

Elsevier Editorial System(tm) for Molecular Immunology
Manuscript Draft

Manuscript Number:

Title: A CD4 homologue in sea bass (*Dicentrarchus labrax*): molecular characterisation and structural analysis

Article Type: Full Length Article

Section/Category:

Keywords: CD4; *Dicentrarchus labrax*; real time PCR; 3D structure; MHC class II.

Corresponding Author: Dr Francesco Buonocore, Ph.D

Corresponding Author's Institution: University of Tuscia

First Author: Francesco Buonocore, Ph.D

Order of Authors: Francesco Buonocore, Ph.D; Elisa Randelli, Ph.D; Daniela Casani; Laura Guerra; Simona Picchiatti, Ph.D; Susan Costantini, Ph.D; Angelo M Facchiano, Ph.D; Jun Zou, Ph.D; Chris J Secombes, Ph.D; Giuseppe Scapigliati, PhD

Manuscript Region of Origin:

Abstract:

Dear Prof. M.R. Daha,

I am submitting you the paper: “A CD4 homologue in sea bass (*Dicentrarchus labrax*): molecular characterisation and structural analysis”, author team: Francesco Buonocore, Elisa Randelli, Daniela Casani, Laura Guerra, Simona Picchietti, Susan Costantini, Angelo M. Facchiano, Jun J. Zou, Chris J. Secombes, Giuseppe Scapigliati.

The paper is focused on the characterisation of a fundamental molecule of the sea bass immune system and will help in studying T cell subsets in this teleost.

Best regards

Dr. Francesco Buonocore, Ph.D
Univ. of Tuscia
Dep. Environmental Sciences

A CD4 homologue in sea bass (*Dicentrarchus labrax*): molecular characterisation and structural analysis

ABSTRACT

CD4 is a transmembrane glycoprotein fundamental for cell-mediated immunity. Its action as a T cell co-receptor increases the avidity of association between a T cell and an antigen-presenting cell by interacting with portions of the complex between MHC class II and TR molecules. In this paper we report the cDNA cloning, expression and structural analysis of a CD4 homologue from sea bass (*Dicentrarchus labrax*). The sea bass CD4 cDNA consists of 2071 bp that translates in one reading frame to give the entire molecule containing 480 amino acids. The analysis of the sequence shows the presence of four putative Ig-like domains and that some fundamental structural features, like a disulphide bond in domain D2 and the CXC signalling motif in the cytoplasmic tail, are conserved from sea bass to mammals. Real time PCR analysis showed that very high levels of CD4 mRNA transcripts are present in thymus, followed by gut and gills. *In vitro* stimulation of head kidney leukocytes with LPS and PHA-L gave an increase of CD4 mRNA levels after 4 h and a decrease after 24 h. Homology modelling has been applied to create a 3D model of sea bass CD4 and to investigate its interaction with sea bass MHC-II. The analysis of the 3D complex between sea bass CD4 and sea bass MHC-II suggests that the absence of a disulfide bond in the CD4 D1 domain could make this molecule more flexible, inducing a different conformation and affecting the binding and the way of interaction between CD4 and MHC-II. Our results will add new insights into the sea bass T cell immune responses and will help in the identification of T cell subsets in teleost fishes to better understand the evolution of cell-mediated immunity from fish to mammals.

**A CD4 homologue in sea bass (*Dicentrarchus labrax*):
molecular characterisation and structural analysis**

**Francesco Buonocore*, Elisa Randelli, Daniela Casani, Laura Guerra, Simona
Picchietti, Susan Costantini¹, Angelo M. Facchiano¹, Jun Zou², Chris J.
Secombes², Giuseppe Scapigliati**

Dipartimento di Scienze Ambientali, University of Tuscia, Largo dell'Università, 01100 Viterbo,
Italy

¹Laboratorio di Bioinformatica e Biologia Computazionale, Istituto di Scienze dell'Alimentazione,
CNR, I-83100 Avellino

²Scottish Fish Immunology Research Centre, Aberdeen University, AB24 2TZ Aberdeen, UK.

*Corresponding author: Dr. Francesco Buonocore, Dipartimento di Scienze Ambientali,
Università della Tuscia, Largo dell'Università s.n.c., I-01100 Viterbo, Italy.

Phone +39-0761-357644; Fax +39-0761-357179; Email: fbuono@unitus.it

ABSTRACT

CD4 is a transmembrane glycoprotein fundamental for cell-mediated immunity. Its action as a T cell co-receptor increases the avidity of association between a T cell and an antigen-presenting cell by interacting with portions of the complex between MHC class II and TR molecules. In this paper we report the cDNA cloning, expression and structural analysis of a CD4 homologue from sea bass (*Dicentrarchus labrax*). The sea bass CD4 cDNA consists of 2071 bp that translates in one reading frame to give the entire molecule containing 480 amino acids. The analysis of the sequence shows the presence of four putative Ig-like domains and that some fundamental structural features, like a disulphide bond in domain D2 and the CXC signalling motif in the cytoplasmic tail, are conserved from sea bass to mammals. Real time PCR analysis showed that very high levels of CD4 mRNA transcripts are present in thymus, followed by gut and gills. *In vitro* stimulation of head kidney leukocytes with LPS and PHA-L gave an increase of CD4 mRNA levels after 4 h and a decrease after 24 h. Homology modelling has been applied to create a 3D model of sea bass CD4 and to investigate its interaction with sea bass MHC-II. The analysis of the 3D complex between sea bass CD4 and sea bass MHC-II suggests that the absence of a disulfide bond in the CD4 D1 domain could make this molecule more flexible, inducing a different conformation and affecting the binding and the way of interaction between CD4 and MHC-II. Our results will add new insights into the sea bass T cell immune responses and will help in the identification of T cell subsets in teleost fishes to better understand the evolution of cell-mediated immunity from fish to mammals.

Key words: CD4; *Dicentrarchus labrax*; real time PCR; 3D structure; MHC class II.

1. INTRODUCTION

The two major subclasses of T cells, helper T (T_H) cells and cytotoxic T (T_C) cells, are characterized in mammals by different responses to antigens. T_H cells stimulate the expression and secretion of cytokines that produce antibody responses or lead to macrophage activation, while T_C cells are involved in the killing of the antigen-bearing cells. The T cell antigen receptor (TR) recognizes a complex formed by a peptide antigen bound to the major histocompatibility complex (MHC) molecules (Wange and Samelson, 1996). The antigens are presented by different types of MHC molecules: antigens presented by class II MHC molecules generally elicit a T_H response, whereas antigens presented by class I MHC molecule give rise to a T_C response. CD8 and CD4 are T cell co-receptor glycoproteins that participate in antigen recognition through interactions with non-polymorphic regions of class I and class II MHC molecules, respectively (Salter et al., 1990; Konig et al., 1992; Cammarota et al., 1992). Therefore, CD4-expressing cells are predominantly of the helper phenotype, while CD8-expressing cells of the cytotoxic phenotype (Germain, 2002).

In mammals, CD4 is a transmembrane molecule, belonging to the immunoglobulin superfamily (IgSF), that contains four Ig-like domains (D1-D4), two of which are V-like domains (D1 and D3) and the other two C-like domains (D2 and D4). Moreover, CD4 has a cytoplasmic tail that associates with a tyrosine protein kinase, p56^{lck}, that provides the first signal for T cell activation (Lin et al., 1998). Producing diffraction-quality crystals of the entire extracellular region of CD4 proved difficult. However, proteolytic analyses showed that this molecule could be cleaved into stable fragments that proved much more amenable to crystallographic studies. Both TR-independent binding assays and T cell functional assays, in conjunction with mutagenesis of human or murine CD4 molecules, suggested that an extended region on the CD4 D1 and D2 domains interact with class II MHC molecule (Konig et al., 1996).

In teleost fish, CD4 cDNAs have been recently obtained from fugu (*Takifugu rubripes*) (Suetake et al., 2004), rainbow trout (*Oncorhynchus mykiss*) (Dijkstra et al., 2006; Laing et al., 2006) and channel catfish (*Ictalurus punctatus*) (Edholm et al., 2007). Some particular features have been

demonstrated. In *fugu* the CD4 contains four extracellular Ig-like domains, as in mammals, but lacks the Cys residue pairs of the first Ig-like domain, and the predicted CD4 gene is composed of 12 exons, compared with 10 in mammals and birds. In rainbow trout two CD4-like genes have been found, one with the classical four Ig-like domains and the other with only two Ig-like domains. Similarly, in channel catfish two CD4-like genes have been found, one with the classical four Ig-like domains and the other with three Ig-like domains. These findings have raised some questions on the functional significance of the cloned CD4 molecules but, nevertheless, other evidence suggests that the T cell system is likely to be similar from teleost fish to mammals. For example, both cytotoxic T like cells (Nakanishi et al., 2002) and T helper cells have been observed upon in vitro incubation of allogeneic leukocytes (Heiger et al., 1977; Caspi et al., 1984; Meloni et al., 2006). However, the lack of suitable markers for specific T lymphocytes has prevented a precise characterization of fish T cell subsets.

