African Journal of Biotechnology Vol. 10(43), pp. 8669-8674, 10 August, 2011
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.871
ISSN 1684–5315 © 2011 Academic Journals

*Full Length Research Paper*

# cDNA cloning and primary structure analysis of invariant chain in Chinese Pengze crucian carp

**Xuelan Liu, Weiyi Yu*, Jinnian Li, Fangfang Chen, Shengjie Liu, Chao Wu and Jiaping Xu**

Key Laboratory of Zoonoses of Anhui Province, Anhui Agricultural University, Hefei, 230036, China.

Accepted 17 June, 2011

**Invariant chain (Ii) plays an important role in MHC class II molecules assembly and exogenous peptide presentation in vertebrates. Although mammalian Ii has been extensively studied, less attention is paid to its fish counterpart. In this study, in order to understand the structure and biological function of Chinese Pengze crucian carp (*Carassius auratus*) Ii, we first cloned Ii gene (*Caau-Ii*) using RT PCR and rapid amplification of cDNA end techniques. Firstly, a conserved DNA fragment was amplified using a pair of degenerate primers from fish's head kidney and spleen total RNA. Then specific primers for RACE were designed to amplify the 5' and 3' ends of *Caau-Ii*. Finally, the full-length cDNA of *Caau-Ii* was amplified, sequenced and analyzed and the results showed that it was 1063 bp in length and contained an open reading frame of 696 bp encoding *Caau-Ii* precursor protein. Amino acid sequence alignment showed that (1) *Caau-Ii* had high homology with that of carp and zebrafish (identify >70%) and low homology with mammalian Ii (30% < identity <40%); (2) *Caau-Ii* had similar domains seen in mammalian Ii, including transmembrane domain, CLIP (class II-associated Ii-derived peptide) and TRIM (trimerization) domain.**

**Key words:** Invariant chain, *Carassius auratus*, RACE, cloning, domain analysis.

## INTRODUCTION

Invariant chain (Ii), also known as leukocyte differentiation antigen CD74, is a type II transmembrane protein that associates with the MHC class II molecule heterodimer formed by an alpha chain (α) and a beta chain (β), and so named because it is non-polymorphic. In mammals, Ii plays an important auxiliary role in the correct folding, assembly and exogenous antigen presentation of MHC class II molecule (Anderson and Miller, 1992; Serwe et al., 1997; Neumann and Koch, 2005; Holst et al., 2008). Ii aids MHC class II molecule in presenting antigen through its CLIP fragment, which occupies the groove formed between α and β chains of MHC class II molecule and thus, prevents the binding of endogenous peptides to it (Cresswell, 1996). Exogenous peptides will finally enter the groove by taking the place of CLIP and then be presented to cell surface (Robertson, 1998). In addition, recent studies showed that Ii may be used as a signal molecule that regulates the maturation of B cells (Matza
.et al., 2003) and the migration of dendritic cells (Faure-André et al., 2008) and that its loss or over expression may result in autoimmune diseases (Ansari, 1993; Bryan et al., 2008).

The mechanism of fish specific humoral immunity is not fully defined. In particular, how pathogen antigens are presented in fish remains under discussion. The MHCs of fish, like those in higher mammals are the major genes responsible for fish disease resistance, and related studies have been conducted in a variety of fish species (Cardenas et al., 2010; Yu et al., 2010). However, there are much fewer reports on Ii, an important chaperon of MHC class II molecules. Moreover, crucian carp is a main economical freshwater fish in China. In this study, using RACE techniques, we cloned for the first time Ii gene (*Caau-Ii*) from head kidney and spleen of crucian carp and used multiple bioinformatic tools to analyze its predicted protein structures and domains, laying the foundation for further studying the interaction between Ii and MHC class II molecules as well as the mechanism of antigen presentation in crucian carp.

*Corresponding author. E-mail: liuxuel03@yahoo.com.cn or yuweiyi@ahau.edu.cn. Tel: +86 551 5786891. Fax: +86 551 5785027.

