Molecular cloning, characterization, and expression pattern of Atlantic halibut (*Hippoglossus hippoglossus* L.) CD4-2, Lck, and ZAP-70

Aina-Cathrine Øvergård [a,b,*], Audun Helge Nerland [a,b] Sonal Patel [a]

[a] Institute of Marine Research, P.O. Box 1870, N-5817 Bergen, Norway

[b] The Gade Institute, University of Bergen, N-5021 Bergen, Norway

* Corresponding author

Telephone: +47 55 90 65 16

Fax: +47 55 23 85 31

Email addresses:

Aina-Cathrine Øvergård: aina-cathrine.oevergaard@imr.no

Audun Helge Nerland: audun.nerland@imr.no

Sonal Patel: sonal.patel@imr.no

Abstract

As known from mammalia, the co-receptors CD4 or CD8 associate with a lymphocyte cell-specific kinase (Lck) upon T-cell activation. Lck phosphorylates tyrosine residues within the CD3 chains, providing docking sites for a 70 kDa zeta-associated-protein (ZAP-70), a tyrosine protein kinase important for T-cell signaling. The sequences of a CD4-like gene (CD4-2), Lck, and ZAP-70 were cloned, characterized, and the relative expression pattern was explored in several organs of Atlantic halibut (*Hippoglossus hippoglossus* L.). Important structural features, as a signal peptide, two Ig-like domains followed by a connecting peptide, a transmembrane region, and a CxC motif within the cytoplasmic tail were conserved within the predicted halibut CD4-2 protein. The deduced halibut Lck protein sequence was found to be composed of a N-terminal Src homology (SH) 4 domain, required for membrane attachment and CD4/CD8 binding, SH3 and SH2 adapter domains, and a SH1 domain followed by a regulatory C-terminal tail (COOH-domain). Tyrosine residues important in Lck activation were conserved within the SH1 and COOH-domain. Structural features of ZAP-70 as tandem SH2 domains and a C-terminal SH1 domain were predicted within the halibut ZAP-70 sequence, having the highest level of conservation within these regions. Several important phosphorylation sites found to play a critical role in T-cell antigen receptor signaling in mammalian were conserved. The overall expression pattern of the three genes was highly similar, showing the highest mRNA level of all three genes in thymus. Some expression was seen in spleen, anterior and posterior kidney, gills, and fin, as seen for other halibut T-cell markers. This study will enable further experiments on halibut T-cell signaling and activation, and enhance understanding about the development of immunological memory T-cells of halibut.

Keywords: lymphocyte cell-specific kinase, 70 kDa zeta-associated-protein, teleost, T-cell

1. Introduction

In mammals, T-cells can be activated when antigenic peptides, presented by major histocompatibility complex (MHC) molecules as peptide MHC (pMHC) complexes, are recognized by the T-cell receptor (TCR). CD4 and CD8 are transmembrane glycoproteins belonging to the Ig superfamily, functioning as co-receptors for the TCR by their interaction with MHC and downstream signal transduction. T-cells can be divided into two major subsets based on the expression of CD4 and CD8, namely $CD4^+CD8^-$ T helper (T_H) cells and $CD4^-$ $CD8^+$ T cytotoxic (T_C) cells. Upon T-cell activation, CD4 or CD8 associates with a lymphocyte cell-specific kinase (Lck), belonging to the Src-family of protein tyrosine kinases (PTKs). The TCR is non-covalently bound to a CD3 complex that carries immune-receptor tyrosine based activation motifs (ITAMs), and Lck phosphorylates tyrosines (Y) within these motifs. This provides docking sites for a cytosolic 70 kDa zeta-associated-protein (ZAP-70), belonging to the Syk/ZAP-70 family of PTKs. The kinase activity of the receptor-associated ZAP-70 becomes activated after tyrosine phosphorylation by Lck, and several tyrosine residues of ZAP-70 becomes phosphorylated creating docking sites for downstream molecules of the activation cascade. The phosphorylation of ZAP-70 occurs either by autophosphorylation both in *cis* and *trans*, or with the help of kinases such as Lck. Also, ZAP-70 directly phosphorylates and activates important signaling molecules of the calcium and Ras pathway.

