *Cyprinus carpio* apoA-I. The amino acid sequences predicted to form amphipathic α-helical (H) and random coils (C) are indicated.

Similar values were found utilizing other secondary structure prediction programs.

The sequence comparison followed by the tracking of its repeating units referring to that of *Takifugu rubripes* apoA-I (Kondo *et al.*, 2005) indicates that the first part of *C. carpio* apoA-Iconstituted of 32 amino acids containing units of 1-3 (Fig. 5). The amino acid residues in this area for both*C. carpio*and *T. Rubripes*sequences are the same. The second part of *C. carpio* sequence includes internal repeats 4 (22 amino acid residues) and 5 (18 amino acid residues). Deletions of four and three amino acid residues were observed in internal repeat 5 and 12 of *C. Carpio*, whereas those repeats of *Takifugu rubripes*sequence consisted of 22 and 11 amino acids, respectively. In both sequences, the number of amino acid for five more internal repeats including repeats 7, 8, 9, 10 and 13 (22 amino acid residues) were the same. Proline was situated in the first position of internal repeats 5, 8, 9, 10 and 11in the both sequences.

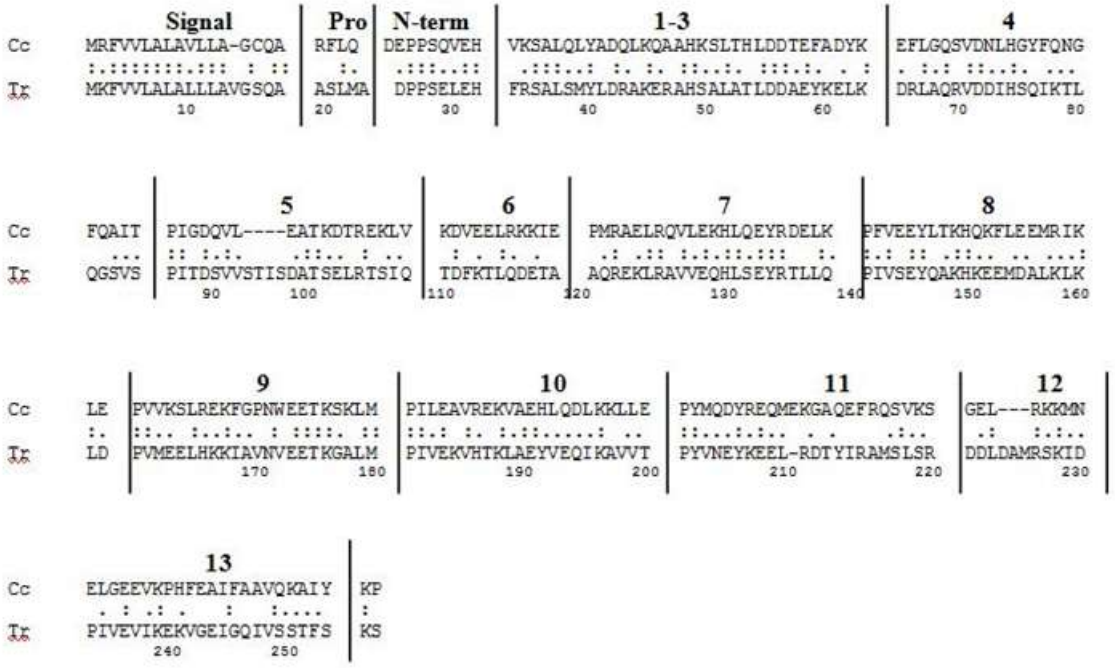


Figure 5: Alignments of amino acid sequences of *Cyprinus Carpio* apoA-I (Cc) with that of *Takifugu rubripes* apoA-I (Tr) counterpart. Repeated structures are separated by vertical bars. Signal peptides (Signal), propeptides (pro), and N-terminal (N-term) regions are also separated by vertical bars. Numbers of the internal repeats are indicated above the top sequence. Gaps represented by dashes were introduced to maximize the alignment. Double dots indicate identities and single dots indicate conservative substitutions between *Cyprinus carpio* and *Takifugu rubripes* sequences.

Discussion

The synthesis and secretion of apoA-I from the skin of the chicken is the only report which has been published by Tarugi and co-workers (1991). There are no other previous studies on this matter in lower vertebrates. The characterization of apolipoprotein family members from the fish species is currently limited to a very few. In fact, the physiological relevance of apoA-I expression in the fish skin has probably been underestimated.