**Table 1.** Primers for *Caau-li* cloning.

| Primer name | Primer sequence |
|---|---|
| li–$F_1$ | 5' GACAGGCGCTGACCGCAKACA 3' |
| li–$R_1$ | 5' GCAGTAYCCSGKKCTGTGCCA 3' |
| 3' RACE GSP1 | 5' TGATGCCGAAGCAGATACAGC 3' |
| 3′ RACE outer primer | 5' TACCGTCGTTCCACTAGTGATTT 3' |
| 3′ RACE inner primer | 5' CGCGGATCCTCCACTAGTGATTTCACTATAGG 3' |
| 5' RACE GSP1 | 5' ACACTGAGGCTTGAAGAACCC 3' |
| 5' RACE GSP2 | 5' TGGTCGCAGTTTAGTGAGAGGAATG 3' |
| 5′RACE outer primer | 5' CATGGCTACATGCTGACAGCCTA 3' |
| 5′RACE inner primer | 5'CGCGGATCCACAGCCTACTGATGATCAGTCGATG3' |
| Full-length li-F | 5' AGAGACGTAAACACACAGAGACAC 3' |
| Full-length li-R | 5' TCATCCATGCTACCTCTAACAACA 3' |

## MATERIALS AND METHODS

### Design of degenerate primers

Based on li gene sequences from the same family such as carp and zebrafish, a pair of degenerate primers, li–$R_1$ and li–$F_1$, was designed to amplify a conserved segment of *Caau-li*. The primers design is shown in Table 1.

### Extraction of total RNA and amplification of a conserved segment of Caau-li

Total RNA was extracted from head kidney and spleen of Chinese Pengze crucian carp using Trizol kit (TaKaRa) following manufacturer's instructions. First-strand cDNA was synthesized using the method described in the reagent kit and then used to perform PCR with specific primers. The PCR conditions were as follows: initial denaturation at 94°C for 4 min; 30 cycles of 94°C for 1 min, 58°C for 1 min and 72°C for 1.5 min and a final extension at 72°C for 7 min. PCR products were analyzed by using 1% agarose gel electrophoresis; gel was recovered using reagent kit and ligated with pMD18-T vector at 4°C overnight. The resulting mixture was used to transform competent *E. coli* Rosseta (DE3) cells, which were then cultured overnight. Positive clones, identified by using specific PCR primers, were sent to Sangon Biotech (Shanghai) Co., Ltd for sequencing.

### RACE amplification

According to the sequence of the conserved segment, forward gene specific primer 1 for 3' RACE (3' GSP1) was designed. With total RNA as template and 3′RACE adaptor as the primer, first-strand cDNA was synthesized by a reverse transcriptase. Using 3' GSP1 as well as 3′ RACE outer primer and 3′RACE inner primer contained in the reagent kit (TaKaRa), the second PCR was performed according to a minor modification of the method described in the kit. The amplified 3' fragment was cloned and sequenced using the same method described earlier.

According to the sequences of the conserved segment and the amplified 3' fragment, reverse outer gene specific primer 1(GSP1) and inner specific gene primer 2(GSP2) for 5' RACE were designed. With total RNA as template, dephosphorylation, decapping, ligation and reverse transcription were performed according to the methods described in the reagent kit (TaKaRa).

Outer PCR was performed with GSP1 and 5′ RACE outer primer and inner PCR with GSP2 and 5′ RACE inner primer. The amplified 5' fragment was cloned and sequenced using the same method described earlier. Primers used in 5' and 3' RACE are shown in Table 1.

### Cloning and identification of full-length *Caau-li*

Based on the sequences of amplified 5' and 3' ends, a pair of primers, full-length li-F and full-length li-R, was designed to amplify the full-length cDNA. The amplified cDNA was cloned and sequenced. The cDNA sequence was subjected to BLAST search against sequence database to determine whether the cloned sequence is full-length *Caau-li*.