Here we describe the cloning of a CD4 homologue cDNA from sea bass (*Dicentrarchus labrax*), a fish of high economic importance in Mediterranean aquaculture. Moreover, we investigate its 3D structure, alone or in combination with MHC class II, by molecular modelling techniques to identify the sites of interaction between these two molecules. With the same species we recently studied CD8 α (Buonocore et al. 2006), and MHC class II (Buonocore et al., 2007a), and, therefore, this sequence will add a new marker to study sea bass T cell subsets.

2. MATERIALS AND METHODS

2.1 Sea bass CD4 cloning and sequencing

Two primers (CD4F1: 5'- TCTGACCTGCATCTGAAGTGG -3' and CD4R1: 5'- TGTAGAATCCTCTGGGTTTGGG -3') corresponding to highly conserved regions of known CD4 genes were used in RT-PCR on total RNA extracted with Tripure (Roche) solution from a juvenile sea bass (150 g of weight) thymus, as described by Scapigliati et al. (2001). RT-PCR was performed using Ready-To-Go RT-PCR Beads (GE Healthcare). For cDNA synthesis, 1 µg of total RNA and 0.5 µg of random primers [pd(N)₆] were used in each reverse transcription reaction in a total volume of 50 µl. Reactions were conducted using a Mastercycler (Eppendorf). The cycling protocol was one cycle of 94°C for 5 min, 35 cycles of 94°C for 45 s, 48°C for 45 s, 72°C for 45 s, followed by one cycle of 72 °C for 10 min. PCR products (15 µl) were visualised on 1% (w/v) agarose gels containing ethidium bromide (10 ng/ml) using hyperladder IV (Bioline) as size marker. Controls for the presence of DNA contamination were performed using the RNA samples as template. DNA amplified by PCR was purified using a QIAquick Gel Extraction Kit (QIAGEN), inserted into the pGEM-T Easy vector (Promega) and transfected into competent JM109 *Escherichia coli* cells. Plasmid DNA from at least ten independent clones was purified using the Wizard Plus SV Minipreps DNA Purification System (Promega) and sequenced using MWG DNA Sequencing Services. Sequences generated were analysed for similarity with other known sequences using the FASTA (Pearson and Lipman, 1988) and BLAST (Altschul et al., 1990) programs.

Further primers were designed based on the initial sea bass CD4 sequence for 5'- and 3'- rapid amplification of cDNA ends (RACE)-PCR (3'CD4FW1: 5'- CTGACCATCACCCCACTC -3' and 3'CD4FW2: 5'- GGAAGATGACACACCTCAG -3'; 5'CD4RW1: 5'- CTGAGGTGTGTCATCTTCC -3' and 5'CD4RW2: 5'- GAGTGGGGTGATGGTCAG -3').

cDNA was synthesised from total thymus RNA with a First-strand cDNA Synthesis kit (GE Healthcare) following the manufacturer's instructions. For 3' RACE-PCR, cDNA was transcribed

using an oligo-dT adaptor primer (5'-CTCGAGATCGATGCGGCCGCT₁₅-3'). PCR was performed initially with the 3'CD4FW1 primer and the oligo-dT adaptor primer, followed by a semi-nested PCR using 3'CD4FW2 primer and the adaptor primer (5'-CTCGAGATCGATGCGGCCGC-3'). For 5' RACE-PCR, cDNA was transcribed from total RNA using the oligo-dT primer, treated with *E. coli* RNase H (Promega), purified using a PCR Purification Kit (QIAGEN), and tailed with poly(C) at the 5' end with terminal deoxynucleotidyl transferase (TdT, Promega). PCR was performed initially with 5'CD4RW1 primer and an Oligo-dG primer (5'-GGGGGGIGGGIIGGGIIG-3'), and then semi-nested with 5'CD4RW2 and the oligo-dG primers. Sequencing and similarity searches were as described above.

The obtained cDNA sequence was analysed for the presence of a signal peptide, using SignalP software (Nielsen et al., 1997), and for N- (with the NetNGlyc 1.0 Server) and O-linked glycosylation sites (Julenius et al., 2005). Alignment of the sea bass CD4 amino acid sequences to their counterparts from other species was carried out using MEGA 3.1 Software (Kumar et al., 2004). A phylogenetic tree was constructed by the “neighbour-joining” method using MEGA 3.1 Software (Kumar et al., 2004) on full-length amino acid sequences and bootstrap values calculated.

2.2 Basal and in vitro CD4 expression analysis

To study CD4 basal expression, six sea bass juveniles (150 g of weight) were sampled and leucocytes from different tissues (peripheral blood (PBL), liver, brain, gut, thymus, head kidney (HK), gills, and spleen) were obtained as described by Scapigliati et al. (2001). Total RNA was isolated from each tissue separately with Tripure (Roche) following the manufacturer's instructions, resuspended in DEPC treated water and used for real-time quantitative PCR with samples from individual analysed separately. Controls for the presence of DNA contamination were performed using β -actin primers that were either side of an intron.

The *in vitro* CD4 expression of HK leucocytes was studied after different stimulating conditions, using cells from six sea bass juveniles as described above. HK leucocytes were adjusted to 1×10^5 cells/ml and incubated at 18 °C for 4 h and 24 h with 5 μ g/ml of lipopolysaccharide (LPS

from *E. coli* 0127:B8, Sigma) in PBS or with 1 µg/ml of lectin from *Phaseolus vulgaris* Leucoagglutinin (PHA-L from Sigma) in PBS. The control samples were stimulated with PBS and analysed at the same time points. Total RNA was isolated with Tripure (Roche) and the samples treated as described above for the basal expression.

For reverse transcription, the BioScript RNase H minus (Bioline) enzyme was used as described by Buonocore et al. (2007b). The expression level of CD4 was determined with a Mx3000P™ real time PCR system (Stratagene) equipped with version 2.02 software using the Brilliant SYBR Green Q-PCR Master Mix (Stratagene) and following the manufacturer's instructions, with ROX as internal passive reference dye. Specific PCR primers were designed for the amplification of about 200 bp products from both CD4 (primers selected to amplify a region between the transmembrane domain and the cytoplasmic tail) and β-actin, used as an house-keeping gene for subsequent normalisation of the data. The primers were: RQCD4FW: 5'-GTGATAACGCTGAAGATCGAGCC-3' and RQCD4RW: 5'-GAGGTGTGTCATCTTCCGTTG-3'; RQACTFR: 5'-ATGTACGTTGCCATCC-3' and RQACTRV: 5'-GAGATGCCACGCTCTC-3', respectively. Approximately 10 ng of cDNA template was used in each PCR reaction. The PCR conditions were 95 °C for 10 min, followed by 35 cycles of 95 °C for 45 s, 52 °C for 45 s and 72 °C for 45 s. Triplicate reactions were performed for each template cDNA and the template was replaced with water in all blank control reactions. The analysis was carried out using the endpoints method option of the Mx3000P™ software that causes the collection of the fluorescence data at the end of each extension stage of amplification. A relative quantitation was performed, comparing the levels of the target transcript (CD4) to a reference transcript (calibrator). For the basal expression analyses the calibrator was the tissue with the lowest CD4 expression (the brain), whereas for the *in vitro* expression the time 0 control was used. A normalizer target (the β-actin transcript) was included to correct for differences in total cDNA input between samples. The results are expressed as the mean ± SD of the results obtained from six fish and, for the *in vitro* expression analyses, the differences from the control were considered significant if $p < 0.05$ using the one-way ANOVA test.

The real-time PCR products from the different tissues were also examined by agarose gel electrophoresis to confirm their specificity and size.

2.3 3D Modelling of sea bass CD4

A three-dimensional model of sea bass CD4 was created using the template model of human CD4 complexed with the mouse MHC I-A^k β chain and an antigen peptide (Wang et al., 2001) (PDB code: 1JL4 chain D). As the sequence identity between the sea bass protein and the homologous template was quite low (22%), an accurate procedure was used to search the best alignment of sequences in agreement with the rules recently reviewed to improve the quality of the modelling results at low target-template sequence similarity (Dalton and Jackson, 2007). The search for sequence similarity within databases was performed with the BLAST program (Altschul et al., 1990). The alignment of CD4 protein sequences from 25 different organisms was made with the 3D-Coffee program (O'Sullivan et al., 2004) and a few manual refinements were included to align the cysteine residues and to account for the position of secondary structures. The MODELLERv7 module (Sali and Blundell, 1993) within the InsightII program (Accelrys, San Diego, CA, USA) was used to build 10 full-atom models of sea bass CD4 by setting 4.0 Angstrom as root mean square deviation (RMSD) among the structures of the template and fully optimized models. The best model among those obtained was selected using the PROCHECK program (Laskowski et al., 1993) to evaluate the stereochemical quality of the models and their structural packing quality and the ProsaII program (Sippl, 1993) was used to check the fitness of sequence to structure and to assign a scoring function. In the selected model the loop regions were refined using the LOOPY module of the Jackal package (Xiang et al., 2002) that appears to produce the most accurate loop conformations (Dalton and Jackson, 2007), as it generates a large number of random conformations via the *ab-initio* method, minimizing each of the random candidates and selecting the best candidates using colony energy. Secondary structures were assigned by the DSSP program (Kabsch and Sander, 1983). A search for structural classification was performed on the CATH database

(Orengo et al., 1997; Pearl et al., 2000). Molecular superimposition, RMSD values and figures were obtained with the InsightII package (Accelrys, San Diego, CA).