Generally, the T-cell system in fish is believed to be very similar to the mammalian system, and several molecules involved in T-cell activation have been identified in several fish species. Molecules identified in halibut believed to have preserved their structural characteristics includes RAG1 (FJ769824); TCRα, TCRβ, CD3γδ, CD3ε, CD3ζ [1]; CD8α, CD8β [2]; and CD4 [3]. In fish, different CD4 like molecules have been identified that shows resemblance to CD4 of higher vertebrates [4-10], one form that has the classical four Ig-like domain (D1-D4) structure, another that has three Ig-domains (D1-D3), and a third form that holds only two Ig-domains (D1-D2) followed by a connecting peptide that is believed to be compensating for the two missing Ig-domains. Seemingly, different fish species are in the possession of only two of the three divergent forms of CD4, the classical CD4 and either the three domain form or the two domain form. However, all CD4-like molecules in fish hold a

transmembrane region and a CxC Lck binding motif in their cytoplasmic tails. The presence of a CxC motif makes them structurally different from the lymphocyte activation gene-3 (LAG-3), another member of the CD4 family [11]. Previously described halibut CD4 molecule has the classical four Ig-domain structure [3].

There is little knowledge about T-cell signaling in teleosts, and the activity of Lck during fish T-cell development and activation is not thoroughly investigated. However, there is some evidence that the function of Lck is conserved between higher vertebrates and teleost. In mammals, Lck is composed of a N-terminal Src homology (SH) 4 domain (unique domain) required for membrane attachment and CD4/CD8 binding, SH3 and SH2 adapter domains, and a SH1 catalytic kinase domain followed by a regulatory C-terminal tail, structural characteristics also found in fish Lck [12-14]. And, as in mammals, the expression of fish Lck could be detected in tissues associated with lymphocyte function and development. In rainbow trout (*Oncorhynchus mykiss*), the expression was restricted to cell surface IgM⁻ anterior kidney lymphocytes and was up-regulated in *in vitro* PHA and PMA stimulated lymphocytes [13]. Also, the basic mechanism that regulates the lymphocyte-specific expression of Lck was shown to be conserved between fugu (*Takifugu rubripes*) and mammals [12]. However, studies in fish regarding the binding affinity of Lck to CD4 or CD8, and its ability to phosphorylate downstream molecules as ZAP-70 are limited. Structural features shared by the Syk/ZAP-70 protein-tyrosin kinases family includes tandem SH2 domains, that interacts with phosphorylated ITAMs within the CD3 chains, and a C-terminal kinase domain, also seen in fish ZAP-70 [15]. However, little is known about the action of ZAP-70 in fish T-cell immunity, as this kinase has not been carefully characterized in fish. Here, we report the description of a halibut ZAP-70 gene, in addition to the characterization of a halibut Lck and CD4 like gene having two Ig-domains (here referred to as CD4-2).

2. Materials and methods

2.1 Fish stocks and sample collection

Four individuals, approximately 1 year old weighing between 70 – 150 g, were obtained from Austevoll Aquaculture Reseach station, Norway. They were reared in 9ºC sea water (salinity of 34.5 ‰), and fed commercial feed twice a day. An overdose of benzocain (The Norwegian

medicine depot) was given before tissue sampling. Samples were collected from thymus, spleen, skin, heart, anterior and posterior kidney, pectoral fin, gills, brain, liver, anterior and posterior gut, and white muscle for total RNA isolation. The organ samples were snap-frozen in liquid nitrogen immediately after dissection and stored at –80ºC until use.

2.2 Isolation of total RNA and cDNA synthesis

Total RNA from organ samples was isolated using TRI reagent (Sigma) according to the Trizol reagent protocol described by Invitrogen, with a few modifications as described previously [2]. The concentration and the purity of the total RNA were assessed with a NanoDrop Spectrophotometer (NanoDrop Technologies), and the quality of random samples was analyzed with an Agilent 2100 Bioanalyzer (Agilent Technologies). Total RNA was reverse transcribed using a Reverse Transcription Core Kit (Eurogentec) and random nonamers as primers in 30 μl reactions with 500 ng total RNA. The cDNA was stored at – 20ºC until use.