### Functional domain analysis of *Caau-li*

DNAStar software was used to analyze amino acid composition of *Caau-li,* TMHMM-2 (http://genome.cbs.dtu/services/TMHMM-2) to detect putative transmembrane domain, SignalP3.0 (http://genome.cbs.dtu/services/SignalP3.0) to detect putative signal peptide, and SWISS-MODEL (http://www.expasy.ch/swissmod/SWISS-MODEL.html) to predict conformation.

## RESULTS

### Amplification of a conserved segment 5′ and 3′ ends of *Caau-li*

Using the primers for a conserved segment of *Caau-li*, after RT-PCR amplification, we obtained a gene fragment of 490 bp (Figure 1, lane 3). BLAST search against NCBI database showed that this fragment shared high homology with animals' li. Based on this, gene specific primers for 5' and 3' ends were designed. Two gene fragments, 411 bp (Figure 1, lane 2) and 610 bp (Figure 1, lane 4), were then amplified by using 5' and 3' RACE techniques, respectively.

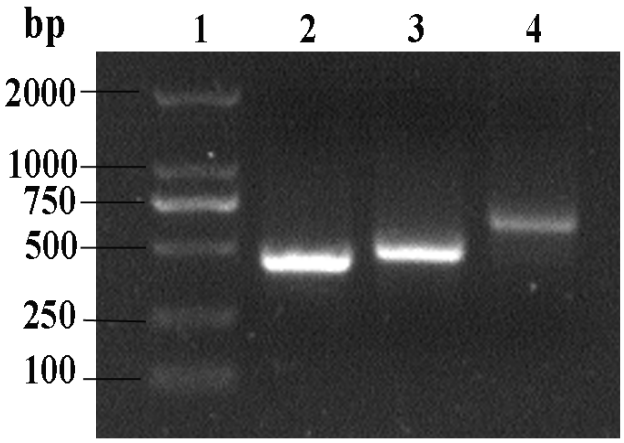


**Figure 1.** Identification of the products of RT-PCR, 3' RACE and 5' RACE. 1, DL2000 DNA marker; 2 to 4, samples from the product of 5' RACE, RT-PCR and 3' RACE, respectively.

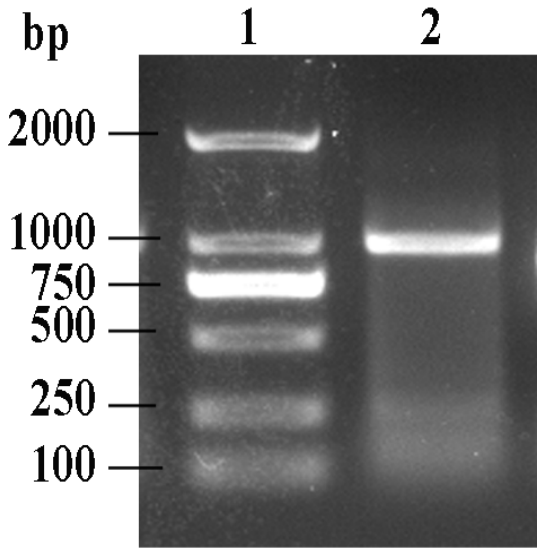


**Figure 2.** Identification of the PCR product of full-length *Caau-Ii*. 1, DL2000 DNA marker; 2, PCR product of full-length *Caau-Ii*.

## Amplification and analysis of full-length *Caau-Ii*

Amplified DNA sequences (3 in total) were assembled and analyzed using DNAMAN6. Based on the assembled sequence, a pair of primers, full-length Ii-F and full-length Ii-R was designed. The full-length *Caau-Ii*, 1063 bp in length was then PCR amplified (Figure 2, lane 2).