2.4 Simulation of the sea bass MHC class II/CD4 complex

The complex between sea bass CD4 and MHC class II β chain was created using as reference the experimental model of the mouse MHC I-A^k β chain complexed with human CD4 (PDB code: 1JL4) (Wang et al., 2001). In detail, we superimposed the coordinates of sea bass CD4 and MHC-II β chain recently modelled by homology (see Buonocore et al., 2007a) to those of the experimental complex. CVFF force field developed for the Insight II software was adopted to assign potentials and charges for the complexed proteins. The complex in sea bass was minimized by using 500 steps of energy minimization under the conjugate gradient algorithm in order to optimise side chain conformations and avoid steric clashes as used in our previous modelling studies (Scapigliati et al., 2004; Costantini et al., 2005; Chambery et al., 2007; Randelli et al., in press). To compare the protein-protein interaction in the complex, the “Protein - Protein Interaction Server” (Jones and Thornton, 1996) and the NACCESS program (Hubbard et al., 1991) were used to evaluate the interface surface area and to identify the amino acids at the protein interface. Lastly, H-bonds were calculated using the Hbplus program (McDonald and Thornton, 1994) and binding free energy using the DCOMPLEX program (Liu et al., 2004).

3. RESULTS

3.1 Sea bass CD4 sequence analysis

PCR with primers CD4F1 and CD4R1 gave a product of the expected size (350 bp, see Fig. 1) that when sequenced showed similarity with other known CD4 molecules (data not shown). 3'-RACE-PCR performed with CD4FW2 (based on the initial 350 bp sequence) and the adaptor primer gave a product of 625 bp (see Fig. 1). 5'-RACE-PCR performed with CD4RW2 (based on the initial 350 bp sequence) and oligo- dG gave products of about 1250 bp (see Fig. 1). The full-length cDNA (EMBL accession number AM849811) is comprised of 2071 bp from the three overlapping products and was confirmed by PCR using primers that amplify the complete coding sequence (data not shown). Finally, the 3'-UTR contained a polyadenylation signal (AATAAA) 15 bp upstream of the poly(A) tail (Fig. 1).

An analysis of the sea bass CD4 sequence (Fig. 1) revealed the presence of a putative 38 amino acid signal peptide (most likely cleavage site between Gly38 and Glu39), two potential N-glycosylation sites and four putative O-glycosylation sites. Comparison of the sea bass CD4 nucleotide and amino acid sequence to its counterparts in other species is shown in Table I. The highest nucleotide and amino acid identity was with Fugu (*Takifugu rubripes*), followed by rainbow trout (*Oncorhynchus mykiss*), whilst the lowest identity was to mouse (*Mus musculus*).

A multiple alignment of the sea bass CD4 amino acid sequence with other known CD4 sequences was assembled (Fig. 2) to investigate the conservation of characteristic amino acid residues involved in structural domains. The sea bass CD4 had four Ig-like domains (D1-D4), with the first and second known to be fundamental for class II MHC binding in mammals (Clayton et al., 1989, Huang et al., 1997, König 2002). The first Ig-like domain of sea bass CD4 (D1) had a cysteine (i.e. Cys114 in sea bass sequence) that is conserved among all known CD4 molecules, but the cysteine residue (C41 in human) that is involved in the formation of a disulfide bridge in mammals and birds was missing. The second Ig-like domain had two cysteine residues: one (Cys193 in sea bass) is present in all known sequences; the other (Cys156 in sea bass) is conserved

in all fish sequences except zebrafish and is quite close to the cysteine (Cys155 in human and Cys159 in mouse) that is involved in the formation of a disulphide bridge. The third Ig-like domain also had two cysteine residues (Cys233 and Cys316 in sea bass) present in all fish sequences that could be involved in the formation of an additional disulphide bond. Difference in the number and location of the disulphide bonds in CD4 molecules has already been observed, as in some species, like chicken, dog and whale no disulphide bridges are formed in the D2 domain (Milde et al., 1993; Romano et al., 1999; Koskinen et al., 2002). In the fourth Ig-like domain two cysteine residues (Cys355 and Cys404 in sea bass) and one N-glycosylation site (Asn352-Leu353-Thr354 in sea bass) are conserved in all sequences, again except in zebrafish. The transmembrane region is not well conserved between the sequences, whereas the cytoplasmic tail shows an interesting conserved feature. In this domain, the CXC motif is present in all sequences, except in zebrafish, and it mediates the binding of the tail to the tyrosine protein kinase p56^{lck} by means of a Zn clasp structure (Lin et al., 1998) to initiate the first signal for T cell activation.

Phylogenetic analysis (Fig. 3) performed using the amino acid sequences shows that sea bass CD4 grouped with other known teleost sequences, with the Fugu CD4 being the closest homologue. A separate cluster is formed by the mammalian and avian sequences.

3.2 Basal and in vitro CD4 expression analysis

The basal expression analysis of CD4 in organs and tissues of un-stimulated sea bass is shown in Figure 4 (Panel A). Real-time PCR products were loaded on agarose gels to exclude the formation of non-specific amplicons and single bands of the expected sizes were obtained. Moreover, to take into consideration the individual genetic variability six different fishes were sampled and analysed separately. The highest CD4 expression was detected in thymus, followed by gut and gills. Lower CD4 levels were observed in spleen, liver and PBL. The brain showed the lowest expression level.

To investigate whether CD4 expression levels could be modulated by LPS and PHA-L, *in vitro* stimulation of HK leucocytes for a short (4 h) and a longer (24 h) time was studied using HK

leukocytes from six individual fish that were analysed separately. Real-time PCR products were loaded on agarose gels and single bands of the expected sizes were obtained. The results are shown in Figure 4 (Panel B). Using both LPS and PHA-L, a significant increase ($p < 0.05$) of CD4 expression was detected after 4 h, whereas after 24 h stimulation a significant decrease was seen using PHA-L.

3.3 3D Modelling of sea bass CD4

BLAST searching within the non-redundant database of all known protein sequences shows that the sea bass CD4 sequence is similar to many other proteins defined as “CD4” (E-value lower than 10^{-26} with fish sequences and about 10^{-13} - 10^{-3} with mammalian and avian sequences). A BLAST search also found an experimentally determined three-dimensional structure of two human CD4 Ig-like domains (D1-D2) that could be considered as a suitable template for comparative modelling of sea bass CD4. The pairwise alignment between sea bass and human sequences returned low amino acid identity. This level of similarity requires careful analysis to build up a 3D model of a protein by comparative modelling. In these cases the main problem is related to the finding of the right sequence alignment, since sequences with low identity can be aligned differently with similar scores, and gaps required to better align the sequences may occur in the middle of secondary structure elements, with the consequence of obtaining a wrong model (Tramontano, 1998; Kopp and Schwede, 2004). Therefore, we applied an accurate alignment procedure recently described (Dalton and Jackson, 2007) restricted to the D1 and D2 Ig-like CD4 domains, for which it was possible to create a 3D model by using the human crystallographic structure as template. Starting from this alignment (see Fig. 5), ten structural models were created for the sea bass CD4 39-212 amino acid region. Fig. 6 shows the best model chosen for sea bass CD4 with its secondary structure elements. This model has a classical organization in two distinct Ig-like domains (D1 and D2) in agreement with the structural classification reported by the CATH database (Orengo et al., 1997; Pearl et al., 2000) for the crystallographic structure of human CD4. The D1 domain is classified as a “V-like domain” and is characterized by an immunoglobulin-like beta-sandwich

made of nine beta-strands (ABCC'C''DEFG). As already observed (see section 3.1) and as in other fish CD4 sequences, it has a conserved cysteine in the F strand but lacks the second cysteine in the B strand for disulfide bridge formation in contrast to mammalian and bird CD4 molecules. The D2 domain is a "C2-like domain" and is characterized by an immunoglobulin-like beta-sandwich made of seven beta-strands (ABCC'EFG) including a pair of cysteine residues (Cys 156-Cys 193) that may stabilize the Ig fold via a disulfide bridge.

Superimposition of structurally conserved regions of sea bass and human models gave an RMSD value of 0.75 Angstroms. This value indicates that the 3D structures are still similar even if their sequence identity is low. Comparison of the secondary structures elements (Fig. 5) revealed that the beta-strands are conserved, with few external residues being added or excluded. However, three short beta-strands and the 3_{10} helix present in the human D2 domain are not conserved in the sea bass CD4 model.