2.3 DNA sequencing and Bioinformatic analysis

Expressed sequence tags (EST) representing CD4-2 (GenBank accession no.: GE31586), Lck (GenBank accession no.: GE630612), and ZAP-70 (GenBank accession no.: GE630252) were identified by blast search in a database created on the basis of Atlantic halibut cDNA libraries [16]. To sequence the 5'end and 3'end of the genes, the SMART^{TM} RACE cDNA amplification kit from Clontech was used as described previously [3]. Amplification and sequencing of cDNA was performed as to confirm the open reading frame of the genes (GenBank accession no.: CD4-2 - GU985449, Lck - FJ769822, and ZAP-70 - GU985452), and genomic sequences were amplified to confirm exon-exon boundaries, as described previously [2]. Only fragments, sufficient for designing a real time RT-PCR assay not detecting genomic DNA, were amplified for the Lck (GenBank accession no.: FJ769823) and ZAP-70 genes (GenBank accession no.: GU985453). The genomic structure of CD4-2 (GenBank accession no.: GU985448) was however more carefully sequenced.

The open reading frames (ORF) were blasted using ExPASy BLAST form (http://ca.expasy.org/tools/blast/), and aligned with ClustalW (www.ebi.ac.uk/Tools/clustalw/index.html). A phylogenetic tree was constructed by the

neighbor-joining algorithm using ClustalW Multiple Alignment (http://www.bioinformatics.nl/tools/clustalw.html) and 2000 bootstrap replications. Location of domains was predicted using InterProScan (http://www.ebi.ac.uk/InterProScan), and physico-chemical parameters were calculated using ProtParam (http://au.expasy.org/tools/protparam.html). Post translational modifications were predicted using the MYR Predictor (http://mendel.imp.ac.at/myristate/SUPLpredictor.htm), CSS-Palm (http://bioinformatics.lcd-ustc.org/css_palm/prediction.php), NetPhos 2.0 Server (http://www.cbs.dtu.dk/services/NetPhos/), NetOGlyc 3.1 Server (http://www.cbs.dtu.dk/services/NetOGlyc/) and NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/).

2.4 Real time RT-PCR assay and data analysis

Primers and probes for real time RT-PCR were designed with Primer express software 3.0 (Applied Biosystems), according to the manufacturer´s guidelines. Either the probe or one of the primers was designed such that they spanned an exon-exon boundary, as to avoid amplification of genomic DNA. Five-point standard curves of 4-fold dilution series (1:1 – 1:256) from pooled cDNA were used for calculation of the PCR efficiency, given by the equation E% = $(10^{1/\text{slope}} – 1)$ x 100 [17], and for revealing PCR poisoning. The primer and probe sequences with corresponding PCR efficiencies are listed in Table 1.

The PCR reaction mix contained 1x TaqMan Fast PCR Master Mix (Applied Biosystems), 900 nM of each primer, 200 nM TaqMan probe and 1 µl cDNA in a final volume of 12.5 µl. The PCR cycling was carried out as follows: 95ºC for 20 sec, 40 cycles of 95ºC for 1 sec followed by 60ºC for 20 sec. Samples were run in duplicate on the 7900 HT Fast Real-Time PCR System (Applied Biosystems), and the mean Ct value for each sample was used for analysis if the deviation was smaller than 5 %. The real time RT-PCR data was normalized using elongation factor 1 alpha (EF1A1) as internal reference gene [18]. Efficiency of each assay was taken into consideration and the relative expression was transformed using the following formula (Efficiency)$^{-\Delta\Delta Ct}$ [19, 20], calibrated against the liver expression.

3. Results

3.1 Characterization of CD4-2, Lck, and ZAP-70

3.1.1 CD4-2

The halibut CD4-2 cDNA was found to be 1167 base pair (bp), containing an ORF of 924 bp encoding 308 amino acids (aa). The predicted protein shared a sequence identity of 33% with rainbow trout, and 32% with green puffer fish (*Tetraodon nigroviridis*) and salmon (*Salmon salar*) CD4 chains having two Ig-like domains; and 30% with zebrafish (*Danio reio*) and 25% with channel catfish (*Ictalurus punctatus*) CD4 chains having three Ig-like domains. The human CD4 chain showed a sequence identity of 26% with the halibut CD4-2, which is more than the classical four domain CD4 molecule of halibut [3]. Both CD4 molecules (CD4 and CD4-2) grouped with other CD4 molecules in fish and higher vertebrates, and not with LAG-3 (Figure 1A).