Using DNAStar software, we found that the ORF of *Caau-Ii* (nucleotide 42-737), encodes 231 amino acids (molecular weight= 25.48 kD and pI= 9.4). Analysis of *Caau-Ii* amino acid sequence by TMHMM-2 software showed that there was a putative transmembrane (TM) domain between amino acid 33 and 55. In addition, SignalP-3.0 software predicted a putative signal peptide (score: 0.052) and a signal anchor (score: 0.914) in the amino acid sequence of *Caau-Ii.*

## Alignment of *Caau-Ii* with the Ii of other vertebrates

*Caau-Ii* amino acid sequence was aligned with that of other vertebrates. The result showed that among these aligned species, the most highly conserved domains were transmembrane and thyroglobulin domain, followed by CLIP, cytoplamic and TRIM domains (Figure 3).

## Pylogenetic analysis

Pylogenetic analysis showed that *Caau-Ii* had the highest homology with that of carp and zebrafish (amino acid identity: 81 and 73%, respectively), a modest homology with that of chicken (amino acid identity: 53%) and a relatively low homology with that of mouse and human (amino acid identity: 34%) (Figure 4).

## DISCUSSION

Hashimoto et al. (1990) first amplified part of carp MHC gene and since then, cloning of MHC gene has been conducted in a wide range of fish. Ii is an important molecular chaperon and signal molecule and its mutations will not only blockMHC class II molecule transport, but also influence the development of B cells (Shachar and Flavell, 1996). Yoder et al. first cloned from zebrafish spleen cDNA library two cDNAs, *Iclp-1* and *Iclp-2*, encoding invariant chain like proteins (Yoder et al., 1999). Till date, Ii has been found in a host of fish such as sea bass (Silva et al., 2007) and rainbow trout (Dijkstra et al., 2003). In this study, using RACE techniques, we first cloned the full-length *Carassius auratus* Ii gene (*Caau-Ii*) and analyzed the structure of its predicted protein.

In the amino acid sequence of *Caau-Ii,* there are six highly conserved cysteines in the Tg domain close to its carboxyl-terminus (Figure 3). The Tg domain of mammalian Ii functions as a protease inhibitor that interacts with cathepsins (Lennon-Dumenil et al., 2002) and this domain is highly conserved in the molecular evolution of Ii. Additionally, at the amino-terminus, there is a highly conserved transmembrane (TM) domain consisting of 23 amino acid residues, of which 17 amino acid residues are identical or similar among the aligned species. Ashman and Miller (1999) suggested that TM domain plays a key role in the self-assembly of Ii into trimer, the formation of which is the basis for the interaction between MHC class II molecules and Ii. Accordingly, we speculate that TM domain of different