3.4 Simulation of sea bass MHC class II/CD4 complex

On the basis of the crystallographic structure of the complex between human CD4 and murine MHC class II β chain we simulated the same interaction in sea bass. For each complex we evaluated the interaction residues, the number of intrachain H-bonds, the interface surface area and the energy of binding (Figure 7). The amino acids of the CD4 at the interface with the MHC-II molecule are located in C' and C'' strands of the D1 domain, and interact with MHC-II amino acids in the $\beta 2$ domain (Figure 8).

The interaction regions between CD4 and MHC-II in the experimentally determined and sea bass complexes are quite well conserved (see also Fig. 5) and, in particular, Phe68 of human CD4, indicated as the most striking feature of the interaction and located at the beginning of the C'' strand (Wang et al., 2001), is conserved in sea bass CD4 although not in Fugu and rainbow trout.

The number of interaction residues is similar in both experimentally determined and sea bass modelled complexes. The value of the interface surface area and the number of intrachain H-bonds is higher for the experimentally determined complex (Figure 7, panel A). Furthermore, for each

complex we calculated the binding energy (Fig. 7, panel B) and a higher value was again seen for the experimentally determined complex. These results could be due to the absence of a disulfide bond in sea bass CD4 D1 domain that could make this molecule more flexible, inducing a different conformation and affecting the binding and the way of interaction between CD4 and MHC class II.

4. DISCUSSION

CD4 and CD8 T cell co-receptors participate in guiding the development and the selection of immature thymocytes (Itano and Robey, 2000) and have been extensively studied in human due to their importance for immune system activation. Recently, CD4-like molecules have been identified in teleost fishes and some interesting differences from mammalian homologues have been found. For example, in rainbow trout in addition to a conventional CD4 molecule cDNA sequences codifying for only two Ig-like domains have been found (Dijkstra et al., 2006; Laing et al., 2006), whilst in catfish the presence of sequences coding for only three Ig-like domains has been found (Edholm et al., 2007).

In this paper, we report the homology cloning of a CD4 homologue from sea bass (*Dicentrarchus labrax*), and present an analysis of its expression and tertiary structure, with particular attention paid to the regions of interactions with the MHC-II molecule.

The sea bass CD4 cDNA consists of 2071 bp that translates in one reading frame to give the entire molecule containing 480 amino acids. An analysis of the sequence shows the presence of four putative Ig-like domains and some fundamental structural features that are conserved from fish to mammals, even though the overall similarity among the sequences is quite low. The first cysteine of the D1 domain is not conserved in sea bass and this is likely to affect the interactions with MHC-II as observed in the modelled complex. However, the Phe residue that in mammalian CD4 binds to MHC-II through the interaction of its phenol ring (Wang et al., 2001) is present in sea bass CD4. The possibility of a disulphide bond formation in the D2 region is confirmed in sea bass, both from the alignment and from the 3D structure. Moreover, the WXC motif in the D2 F strand is conserved in all fish sequences and it may represent a unique CD4 feature (Triebel et al., 1990). The CXC motif in the cytoplasmic tail, that interacts with the tyrosine protein kinase p56^{lck}, is also present in sea bass, including several basic (Arg) residues in the helical region immediately preceding the motif itself. A dileucine motif (Leu-Leu) and two Ser residues associated with CD4 internalization are absent in the sea bass molecule, as already observed for Fugu and trout (Laing et al., 2006). The

regulation of mammalian CD4 has been linked to the phosphorylation of one of these serine residues by protein kinase C (Sleckman et al., 1992), but as they are also absent in duck and chicken (Luhtala, 1998) sequences this may differ in other vertebrate groups. At the moment, no sequences corresponding to CD4 molecules with less than four Ig-like domains have been found in sea bass.

Real-time PCR analysis detected high levels of CD4 expression in thymus, as already observed in rainbow trout (Dijkstra et al., 2006; Laing et al., 2006), human and mouse (Maddon et al., 1987), and chicken (Koskinen et al., 2002). Moderate expression was seen in spleen, as observed in rainbow trout (Dijkstra et al., 2006; Laing et al., 2006), in gut and in gills. The gut is a tissue deeply involved in immune responses and leukocytes ontogeny in sea bass (Rombout et al., 2005) and has a quite high percentage of T cells (Romano et al., 2007). The gills in fish are constantly exposed to a plethora of water born antigens and some pathogens use them as a portal of entry into the host (Morris et al., 2000) and so are also a key immune tissue. Finally, it is interesting to note that sea bass CD8 α (Buonocore et al., 2006) and MHC-II (Buonocore et al., 2007a) basal expression in the different organs and tissues examined matches that seen for CD4 in this study.

In vitro CD4 expression was studied using sea bass head kidney leukocytes with the stimulant LPS, to simulate a pathogen infection, or the T cell mitogen PHA-L. Both LPS and PHA-L induced high CD4 expression after 4 h, although a decrease was seen at 24 h with PHA-L. LPS results are in agreement with recent mammalian data (McAleer et al., 2007), where bacterial LPS can act as a natural adjuvant that produces profound effects on T cell clonal expansion, effector differentiation and long-term cell survival. PHA is a typical T cell mitogen and so will act to expand the T cell population present.

The sea bass CD4 3D analysis showed the presence of a putative disulphide bond in the D2 domain, as observed in mammals, and allowed the antigenicity of this region to be studied. The segment that shows the highest value of immunogenicity (Parker et al., 1986) is between amino acid residues 164-190 of the D2 domain, in a well exposed and not structured region. Therefore, it could

be used to design a peptide to try to raise monoclonal antibodies in future studies. The analysis of the interaction region in the complexes between sea bass CD4 and the six MHC-II β -chains characterised in a previous paper (see Buonocore et al., 2007a), revealed that the amino acid residues are located in a highly conserved region of the MHC-II sequences (data not shown). It may be possible to design a peptide, based on these conserved amino acids, that can block complex formation and that, therefore, should inhibit T cell activation *in vitro* and permit interesting functional analyses, as already performed in mammals (Quintana et al., 2007).

In conclusion, the availability of this new sea bass CD4 sequence, together with the available sequences of CD8 α and MHC-II, gives the opportunity to study in more detail T cell immune responses in this species and to try to assess if this teleost possesses a T cell subset comparable with Th cells of mammals.

ACKNOWLEDGEMENTS

Authors are indebted to Dr. C. Magugliani and Dr. E. Vela (Nuova Azzurro, Civitavecchia) for the supply of fish. This work was supported by the European Commission within the project IMAQUANIM (EC contract number FOOD-CT-2005-007103).

- 468 Altschul S.F., Gish W., Miller W., Myers E., Lipman D.J., 1990. Best local alignment search tool. J.
469 Mol. Biol. 215, 403-410.
- 470
- 471 Buonocore, F., Randelli, E., Bird, S., Secombes, C.J., Costantini, S., Facchiano, A., Mazzini, M.,
472 Scapigliati, G., 2006. The CD8 α from sea bass (*Dicentrarchus labrax* L.): cloning, expression and
473 3D modelling. Fish Shellfish Immunol. 20, 637-646.
- 474
- 475 Buonocore, F., Randelli, E., Bird, S., Secombes, C.J., Facchiano, A., Costantini, S., Scapigliati, G.,
476 2007b. Interleukin-10 expression by real-time PCR and homology modelling analysis in the
477 European sea bass (*Dicentrarchus labrax* L.). Aquaculture 270, 512-522.
- 478
- 479 Buonocore, F., Randelli, E., Casani, D., Costantini, S., Facchiano, A., Scapigliati, G., Stet, R.J.,
480 2007a. Molecular cloning, differential expression and 3D structural analysis of the MHC class-II β
481 chain from sea bass (*Dicentrarchus labrax*). Fish Shellfish Immunol. 23, 853-866.
- 482
- 483 Cammarota, G., Scheirle, A., Takacs, B., Doran, D.M., Knorr, R., Bannwarth, W., Guardiola, J.,
484 Sinigaglia, F., 1992. Identification of a CD4 binding site on the β_2 domain of HLA-DR molecules.
485 Nature 356, 799-801.
- 486
- 487 Caspi, R.R., Avtalion, R.R., 1984. The mixed leukocyte reaction (MLR) in carp: bidirectional and
488 unidirectional MLR responses. Dev. Comp. Immunol. 8, 631-637.
- 489
- 490 Chambery, A., Pisante, M., Di Maro, A., Di Zazzo, E., Costantini, S., Colonna, G., Parente, A.,
491 2007. Invariant Ser211 is involved in the catalysis of PD-L4, type I RIP from (*Phytolacca dioica*)
492 leaves. Proteins: Structure, Function and Bioinformatics 67, 209-218.
- 493
- 494 Clayton, L.K., Sieh, M., Pious, D.A., Reinherz, E.L., 1989. Identification of human CD4 residues
495 affecting class II MHC versus HIV-I gp120 binding. Nature 339, 548-551.
- 496
- 497 Costantini, S., Colonna, G., Rossi, M., Facchiano, A.M., 2005. Modelling of HLA-DQ2 and
498 simulations of its interaction with gluten peptides to explain molecular recognition in celiac disease.
499 J. Mol. Graph. Model. 23, 419-431.
- 500
- 501 Dalton, J.A.R., Jackson, R.M., 2007. An evaluation of automated homology modelling methods at
502 low target-template sequence similarity. Bioinformatics 23, 1901-1908.
- 503
- 504 Dijkstra, J.M., Somamoto, T., Moore, L., Hordvik, I., Ototake, M., Fischer, U., 2006. Identification
505 and characterization of a second CD4-like gene in teleost fish. Mol. Immunol. 43, 410-419.
- 506
- 507 Edholm, E.-S., Stafford, J.L., Quiniou, S.M., Waldbieser, G., Miller, N.W., Bengten, E., Wilson,
508 M., 2007. Channel catfish, *Ictalurus punctatus*, CD4-like molecules. Dev. Comp. Immunol. 31,
509 172-187.
- 510
- 511 Germain, R.N., 2002. T-cell development and the CD4-CD8 lineage decision. Nat. Rev. Immunol.
512 2, 309-322.
- 513
- 514 Heiger, H.M., Hodgins, H.O., Chiller, J.M., 1977. Evolution of the lymphoid system. II. Evidence
515 for immunoglobulin determinants on all rainbow trout lymphocytes and demonstration of mixed
516 leukocyte reaction. Eur. J. Immunol. 7, 881-887.