Despite the low level of sequence identity between different CD4 molecules, some important structural features were conserved in the predicted halibut CD4-2 protein. A signal peptide is likely to be cleaved at amino acid 18, giving a mature protein of 32 kD. The CD4-2 gene was found to have two Ig-like domains, both having two cysteines for possible formation of disulfide bridges (Figure 2). Also, tryptophan residues believed to stabilize the Ig-like domains were found. The halibut CD4-2 D1 showed a higher sequence identity with D1 of human and halibut CD4 molecules than D3, having an important residue (corresponding to human F^{43}) for MHC interaction. The D2 domain of halibut CD4-2 aligned with the D2 of human CD4 and the D4 of the classical halibut CD4 molecule. The two Ig-like domains were found to be followed by a connecting peptide region having a CxxC motif that may be involved in dimer formation, a region of 23 aa likely to form a transmembrane helix, and a longer cytoplasmic tail compared to the classical CD4 molecules. The CxC motif believed to interact with Lck through a zinc clasp structure [21, 22], was found to be conserved in the cytoplasmic tail of halibut CD4-2. No N-linked glycosylation sites were predicted within the halibut CD4-2 sequence, while three possible O-linked glycosylation sites were found; one within D2 and two within the connecting peptide. A palmitoylation site was predicted in the juxtratransmembrane region.

A genomic sequence of 3984 bp, corresponding to the halibut CD4-2 cDNA sequence, was amplified in overlapping segments and assembled. The halibut CD4-2 gene was found to be separated into 9 exons by 8 introns (Figure 3) with a conserved GT-AG splicing pattern. The first intron divides the 5'UTR over the first two exons. Exon 2 also encodes the first part of the signal peptide, while the second part lies in exon 3. D1 was encoded for by exon 3 and 4, while exon 5 encoded D2. The connection peptide was found within exon 6 and 7. Exon 7 was also found to encode the transmembrane region and the first part of the cytoplasmic tail found to span the entire exon 8 and the first 82 bp of exon 9 that also encoded the 3'UTR.

3.1.2 Lck

The halibut Lck cDNA was found to be 1984 bp, containing an ORF of 1530 bp encoding 509 aa. The predicted protein sequence was estimated to have a molecular weight of about 58 kDa, and shared a sequence identity of 85%, 81%, 76%, and 70% with fugu, rainbow trout, zebrafish, and chicken (*Gallus gallus*) Lck, respectively; and 68 % with human and mouse Lck. The halibut Lck also showed similarities with other Src tyrosine kinase family members, with identities from 59 – 66 % to mammalian and fish non-Lck Src kinases. However, halibut Lck grouped with other fish Lck sequences with high bootstrap support (Figure 1B).

Important structural characteristics within the four Src homology (SH) domains of the Lck gene were identified in the predicted halibut Lck protein (Figure 4). Residues known to be important for lipid modification and membrane localization of Lck in a CD4/CD8 independent manner [23-25], were conserved in the N-terminal unique (SH4) domain of halibut Lck, the domain that distinguishes Lck from other Src-kinases. The post-transcriptional modifications of these motifs were predicted equally to what is seen in mammalian Lck, with myristoylation of halibut G^2 and palmitoylation of C^3 and C^5. Also, the dicysteine motif (CxxC), found to interact through a zinc clasp structure with a corresponding dicysteine motif (CxC) in the CD4 and CD8α chains of human [21, 22], was conserved in the SH4 domain of halibut Lck. The highest degree of conservation was found in the adaptor protein SH2 and SH3 domains and in the catalytic kinase domain. Tyrosine residues believed to be phosphorylated (corresponding to human Y^{394}) and dephosphorylated (corresponding to human Y^{505}) upon Lck activation, were found within the SH1 domain and COOH-domain of halibut Lck sequence, respectively (Figure 4).

3.1.3 ZAP-70

The halibut ZAP-70 cDNA was found to be 2363 bp, containing an ORF of 1839 bp encoding 612 aa. The predicted protein was estimated to have a molecular weight of about 69 kDa, and shared the highest similarity to other ZAP-70 sequences with a sequence identity of 64% to African clawed frog (*Xenopus laevis*) and 63% to human, mouse, and rat. The halibut ZAP-70 sequence did also show similarity to the spleen tyrosine kinase (Syk), with identities of 49 – 51% to mammals and 48% identity to salmon and zebrafish Syk. Phylogenetic analysis indicated, however, that halibut ZAP-70 have a closer relationship to other ZAP-70 sequences of fish and higher vertebrates than to Syk (Figure 1C).