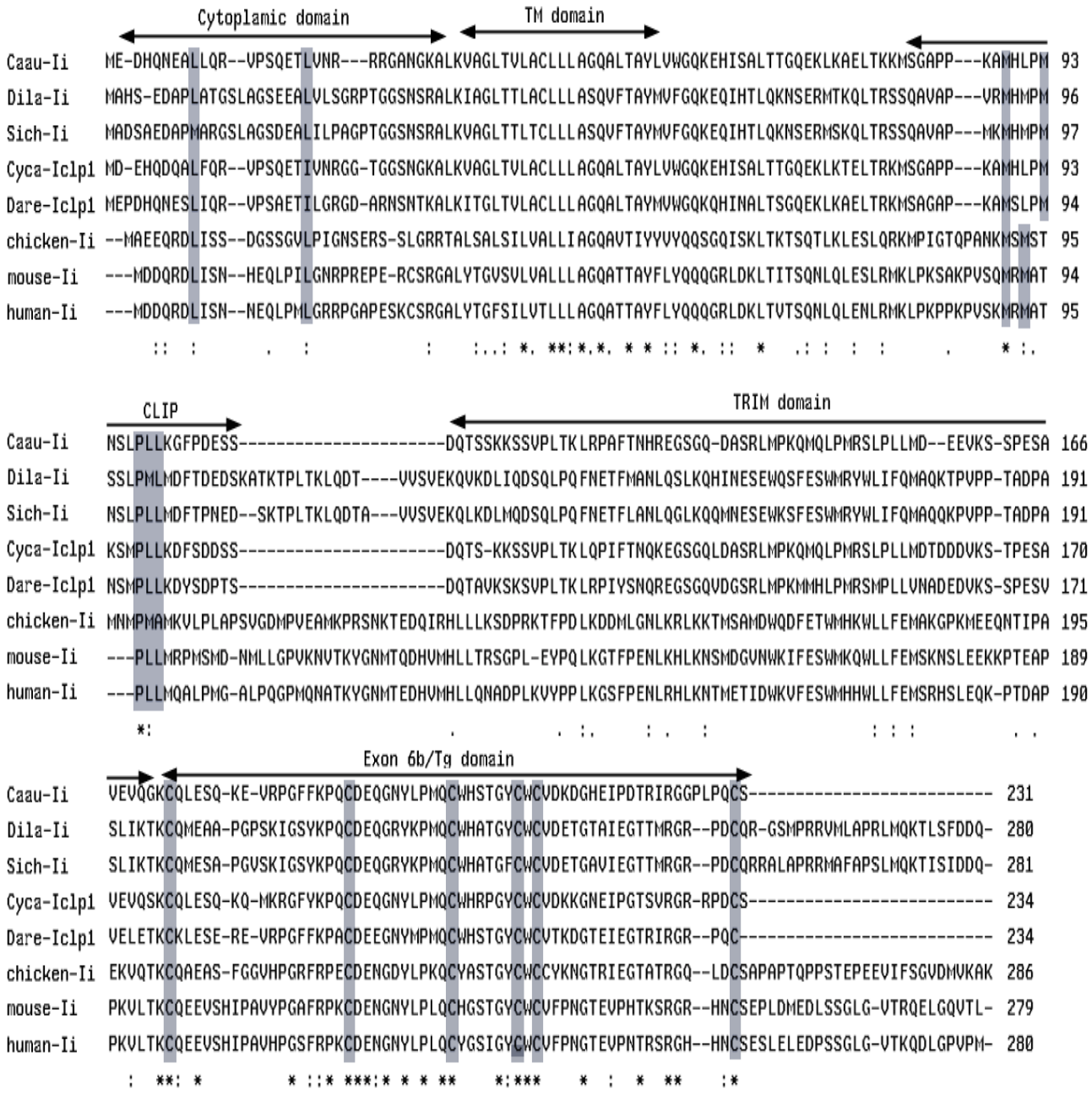


**Figure 3.** Alignment of *C. auratus* Ii cDNA with representative homologues of the Ii of other vertebrates. GenBank accession numbers not previously mentioned are: *Cyprinus carpio* ( Cyca-Iclp1), BAC53767; *D. labrax* (Dila-Ii), ABH09445; *Siniperca chuatsi* (Sich-Ii), AAS77256; Danio rerio (Dare-Iclp1), AAD24542; *Gallus gallus* (chicken-Ii), CAC27415; *Homo sapiens* (human-Ii), P04233; *Mus musculus* (mouse-Ii), P04441. Different domains, including cytoplamic, TM (transmembrane), CLIP (class II-associated Ii-derived peptide), TRIM (trimerization) and Tg (thyroglobulin) structures, are indicated by arrows.

species' Ii plays the same role in the formation of Ii trimer.

After the formation of its trimer, mammalian Ii aids MHC class II molecules in presenting antigens and performing other biological functions (Peterson and Miller, 1990; Bremnes et al., 2000). In the endoplasmic reticulum (ER), Ii trimers associate with three MHC class II heterodimers and prevent the binding of endogenous peptides to the class II molecules. Silva et al. (2007) cloned the Ii gene of *Dicentrarchus labrax* and by using Swiss-model and other bioinformatics softwares, constructed a 3D model of its amino acid structure, which through its comparison with human Ii (p41) structure showed that the Ii of *D. labrax* (Dila-Ii) has a similar