- Huang, B., Yachou, A., Fleury, S., Hendrickson, W.A., Sekaly, R.P., 1997. Analysis of the contact sites on the CD4 molecule with class II MHC molecule: co-ligand versus co-receptor function. *J. immunol.* 158, 216-225.
- Hubbard, S.J., Campbell, S.F., Thornton, J.M., 1991. Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. *J. Mol. Biol.* 220, 507-530.
- Itano, A., Robey, E., 2000. Highly efficient selection of CD4 and CD8 lineage thymocytes supports an instructive model of lineage commitment. *Immunity* 12, 383-389.
- Jones, S., Thornton, J.M., 1996. Principles of protein-protein interactions derived from structural studies. *Proc. Natl. Acad. Sci. USA* 93, 13-20.
- Julenius, K., Molgaard, A., Gupta, R., Brunak, S., 2005. Prediction, conservation analysis and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 15, 153-164.
- Kabsch, W., Sander, C., 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577-2637.
- Konig, R. 2002. Interactions between MHC molecules and co-receptors of the TCR. *Curr. Opin. Immunol.* 14, 75-83.
- Konig, R., Fleury, S., Germain, R.N., 1996. The structural basis of CD4-MHC class II interactions: coreceptor contributions to T cell receptor antigen recognition and oligomerization-dependent signal transduction. *Curr. Top Microbiol. Immunol.* 205, 19-46.
- Konig, R., Huang, L., Germain, R.N., 1992. MHC class II interaction with CD4 mediated by a region analogous to the MHC class I binding site for CD8. *Nature* 346, 796-798.
- Kopp, J., Schwede, T. 2004. Automated protein structure homology modelling: a progress report. *Pharmacogenomics* 5, 405-416.
- Koskinen, R., Salomonsen, J., Tregaskes, C.A., Young, J.R., Goodchild, M., Bumstead, N., Vainio, O., 2002. The chicken CD4 gene has remained conserved in evolution. *Immunogenetics* 54, 520-525.
- Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated Software for Molecular Evolutionary Genetics Analysis and Sequence Alignment. *Briefings in Bioinformatics* 5, 150-163.
- Laing, K.J., Zou, J.J., Purcell, M.K., Phillips, R., Secombes, C.J., Hansen, J.D., 2006. Evolution of the CD4 family: teleost fish possess two divergent forms of CD4 in addition to lymphocyte activation gene-3. *J. Immunol.* 177, 3939-3951.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK - A program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 26, 283-291.
- Lin, R.S.C., Rodriguez, A., Veillette, Lodish, H.F., 1998. Zinc is essential for binding of p56(lck) to CD4 and CD8 α . *J. Biol. Chem.* 273, 32878-32882.

- Liu, S., Zhang, C., Zhou, H., Zhou, Y., 2004. A physical reference state unifies the structure-derived potential of mean force for protein folding and binding. *Proteins* 56, 93-101.
- Luhtala, M., 1998. Chicken CD4, CD8 $\alpha\beta$, and CD8 $\alpha\alpha$ T cell co-receptor molecules. *Poultry Sciences* 77: 1858-1873.
- Maddon, P.J., Molineaux, S.M., Maddon, D.E., Zimmerman, K.A., Godfrey, M., Alt, F.W., Chess, L., Axel, R., 1987. Structure and expression of the human and mouse T4 genes. *Proc. Natl. Acad. Sci. USA* 84, 9155-9159.
- McAleer, J.P., Zammit, D.J., Lefrancois, L., Rossi, R.J., Vella, A.T., 2007. The lipopolysaccharide adjuvant effect on T cells relies on nonoverlapping contributions from the MyD88 pathway and CD11c+ cells. *J. Immunol.* 179, 6524-6535.
- McDonald, I.K., Thornton, J.M., 1994. Satisfying hydrogen bonding potential in proteins. *J. Mol. Biol.* 238, 777-793.
- Meloni, S., Zarletti, G., Benedetti, S., Randelli, E., Buonocore, F., Scapigliati, G., 2006. Cellular activities during a mixed leucocyte reaction in the teleost sea bass *Dicentrarchus labrax*. *Fish Shellfish Immunol.* 20, 739-749.
- Milde, K., Conner, G.E., Mintz, D.H., Alejandro, R., 1993. Primary structure of a canine CD4 antigen. *Biochim. Biophys. Acta* 1172, 315-318.
- Morris, D.J., Adams, A., Richards, R.H., 2000. *In situ* hybridisation identifies the gill as a portal of entry for PKK (Phylum Myxozoa), the causative agent of proliferative kidney disease in salmonids. *Parasitol. Res.* 86, 950-956.
- Nakanishi, T., Fisher, U., Dijkstra, J.M., Hasegawa, S., Somamoto, T., Okamoto, N., Ototake, M., 2002. Cytotoxic T cell function in fish. *Dev. Comp. Immunol.* 26, 131-139.
- Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* 10, 1-6.
- Orengo, C.A., Michie, A.D., Jones, S., Jones, D.T., Swindells, M.B., Thornton, J.M., 1997. CATH-A Hierarchic Classification of Protein Domain Structures. *Structure* 5, 1093-1108.
- O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D.G., Notredame, C., 2004. 3DCoffee: combining protein sequences and structures within multiple sequence alignments. *J. Mol. Biol.* 340, 385-395.
- Parker, J.M.R., Guo, D., Hodges, R.S., 1986. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray derived accessible sites. *Biochemistry* 25, 5425-5432.
- Pearl, F.M.G., Lee, D., Bray, J.E., Sillitoe, I., Todd, A.E., Harrison, A.P., Thornton, J.M., Orengo, C.A., 2000. Assigning genomic sequences to CATH. *Nucl. Acids Res.* 28, 277-282.
- Pearson, W.R., Lipman, D.J., 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* 85, 2444-2448.

Quintana, F.J., Gerber, D., Bloch, I., Cohen, I.R., Shai, Y., 2007. A structurally altered D,L-amino acid TCRalpha transmembrane peptide interacts with the TCRalpha and inhibits T-cell activation in vitro and in an animal model. *Biochemistry* 46, 2317-2325.

Randelli, E., Scala, V., Casani, D., Costantini, S., Facchiano, A., Mazzini, M., Scapigliati, G., Buonocore, F. T cell receptor beta chain from sea bream (*Sparus aurata*): molecular cloning, expression and modelling of the complexes with MHC class I. *Mol. Immunol.*, in press.

Romano, T.A., Ridgway, S.H., Felten, D.L., Quaranta, V., 1999. Molecular cloning and characterization of CD4 in an aquatic mammal, the white whale *Delphinapterus leucas*. *Immunogenetics* 49, 376-383.

Romano, N., Rossi, F., Abelli, L., Caccia, E., Piergentili, R., Mastrolia, L., Randelli, E., Buonocore, F., 2007. Majority of TcRβ(+) T-lymphocytes located in the thymus and midgut of the bony fish, *Dicentrarchus labrax* (L.). *Cell Tissue Res.* 329, 479-489.

Rombout, J.H., Huttenhuis, H.B., Picchiatti, S., Scapigliati, G., 2005. Phylogeny and ontogeny of fish leucocytes. *Fish Shellfish Immunol.* 19, 441-455.

Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779-815.

Salter, R.D., Benjamin, R.J., Wesley, P.K., Buxton, S.E., Garret, T.P.J., Clayberger, C., Krensky, A.M., Norment, A.M., Littman, D.R., Parham, P., 1990. A binding site for the T-cell co-receptor CD8 on the α3 domain of HLA-A2. *Nature* 345, 41-46.