Structural features of ZAP-70 as a tandem SH2 domains and a C-terminal kinase domain were predicted within the halibut ZAP-70 sequence, having the highest level of conservation within these regions (Figure 5). Several of the predicted phosphorylation sites were conserved between halibut and mammals, as the tyrosines corresponding to human Y^{292}, Y^{319}, Y^{492}, Y^{493}, and Y^{598} [26-30], were present in halibut ZAP-70 (Figure 5). However, other tyrosine residues found to play a critical role in T-cell antigen receptor signaling, as Y^{315} and Y^{474} [31, 32], were not seen in the predicted protein sequence of halibut ZAP-70.

3.2 mRNA levels in different organs of one year old halibut

The overall expression pattern of the three genes analyzed by real time RT-PCR was highly similar, showing the highest mRNA level of all three genes in thymus (Figure 6). Some expression was seen in spleen, anterior and posterior kidney, gills, and fin. Also, some expression of the two kinases was observed in skin, however relatively little CD4-2 mRNA in skin compared to liver could be seen.

4. Discussion

Since the first discovery of a non-classical CD4 like gene in teleost, it has been speculated upon which roles the different CD4 forms play in fish immune responses. In many ways, the two-domain CD4 molecule is similar to the mammalian CD4 molecule, of course with the exception of the missing Ig-domains. The genomic organization of the halibut CD4-2 gene shows a close relationship with CD4 and LAG-3 genes, as the first domain D1 was shown to be dispersed between two exons. This is a special feature seen in the CD4 and LAG-3 genes

of higher vertebrates, and found in all the teleost CD4 genes characterized to date [3-8]. Interestingly, a sea lamprey (*Petromyzon marinus*) two-domain CD4-like gene was found to have the D1 domain in a single exon [33], and the three-domain CD4-like gene in channel catfish have both the D1 and D3 domain separated in two exons [5].

An important feature of the D1 of CD4 in mammals is the binding of non-polymorphic residues of the MHC II molecule, bringing Lck to the site of immune recognition. A crystal structure of the human CD4 in complex with MHC II proposed that the major binding energy was provided by the aromatic ring of F^{43} in the mature CD4 molecule surrounded by the hydrophobic residues from the α2 and the β2 domain of MHC class II [34]. Surprisingly, the F^{43} was conserved in halibut CD4-2 (Figure 2), and not in the classical four-domain halibut CD4 [3]. The two-domain CD4 of fugu is also in the possession of a phenylalanine corresponding to human F^{43} [4], giving the consensus FxxK in halibut, fugu, and human but not in other CD4 like genes [3-10]. Other features common in mammalian and fish CD4 Ig-domains are the presence of cysteines for intra-chain disulfide bridging and a core tryptophan forming an essential structural triad thought to stabilize most Ig-domains [35]. Fish non-classical molecules having two or three Ig-domains, are all in the possession of a C-CW triad for possible disulfide bridging in D1 [4-7, 33], as seen in human CD4 [36]. This C-CW triad is also seen in D2 of all teleost CD4-like sequences although the positions are not conserved between fish and human. The first cysteine in human D2 is located noncanonically, and is substituted in many vertebrates. In murine CD4, indications are found for the formation of a disulfide linked dimer formed by domain swapping of D2 giving disulfide-bonds between C^{130} in one monomer and C^{159} in the other one [37]. In human CD4, K^{318} and Q^{344} were identified as important residues for a non-covalent dimer association between D4 of adjacent CD4 molecules [38], and it has been speculated that this aligning of two CD4 monomers gives the opportunity for efficient forming of a covalently disulfide linked CD4 dimers [37]. K^{318} and Q^{344} are poorly conserved in the classical four-domain CD4 molecules of fish (Figure 2). However, the two-domain CD4 in teleost are all in the possession of a CxxC motif within the connecting peptide between D2 and the transmembrane domain, giving the opportunity to form disulfide linked dimers. This can be an important feature of CD4-2, as it has been demonstrated that CD4 dimerization is required for CD4-mediated T-cell activation [38, 39], and it has been indicated that the preferred MHC class II co-receptor is disulfide

linked dimeric CD4 [37]. Also, the connecting peptide region of the two-domain CD4-like molecules can be important for MHC binding, as a high percentage of prolines (about 21% in halibut CD4-2) may allow an extended structure so that the D1 of CD4-2 can reach the MHC molecule.