It is unlikely that apolipoproteins are abundantly synthesized and secreted in the fish epidermis only to take part in the local homeostasis of cholesterol, particularly taking into consideration that substantial amounts of this protein would be lost continuously from the mucus coat to the water surroundings. In a previous study it was shown that HDL and its major apolipoproteins, ApoA-I and ApoA-II display antimicrobial activity in the common carp *C. carpio* (Concha *et al.,* 2003). Therefore, it has been proposed that apolipoproteins could also play an important role in the protection of the skin of the fish by means of its antimicrobial activity. The use of genetic information for fish apolipoproteins is considered important not only to develop a better understanding of lipid metabolism but also to figure out the importance of the role of this protein in defensive functions in the teleost family.

In this study, *C. carpio* apoA-I with low identity to mammalian apoA-Is shows a higher signal peptide probability (0.912) than other apolipoprotein family members (except for mud carp*Cirrhinus molitorella* apoA-I which is 0.929). Therefore, it can be assured that the predicted signal peptide in common carp apoA-I is actually cleaved between positions 17 and 18. However, other apolipoproteins family members, mammalian apoM has been reported to retain an uncleaved N-terminal signal peptide that may serve as a phospholipid anchor into a single lipid layer of high-density lipoprotein (Axler *et al.,* 2008). *C. Carpio* preproapolipoprotein A-I had predicted molecular weights of 29.967, and predicted isoelectric points of 6.13**,** respectively. These values are consistent with the results of recent studies showing that an apolipoprotein of *H. mylodon* which had a molecular weight of approximately 29.970 (6, 8, 9) and migrated in the pH 5.78 region of an electrofocusing gel (Keun-Yong *et al.,* 2009).

Comparison of *C. carpio*cDNA sequence against GenBank database including BLASTx and tBALSTN options clearly confirmed *C. carpio* apoA-I to be specific to fish. Based on the database search, apoA-I sequences was identified in several species belonging to various organisms indicating that this apolipoprotein gene is widely distributed in the teleost family. As expected, *C. carpio* apoA-I sequence isolated in this study shows higher levels of identity with their corresponding orthologs from teleosts

than from terrestrial vertebrates. This is indicated that teleosts are genetically separated from terrestrial animals. Multiple alignments of apoA-I deduced amino acid sequence shows that the primary structure of this protein is poorly conserved among vertebrates; however, the predicted secondary structure of these proteins is surprisingly similar (high content of amphipathic a-helix). In fact, their overall secondary structures were generally conserved among vertebrates despite modest or low levels of amino acid sequence identity across taxa. In particular, *C. carpio* apoA-I, was well conserved in comparison to previously reported vertebrate counterparts in terms of the characteristic tandem repeat units forming amphipathic helices (Kondo *et al.,* 2005). For example, the comparison of *C. carpio* and human apoA-I amino acid sequences revealed a high degree of resemblance in terms of the characteristic tandem repeat units forming amphipathic helices despite the large evolutionary distance between the two species. Many previous studies suggested that fish apolipoproteins such as apo-14 kDa (Keun-Yong *et al.,* 2009) have experienced a different evolutionary history from those of mammalian counterparts. However, the similarity of the repeat pattern in *C. carpio* and human apoA-Is suggests that the different apolipoprotein genes arose from a common ancestral gene prior to the teleost fish-mammal split, some 400 million years ago. Therefore, we speculated that in spite of the low sequence similarities that exist between mammalian and teleost apolipoprotein A-I, its conserved overall structure would be responsible for maintaining these defensive functions during evolution.

The results of this study will be useful as fundamental baseline data to gain a better understanding of the specific roles of apoA-I in *C. carpio* and perhaps other teleost fish skin as an important innate immunity effector.

Acknowledgments

This study was financially supported by a research grant number 636 from the Vice President ofResearch Affairs Office at the Shahid Chamran University of Ahvaz, Ahvaz, Iran.

References

Axler, O., Ahnstrom, J. and Dahlback, B., 2008. Apolipoprotein M associates to lipoproteins through its retained signal peptide. *FEBS Letters*, 582, 826-828.