Scapigliati, G., Buonocore, F., Bird, S., Zou, J., Pelegri, P., Falasca, C., Prugnoli, D., Secombes, C.J., 2001. Phylogeny of cytokines: molecular cloning and expression analysis of sea bass *Dicentrarchus labrax* interleukin-1 beta. *Fish Shellfish Immunol.* 11, 711-726.

Scapigliati, G., Costantini, S., Colonna, G., Facchiano, A., Buonocore, F., Bossù, P., Holland, J.W., Secombes, C.J., 2004. Modelling of fish interleukin 1 and its receptor. *Dev. Comp. Immunol.* 28, 429-441.

Sippl, M.J., 1993. Recognition of errors in three-dimensional structures of proteins. *Proteins* 17, 355-362.

Sleckman, B.P., Shin, J., Igra, V.E., Collins, T.L., Strominger, J.L., Burakoff, S.J., 1992. Disruption of the CD4-p56lck complex is required for rapid internalization of CD4. *Proc. Natl. Acad. Sci. USA* 89, 7566-7570.

Suetake, H., Araki, K., Suzuki, Y., 2004. Cloning, expression, and characterization of the fugu CD4, the first ectothermic animal CD4. *Immunogenetics* 56, 368-374.

Tramontano, A., 1998. Homology modelling with low sequence identity. *Methods* 14, 293-300.

Triebel, F., Jitsukawa, S., Baixeras, E., Roman-Roman, S., Genevee, C., Viegas-Pequignot, E., Hercend, T., 1990. LAG-3, a novel lymphocyte activation gene closely related to CD4. *J. Exp. Med.* 176, 327-337.

Wang, J.H., Meijers, R., Xiong, Y., Liu, J.H., Sakihama, T., Zhang, R., Joachimiak, A., Reinherz, E.L., 2001. Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule. *Proc. Natl. Acad. Sci. USA* 98, 10799-10804.

Wange, R.L., Samelson, E., 1996. Complex complexes: signaling at the TCR. *Immunity* 5, 197-205.

Xiang, Z., Soto, C.S., Honig, B., 2002. Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. *Proc. Natl. Acad. Sci. USA* 99, 7432-7437.

FIGURE CAPTIONS

Figure 1. The cDNA and encoded amino acid sequence of sea bass CD4. The primers used for the cloning are indicated. The start and stop nucleotide sequences are in bold, the putative signal peptide and the primers are in italics, the two putative N-glycosylation sites are underlined, the four putative O-glycosylated tyrosine residues are in bold and the polyadenylation signal is in bold and italics.

Figure 2. Alignment of the predicted sea bass CD4 amino acid sequence with other known CD4 molecules. Regions corresponding to the putative signal peptide, the four Ig-like domains (D1-D4), the transmembrane region (TM) and cytoplasmic tail are shown above the sequences according to the human CD4 protein. Conserved residues are shown in bold and conserved cysteines are indicated with an asterisk below the sequences. Accession numbers: sea bass (*Dicentrarchus labrax*) AM849811; Fugu (*Takifugu rubripes*) BAD37153; rainbow trout (*Oncorhynchus mykiss*) AAY42070; channel catfish (*Ictalurus punctatus*) ABD93351; common carp (*Cyprinus carp*) ABD58988; zebrafish (*Danio rerio*) XP_001340275; duck (*Anas platyrhynchos*) AAW63061; human (*Homo sapiens*) NP_000607; mouse (*Mus musculus*) NM_013488.

Figure 3. Phylogenetic tree showing the relationship between sea bass CD4 sequence with other known CD4 molecules. The rooted tree was constructed by the “neighbour-joining” method and was bootstrapped 10000 times. 0.2 indicates the genetic distance.

Figure 4. Basal and *in vitro* CD4 expression analyses. Panel A: Sea bass CD4 basal expression in different tissues. CD4 mRNA levels were expressed as a ratio relative to β -actin levels in the same samples after real-time PCR analysis using the tissue with the lowest expression (brain) as calibrator. Panel B: *In vitro* sea bass CD4 expression analysis. LPS: CD4 mRNA levels expressed as a ratio relative to β -actin levels in the same samples after real-time PCR analysis of HK leucocytes stimulated with PBS (control) and with 5 μ g/ml LPS for 4 and 24 h and normalised against the non-stimulated 0 h

control. PHA: CD4 mRNA levels expressed as a ratio relative to β -actin levels in the same samples after real-time PCR analysis of HK leucocytes stimulated with PBS (control) or with 1 μ g/ml PHA-L for 4 and 24 h and normalised against the non-stimulated 0 h control. Controls for 4 and 24 h of incubation with PBS only are also shown in the graphs. Data were expressed as the mean $\pm$ SD and asterisks indicates when $p < 0.05$ with respect to the time 0 control.

Figure 5. Primary and secondary structures of human and sea bass CD4 DI and D2 domains, aligned for the modelling procedure. The cysteine residues are in italics and bold. Beta-strands are marked with continuous line boxes and 3_{10} helices with dashed line boxes. The interaction residues of human and sea bass CD4 with the MHC class II β chain are shown in bold.

Figure 6. 3D model of sea bass CD4 obtained by homology modelling. The backbone ribbon and the secondary structure topology are shown: yellow arrows represent beta strands, red cylinders represent helices. Green and yellow sticks indicate the possible presence of a disulphide bond in the CD4 D2 domain.

Figure 7. Sea bass and human CD4-MHC-II complexes. Panel A: Analysis of the complexes in terms of interface surface area, intrachain H-bonds and number of interaction residues. Panel B: The bars represent the energies of binding (expressed in kcal/mol) evaluated for the simulated sea bass complex and for the experimentally determined complex.

Figure 8. 3D model of sea bass CD4 -MHC-II complex. The backbone ribbon diagram of the complex between sea bass MHC-II (blue) and CD4 (green) is shown: yellow arrows represent beta strands, red cylinders represent helices.. Labels indicate the CD4 D1 and D2 domains and the MHC-II β 1 and β 2 domains.

TABLE CAPTIONS

Table I. Nucleotide identity and amino acid identity and similarity of sea bass CD4 with other known CD4 sequences.

acatcttttcagcctgcagc**atg**agcgtgtgttttggggacacctgcactcagcaacattgtta
M G S C V L G T L C T Q Q L L
catgtaacagagatgaagaactttatccaatccgctcctcatttctcatcactgtggctgatg
H V T E M K N F I Q S V L I L I T V V M
tcagcagcaggggaaagaactgatatactgtcaggaggacacgtttacccttgaaactc
S A A G E E L I Y A E E G G Q M V T L N P
ccagcagtgtactaatccacaacacattattctcactgggtgttcaatggcaatcaaatt
P A V T N P Q T H Y S Y W V F N G N Q I
gctgtgctaactccttttccgaaagggggtcaatgataaagattcattgtccttgtact
A W R N P F S G K G V N D K D S L S L T
gacggctcagtgctgtcatcacaaaacatccaacaaaattctcttgggaacttttacctgt
D G S V L V I T A N I Q Q N L F G T F T C
caaatttatcagagtgggaatcgtagacacccgtggacacaaccacatataaaaaatctc
Q I Y T S G N R D T P V D T T T Y K I L
aaactcagtggtgactatggaccaccctctcctctcgtcgtcgtgggaagatctgtctctg
K L S V T M D P P S P L L P G G E D L S L
aactgcaatgcaggggagaaatccaagatatcactggctgaatccccaggggacagaaaaa
N C N A G R N P K I H W L N P Q G G Q K I
aacagccaaagaattcaacagaaagccacgggccaaagacagagagtggaactgtgtg
N S Q R I Q Q K A T G Q D K G E W T C V
gtgacatatcttaataaagagagcaaaagccaaaatttctgtcacactgtgtggacctacc
V T Y S N K E S K A K I S V T L V D L T
acagctcttcacatcttcaatatcgtctaaatctcgcctctcaccatcccctgttcc
T A P S H L Q Y T S K S S P L T I P C S
actaccgcccactgtggacacatacaaaactaggacattgaggaggaaaggcgggaactc
T T A T W D T I K T R D I E G G K R E F
tttgctaagccaggcgcgagtgctaaaatctaattgtctccacagatgctcttctccctctc
F A K P G A S L K S N A P Q M L F S L S
ctggatccgctaaagtggagcgcagaccaagaacataactgcgttatgacctaaactl
L D P L K W T P D Q D N E L R Y D P K L
caaaagggaattctgtctttgaagagaaatcaagggaagatggagacagtgaggactat
Q K G I L S L K R N Q G K D G D S G D Y
gtgtgctctctcaaatttaaaatgtgtgaactcggaacaggagataacagtcagtattgy
V C S L K F K N G V T L N R T V T V H V
ctgcaaatcacctcctccccgggaatagagtttaatttcaggccaacaggtcaacctgact
L Q I T S S P G I E L I S G Q Q V N L T
→ CD4F1
tgtgatttgggcagtcgcgtgccctctgcactgcagctgaaatggttcccacctggacaa
C D L G S P L P S D L Q L K W F P P G Q
→ 3' CD4FW1
tcgtccctgccgtctctgaaatctgacctacccccactcatctcatcatcccggaaagt
S S L P S L K S D H H P T H L I I P E V
5' CD4RW2 ←
ggatgggagatggcggaagtggaggtgtgagctgtggcggaacagcacacggcttagc
G M G D G G K W R C E L W R N S T R L T
→ RQCD4FW
tcagctgtgataacgtcgaagatcgagcccaagattactgtgtggatgatagtgtcata
S A V I T L K I E P K I T V W M I V L I
tgcaagtgtcacagtcattgcaatcctcctcctcgtcgtcgtttcatcctcaccagcgc
C S V T V I A I L L L V V A F I L Y R R
→ 3' CD4FW2
agacaacggaagatgacacacctcaggcatcgactctgccgatgcaaaaacccccagccc
R Q R K M T H L R H R L C R C K N P Q P
5' CD4RW1 ←
RQCD4RW ←
agaggattctcacagatcataataaactcccagaagacattttgaaaggagcatcccaacc
R G F Y R S -
CD4R1 ←
aattgaatttctgtccggataagatggagttttctgtcatgtctactacactgtgatcct
ctgcttggtgtgttagatgaggtctgaaagttaaacaagaattcccatcaattgacgaaa
ttttcacaataatgtattttacagtgtacatatcctttgattgtttatatattgtgtaaa
atgtaatcgatttcatccagacttcacatttttaattattgttgtttatgattattactgt
taataattgttatactctcttcatatttattttctcagatcattttgtcatgtatgtta
tgtatttttatgctgttctatgatattgatgccttgatagacttggttacttgctttaatg
aatgttacaataatataattttaagaatgcgcaaaatactgtattactaaaaaaaaaac
aactcttttgtctagaaattgtttattgttagattgaccactaacatttttcaataaaga
atcaataaactttcaaaaaaaaaaaaaaaaaaaaa
Oligo-dT adaptor ←