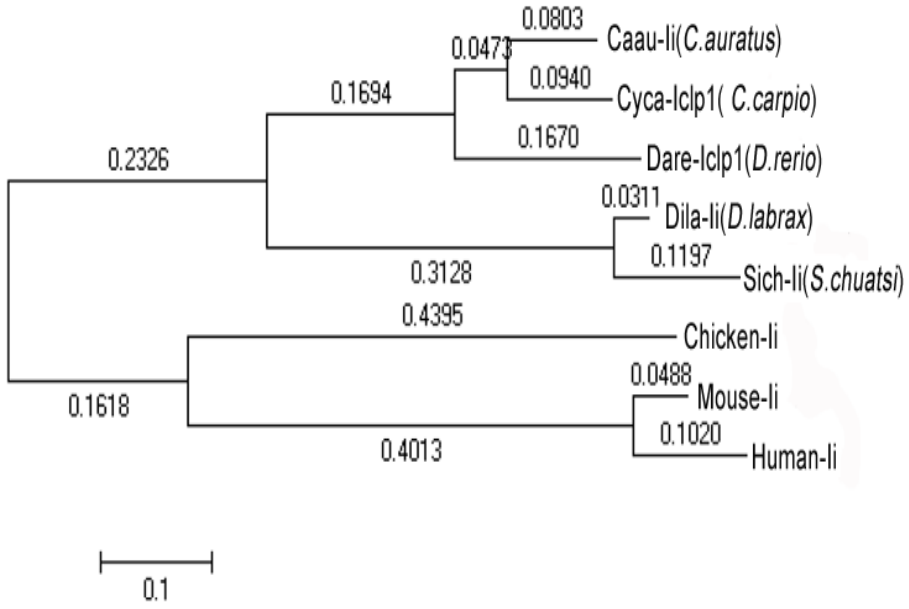


**Figure 4.** Neighbour-joining tree of Ii amino acid sequences generated by MEGA4 program.

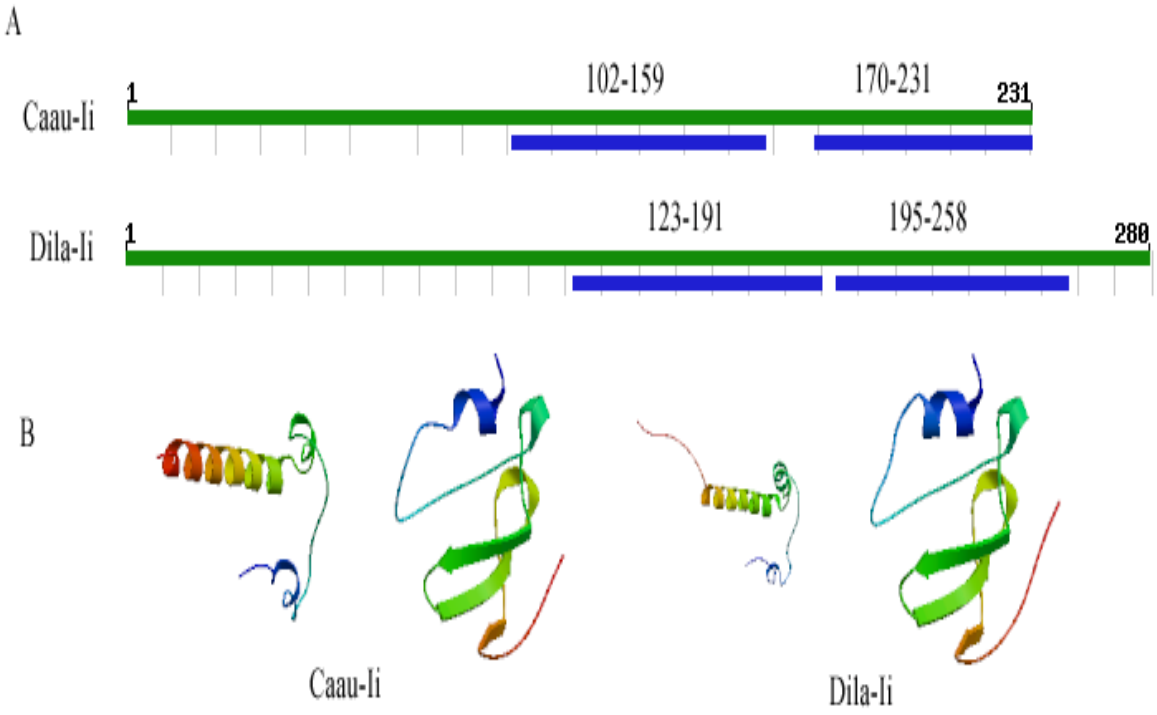


**Figure 5.** Homology modeling of the conformation of *Caau-Ii* (*C. auratus* Ii) based on the structure of *Dila-Ii* (*D. labrax* Ii). (A) There are two main domains of CLIP and TRIM; (B) space structure alignment of domains of *Caau-Ii* and *Dila-Ii.*

space conformation with mammalian Ii. Before their work, there had been an argument whether fish Ii could form the same trimer structure as mammalian Ii. In this study, by comparing the 3D models of *Caau-Ii* and *Dila-Ii* constructed using Swiss-model software, we showed that these two species' Ii are highly similar in conformation (Figure 5). However, it remains to be further studied whether *Caau-Ii* can form a trimer that can associate with MHC class II molecules, thereby preventing the binding of endogenous peptides to the class II molecules.

Moreover, the cytoplasmic domain of *Caau-Ii* contains a highly conserved di-leucine sequence ($L^{9}$ and $L^{19}$), which like its mammalian counterpart, may play a role in targeting endosomes (Silva et al., 2007; Johnson et al., 2001). The fact that the di-leucine sequence is also highly conserved in other fish species and poultry species suggests that this sequence might be phylogenetically conserved.

Currently, the structure of fish Ii's CLIP domain and the properties of its binding target, MHC II molecules, remain undefined. By using alignment, we found that in the center region of CLIP domain of the majority of fish species and mammals, there is a conserved motif PLL ($P^{97}$, $L^{98}$ and $L^{99}$ in Crucian carp, $P^{96}$, $L^{97}$ and $L^{98}$ in human) and a methionine residue ($M^{89}$ in crucian carp and $M^{91}$in human). It has been demonstrated in mammals that $M^{91}$, $P^{96}$ and $L^{97}$ play a major role in CLIP occupying the molecular groove of MHC class II and preventing the binding of endogenous peptides to the groove (Ghosh et al., 1995). Additionally, in this domain and TRIM domain, there is some degree of variation in amino acid sequence, which according to Fujiki et al. (2003) is consistent with Ii function in association with different MHC molecules.

## ACKNOWLEDGEMENTS

We are grateful to Dr. Su Wujie for kindly giving suggestions on the manuscript. The study was supported by the National Natural Science Foundation (Project number: 10975002) and Anhui Provincial Natural Science Foundation (Project number: 11040606M90) of China.

### REFERENCES

Anderson M, Miller J (1992). Invariant chain can function as a chaperone protein for class II major histocompatibility complex molecules. Proc. Natl. Acad. Sci. 89: 2282-2286.

Ansari AA (1993). A possible role of the MHC-associated in-variant chain in rheumatoid arthritis. Semin. Arthritis. Rheum. 23(3): 193-197.

Ashman JB, Miller J (1999). A role for the transmembrane domain in the trimerization of the MHC class II-associated invariant chain. J. Immunol. 163: 2704-2712.

Bremnes B, Rode M, Gedde-Dahl M, Nordeng TW, Jacobsen J, Ness SA, Bakke O (2000). The MHC class II-associated chicken invariant chain shares functional properties with its mammalian homologs. Exp. Cell Res. 259: 360-369.

Bryan KJ, Zhu XW, Harris PL, Perry G, Castellani RJ, Smith MA, Casadesus G (2008). Expression of CD74 is increased in neurofibrillary tangles in Alzheimer's disease. Mol. Neurodegener, 3: p. 13.