Babin, P.J., Thisse, C., Durliat, M., Andre, M., Akimenko, M.A. and Thisse, B., 1997. Both apolipoprotein E and A-I genes are present in a nonmammalian vertebrate and are highly expressed during embryonic development. Proc. Natl. Acad. Sci. USA 94, 8622c. Nat

Cheung, P. and Chan, L., 1983. Nucleotide sequence of cloned cDNA of human apolipoprotein A-I.

Nucleic Acids Research, 11, 3703-3715

Concha, M.I., Molina, S., Oyarzun, C., Villanueva, J. and Amthauer, R., 2003. Local expression of apolipoprotein A-I gene and a possible role for HDL in primary defence in the carp skin. *Fish and Shellfish Immunology*, 14, 259-73.

Concha, M.I., Smith, V.J., Castro, K., Bastías, A., Romero, A. and Amthauer, R.J., 2004. Apolipoproteins A-I and A-II are potentially important effectors of innate immunity in the teleost fish *Cyprinus carpio*. *European Journal of Biochemistry*, 271, 2984-2990.

Durliat, M., Andrea, M. and Babin, P.J., 2000. Conserved protein motifs and structural organization of a fish gene homologous to mammalian apolipoprotein. *European Journal of Biochemistry*, 267, 549-559

Keun-Yong, K., Young Sun, C., In-Chul, B. and Yoon Kwon, N., 2009. Isolation and characterization of the apolipoprotein multigene family in *Hemibarbus mylodon* (Teleostei: Cypriniformes). *Comparative Biochemistry and Physiology*, 152, 38-46.

Kim, K.Y., ChoY, S., Bang, I.C. andNam,Y.K., 2009. Isolation and characterization of the apolipoprotein multigene family in *Hemibarbus mylodon* (Teleostei: Cypriniformes). *Journal of Biochemistry and Molecular Biolology*, 152, 38-46.

Kondo, H., Kawazoe, I., Nakaya, M., Kikuchi, K., Aida, K. and Watabe, S., 2001. The novel sequences of major plasma apolipoproteins in the eel *Anguilla japonica*. *Biochimica et Biophysica Acta*, 1531, 132-142

Kondo, H., Morinaga, K., Misaki, R., Nakaya, M. and Watabe, S., 2005. Characterization of the pufferfish *Takifugu rubripes* apolipoprotein multigene family. *Gene*, 346, 257-266.

Kozak, M., 1981. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Research*, 12, 857-872.

Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codonthatmodulate translation by eukaryotic ribosomes. *Cell*, 44**,** 283-292.

Li, W.H., Tanimura, M., Luo, C.C., Datta, S. and Chan, L., 1988. The apolipoprotein multigene family: biosynthesis, structure, structure–function relationships, and evolution. *The Journal of Lipid Research*, 29, 245- 271.

Luo, C.C. and Li, W.H., 1986. Structure and evolution of the apolipoprotein multigene family. *Journal of Molecular Biology*, 187, 325-340.

Shen, Y., Lindberg, A. and Olivecrona, G., 2000. Apolipoprotein CII from rainbow trout (*Oncorhynchus mykiss*) is functionally active but structurally very different from mammalian

apolipoprotein CII. *Gene*, 254, 189-198

Tada, N., Sakamoto, T., Kamgami, A., Mochizuki, K. and Kurosaka, K., 1993. Antimicrobial activity of lipoprotein particles containing apolipoprotein A-I. *Molecular and Cellular Biochemistry*, 119, 171-178.

Tarugi, P., Albertazzi, L., Nicolini, S., Ottaviani, E. and Calandra, S., 1991. Synthesis and secretion of apolipoprotein A-I by chick skin. *Journal of Molecular Biology*, 266, 7714-7720.

Thompson, J.D., Higgins, D.G. and Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22, 4673-4680.

Villarroel, F., Bastías, A., Casado, A., Amthauer, R. and Concha, M.I., 2007. Apolipoprotein A-I, an antimicrobial protein in *Oncorhynchus mykiss*: evaluation of its expression in primary defense barriers and plasma levels in sick and healthy fish. *Fish and Shellfish Immunology*, 23, 197-209.

Von Heijne, G., 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research*, 14**,** 4683-4690

Zhou, L., Wang, Y., Yao, B., Li, C.J., Ji, G.D. and Gui, J.F., 2005. Molecular cloning and expression pattern of 14 kDa apolipoprotein in orange-spotted grouper, *Epinephelus coioides*. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 142, 432-437