Figure 3

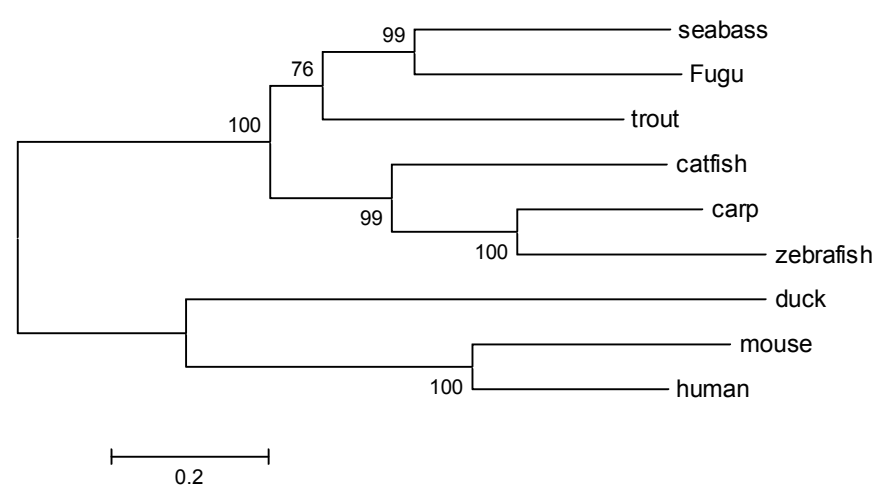


Figure 4

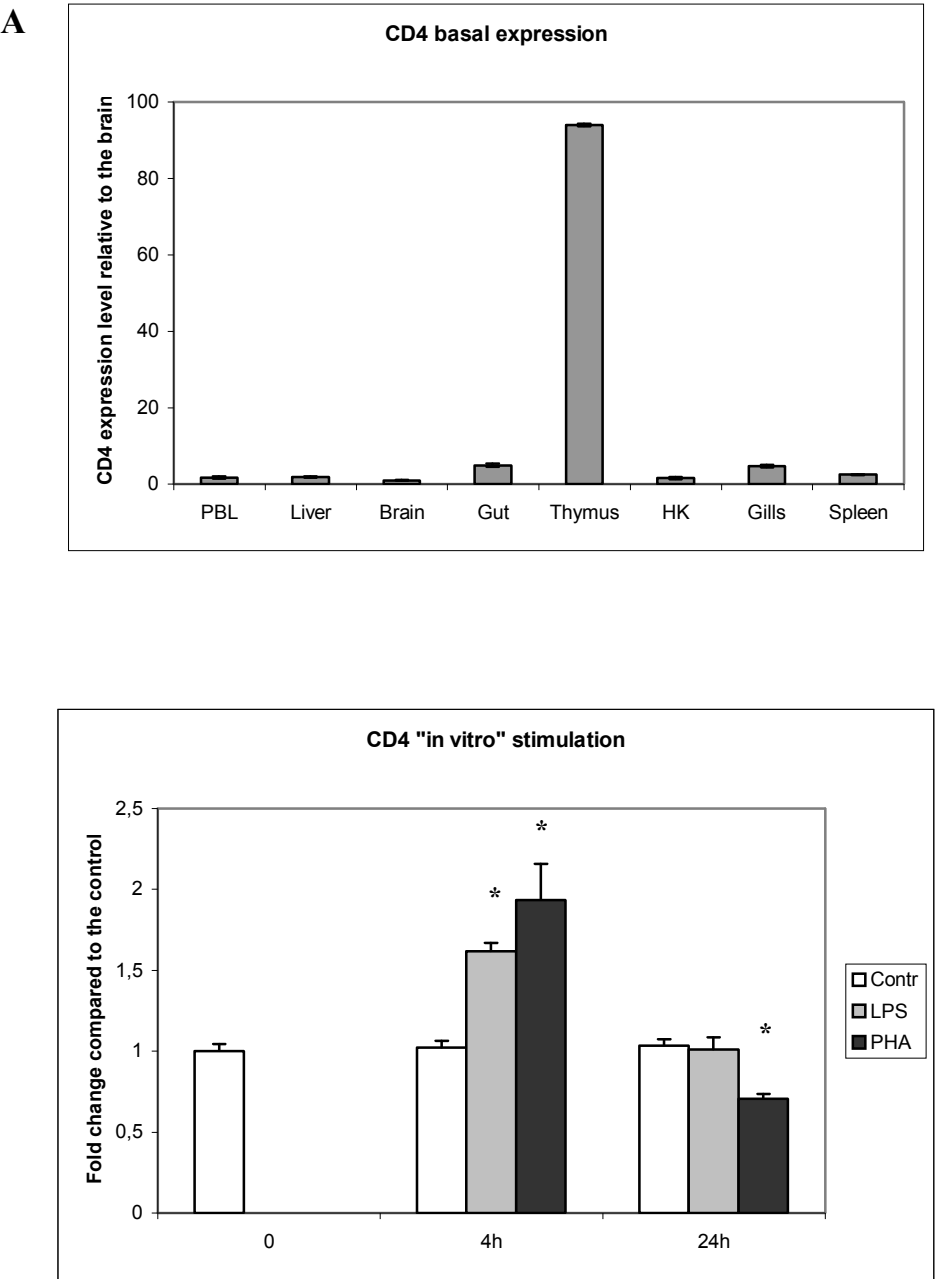


Figure 5

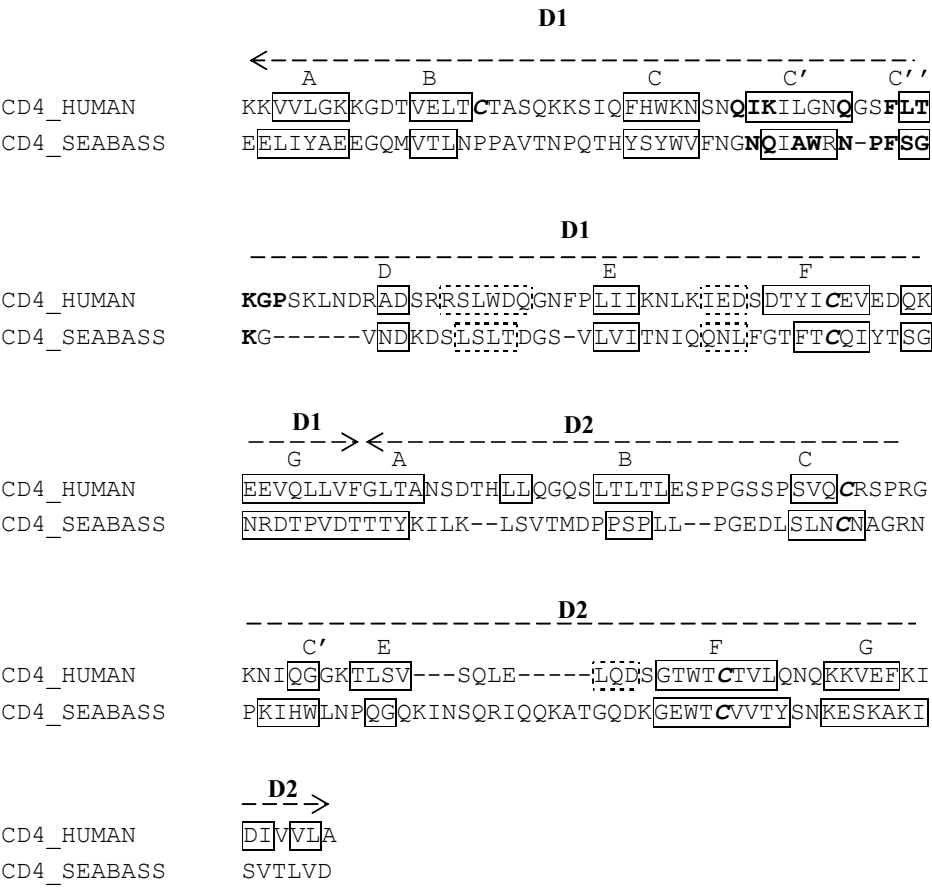


Figure 6

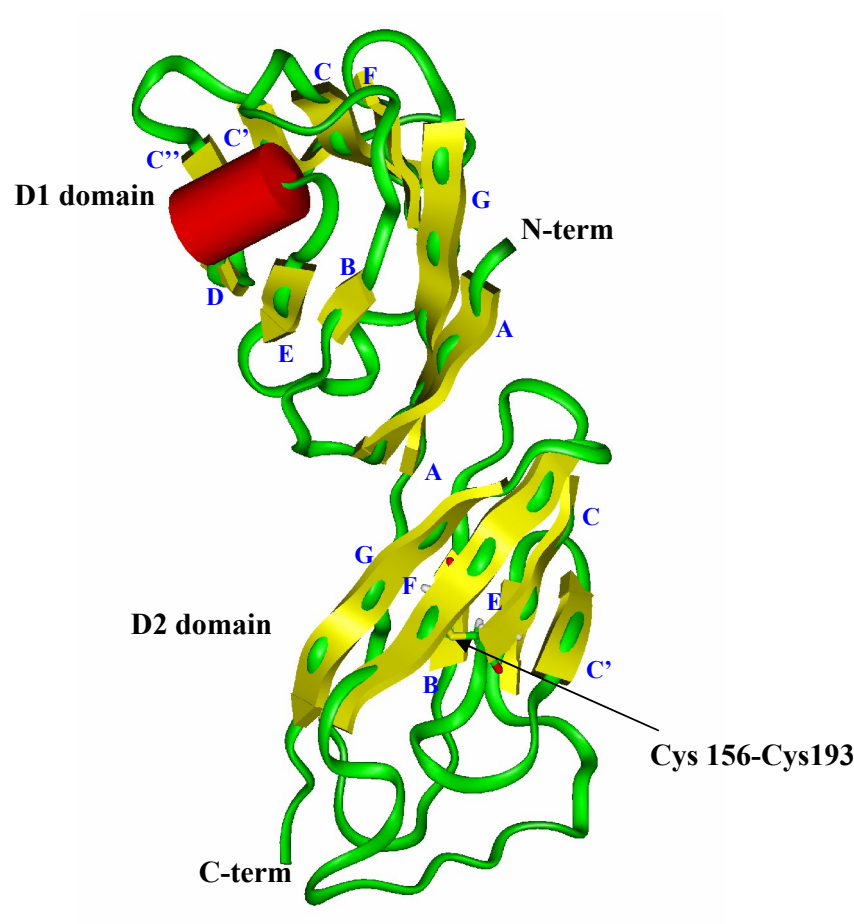


Figure 7

A

	Interface Surface Area	Interchain H-bonds	Interaction residues
MHC-II			
murine	503.96	4	15
CD4 human	537.95	4	10
MHC-II sea			