The cytoplasmic tail of mammalian CD4 is in the possession of residues important for lipid raft localization and internalization. Two palmitoylation sites located in the junction between the transmembrane domain and cytoplasmic tail of CD4 (CxxC) was found to be important for lipid raft localization [40, 41]. This palmitoylation site was found to be poorly conserved in the halibut four-domain CD4-like molecules [3]. Though, in the two-domain CD4-like molecule a single palmitoylation site was predicted in the juxtratransmembrane region indicating a preserved function (Figure 2). A dileucine motif has been found to be essential for endocytosis of CD4 through clathrin-coated pits, however only when S^{408} and S^{415} are phosphorylated [42]. Two serines were indicated to be phosphorylation sites in the cytoplasmic tail of halibut CD4-2, not present in the classical four-domain CD4 of halibut [3]. This is a general trend where the fish two-domain and three-domain CD4-like molecules have possible phosphoserines, but lack the dileucin motifs [4-10, 33], besides the fugu CD4L-2 [4] and the green spotted puffer CD4-2 protein (Figure 2) that both have a dileucine motif with one preceding serine. Interestingly, the chicken CD4 chain lacks the phosphoserines [43], indicating another regulation mechanism in chicken and fish.

In human, Lck is found to interact through a dicysteine motif (CxxC) within the SH4 domain, forming a zinc clasp structure with a corresponding dicysteine motif (CxC) in the cytoplasmic tails of CD4 and CD8α [21]. This motif is conserved in all CD4-like molecules [3-10] and Lck sequences [12-14] characterized in teleost, including halibut CD4-2 and Lck. Interestingly, a third cycteine is seen in all the teleost Lck sequences giving the motif CxxCxC, followed by a proline. In human Lck, the CxxC motif is positioned for metal coordination by a β hairpin [21], and the proline that is also seen in the mammalian Lck homologues, could be conserved as to give this β hairpin a kink. The interaction between human Lck and CD4 was shown to include short amphipathic α- helixes at the N-terminal of the CxC motifs of both molecules packed together with a hydrophobic interface [21]. The different CD4-like molecules in fish are, as mentioned before, in the possession of a CxC

motif, but with different surroundings. The classical four-domain CD4 molecules have a relatively short cytoplasmic tail, with a large number of basic residues. The CD4-like molecules with two domains have a longer cytoplasmic tail with amino acid residues in a pattern that are suited for the formation of an amphipatic α helix as seen in human (Figure 2). The Lck of halibut and other teleosts is also in possession of residues that can make up a short amphipatic α helix in a similar pattern as seen in human Lck (Figure 4). The interface between human CD8α and Lck was found to be much smaller than that seen between CD4 and Lck, but with similar affinity likely as the metal coordination contributes to a large part of the binding energy [21]. In fish, the CxC motif of CD8α is found as CxH in both CD8α and CD8β, and it has been suggested that both molecules would be able to interact with Lck, with contribution from histidine for metal coordination [2, 44]. Other feature of the unique domain of Lck is the presence of a GCxCS motif. This motif is believed to be important for membrane localization of Lck in a CD4/CD8 independent manner, due to palmiotylation of C^3 and C^5 dependent on the myristoylation of G^2 [23-25]. This motif was identified within the halibut Lck sequences (Figure 4), indicating a preserved lipid raft localization strategy between halibut Lck and Lck of higher vertebrates.