Cardenas C, Bidon-Chanal A, Conejeros P, Arenas G, Marshall S, Luque FJ (2010). Molecular modeling of class I and II alleles of the major histocompatibility complex in Salmo salar. J. Comput. Aided. Mol. Des. 24(12): 1035-1051

Cresswell P (1996). Invariant chain structure and MHC class II function. Cell. 84: 505–507.

Dijkstra JM, Kiryu I, Kollner B, Yoshiura Y, Ototake M (2003). MHC class II invariant chain homologues in rainbow trout (*Oncorhynchus mykiss*). Fish Shellfish Immun. 15(2): 91-105.

Faure-André G, Vargas P, Yuseff MI, Heuzé M, Diaz J, Lankar D, Steri V, Manry J, Hugues S, Vascotto F, Boulanger J, Raposo G, Bono MR, Rosemblatt M, Piel M, Lennon-Duménil AM (2008). Regulation of dendritic cell migration by CD74, the MHC class II-associated invariant chain. Science, 322: 1705-1710.

Fujiki K, Smith CM, Liu L, Sundick RS, Dixon B (2003). Alternate forms of MHC class II-associated invariant chain are not produced by alternative splicing in rainbow trout (*Oncorhynchus mykiss*) but are encoded by separate genes. Dev. Comp. Immunol. 27(5): 377-391.

Ghosh P, Amaya M, Mellins E, Wiley DC (1995). The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. Nature, 378: 457-462.

Hashimoto K, Nakanishi T, Kurosaw Y (1990). Isolation of carp genes encoding major histocompatibility complex antigens. Proc. Natl. Acad. Sci. 87: 6863-6867.

Holst PJ, Sorensen MR, Mandrup Jensen CM, Orskov C, Thomsen AR, Christensen JP (2008). MHC class II-associated invariant chain linkage of antigen dramatically improves cell-mediated immunity induced by adenovirus vaccines. J. Immunol. 180(5): 3339–3346.

Johnson AO, Lampson MA, McGraw TE (2001). A di-leucine sequence and a cluster of acidic amino acids are required for dynamic retention in the endosomal recycling compartment of fibroblasts. Mol. Biol. Cell. 12: 367-381.

Lennon-Dumenil AM, Bakker AH, Wolf-Bryant P, Ploegh HL, Lagaudriere-Gesbert C (2002). A closer look at proteolysis and MHC-class-II-restricted antigen presentation. Curr. Opin. Immunol. 14: 15-21.

Matza D, Kerem A, Shachar I (2003). Invariant chain, a chain of command. Trends Immunol. 24: 264-268.

Neumann J, Koch N (2005). Assembly of major histocompatibility complex class II subunits with invariant chain. FEBS. Lett. 579: 6055-6059.

Peterson M, Miller J (1990). Invariant chain influences the immunological recognition of MHC class II molecules. Nature, 345: 172-174.

Robertson M (1998). Antigen presentation. Curr. Biol. 8: 829-831.

Serwe M, Reuter G, Sponaas A, Koch S, Koch N (1997). Both invariant chain isoforms Ii 31 and Ii 41 promote class II antigen presentation. Int. Immunol. 9: 983-991.

Shachar I, Flavell R (1996). Requirement for invariant chain in B cell maturation and function. Science., 274: 106-108.

Silva DSP, Reis MIR, Nascimento DS, Vale AD, Pereira PJB, Santos NMS (2007). Sea bass (*Dicentrarchus labrax*) invariant chain and class II major histocompatibility complex: sequencing and structural analysis using 3D homology modeling. Mol. Immunol. 44: 3758-03776.

Yoder JA, Haire RN, Litman GW (1999). Cloning of two zebrafish cDNAs that share domains with the MHC class II-associated invariant chain. Immunogenetics, 50(1-2): 84-88.

Yu SH, Ao JQ, Chen XH (2010). Molecular characterization and expression analysis of MHC class II alpha and beta genes in large yellow croaker (*Pseudosciaena crocea*). Mol. Bio. Rep. 37(3): 1295-307.