bass	415.30	3	14
CD4 sea bass	456.65	3	10

B

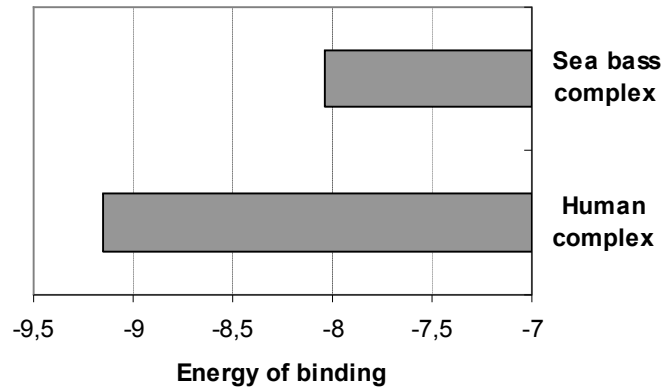


Figure 8

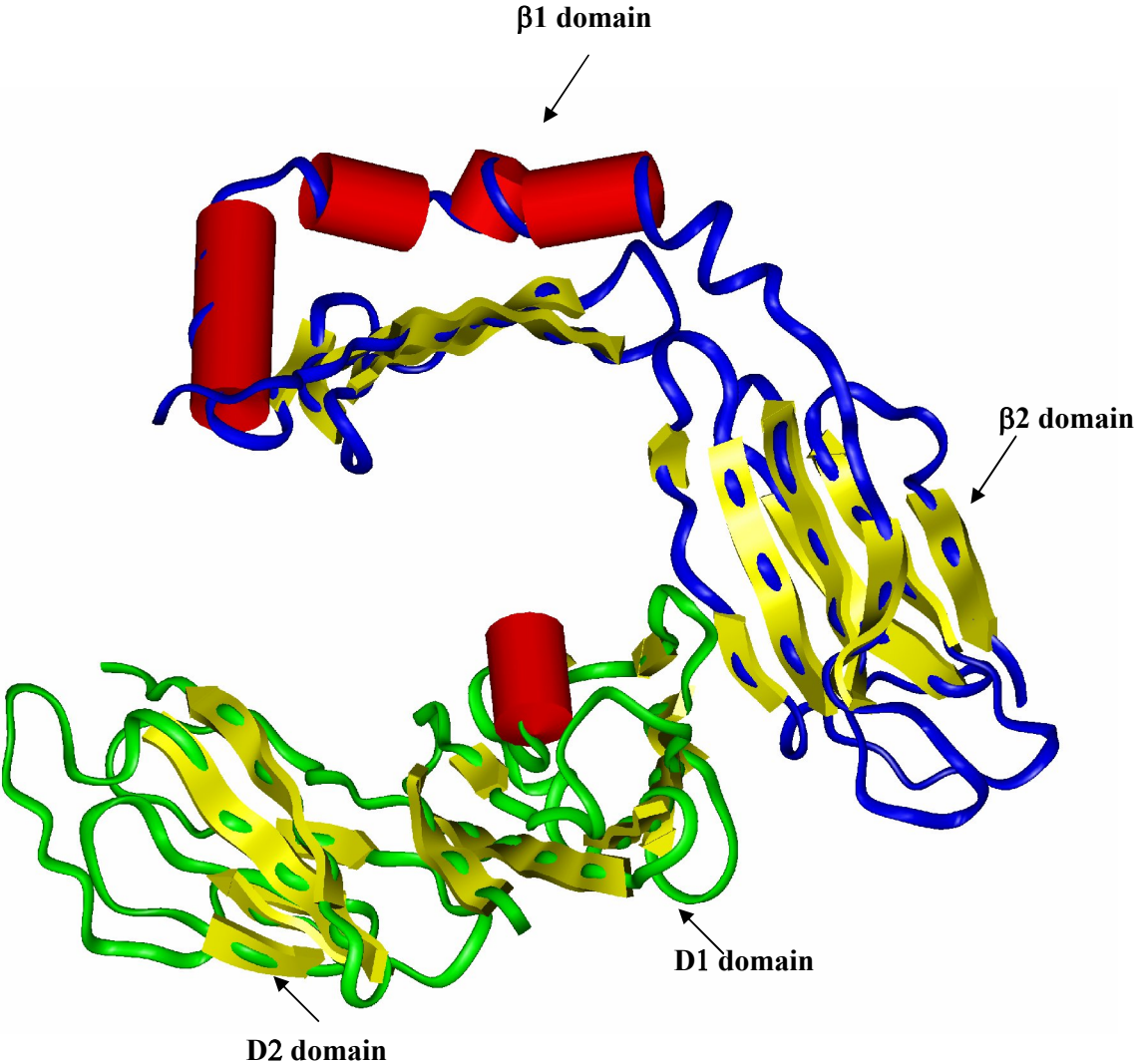


Table I

	Nucleotide identity	Amino acid identity	Amino acid similarity
Fugu	61	46	63
rainbow trout	56	40	55
catfish	52	33	51
carp	49	30	47
zebrafish	49	27	43
duck	47	21	38
human	46	23	40
mouse	43	19	38