The regulatory SH3 and SH2 protein interaction domains and the SH1 kinase domain of halibut Lck shows the highest sequence conservation (Figure 4), naturally as these domains comprises several important functions of Lck. Soon after T-cell activation, tyrosin residues within numerous of proteins involved in T-cell signaling are phosphorylated, many dependent on the presence of Lck. Naturally a close regulation of the Lck kinase activity is important. A regulatory tyrosin (corresponding to human Y^{505}) within the COOH-domain of mammalian Lck is essential in this regulation, phosphorylated by a C-terminal Src kinase (Csk) [45], and dephosphorylated by the phosphatase CD45 [46-48]. In the Src family of PTKs the phosphorylated form of the regulatory tyrosin (corresponding to Y^{505} in Lck) binds to the SH2 domain, keeping the kinase in an inactive conformation, helped by the association between the SH3 domain and a PxxP motif within the linker sequence [49]. The kinase domain becomes accessible again as CD45 dephosphorylates Y^{505} and it dissociates from SH2 [46, 48]. A tyrosin within the activation loop (corresponding to human Y^{394}) becomes phosphorylated in the active form of the kinase, mainly by autophosphorylation [50], causing a conformational change essential for its kinase activity. This mechanism for regulation is

likely to be preserved between mammals and fish, as these tyrosine residues are found in the fish Lck sequences including halibut Lck (Figure 4). Also, the PxxP motif within the linker is seen in fish Lck (Figure 4), however, the consensus sequence of the regulatory COOH-tail, TATExQYQxQ/G, is not conserved in halibut Lck. The last glutamine (Q) in halibut Lck is substituted by an aspartic acid (D), and for unknown reasons the tail is five amino acids longer than any other Lck sequence available. Probably, this would not affect the regulatory function of the tail, as this is an important feature of the kinase. Another mechanism for down-regulation is proposed by the phosphorylation of serine residues, especially S^{59} as a substrate for the mitogen-activated protein kinase (MAPK) [51-53]. After T-cell stimulation, the phosphorylation of S^{59} was found to decrease enzyme activity and regulate SH2 binding specificity [51, 54]. S^{59} is found within the unique domain of Lck as a part of a MAPK target sequence (PxPS), and is conserved between mammals and fish (Figure 4). Considering the low level of conservation otherwise in the unique domain, the regulatory function of this serine is probably preserved.

An important task of Lck is the phosphorylation of ITAMs within the CD3 chains upon T-cell activation, creating docking sites for the tandem SH2 domains of ZAP-70 [55, 56]. The ITAMs within the CD3 chains of halibut and other fish species confirms mainly to the consensus [1, 57-63], indicating a preserved binding mechanism of ZAP-70 between mammals and fish. The high conservation of the tandem SH2 domains of halibut ZAP-70 (Figure 5) agrees with this. When ZAP-70 binds to the phosphorylated ITAMs of the CD3 chains, its SH1 kinase domain becomes activated. Further, ZAP-70 can be phosphorylated creating docking sites for other SH2-containing proteins [64], or directly phosphorylate substrates as LAT and SPL-76 and thereby activate signaling pathways such as the calcium and Ras pathways [65, 66]. The phosphorylation of ZAP-70 Y^{493} by Lck is required for the activation of the ZAP-70 catalytic activity [26-28]. Y^{493} is conserved within the activation loop of halibut ZAP-70, indicating a conserved mechanism for regulation through evolution. The SH2 domain of Lck is required for efficient TCR-mediated signaling in higher vertebrates [67], as it is found to interact with ZAP-70 following T-cell stimulation [68]. The SH2 domain of Lck was found to use the sequence Y^{319}SDP within the interdomain B of ZAP-70 for binding [69, 70], and mutation analysis identified Y^{319} as an important regulator for TCR-dependent phospholipase Cγ1 (PLCγ1) and Ras activation [30, 70]. Mutational

analysis of Y^{315} within the interdomain B of ZAP-70 showed that this tyrosine binds to a SH2 containing guanine nucleotide exchange factor Vav [71], and also allosterically regulates the interaction of the ZAP-70 SH2 domain to ITAMs [31]. Both tyrosines are followed by a proline at the +3 position (Y^{315}ESPY^{319}SDP). The interdomain B is the region showing the least conservation between fish and higher vertebrates, and the YESPYSDP sequence is not preserved in the halibut ZAP-70 (Figure 5). However, a possible tyrosine phosphorylation site is seen within the same region of the interdomain B of fish ZAP-70 having a proline at position +3 (Y^{308}HNP in halibut). The phosphorylation of halibut ZAP-70 Y^{308} can thereby have a similar function as Y^{315} and/or Y^{319}. In mammals, the engagement of a CTL-associated Ag-4 (CTLA-4) was suggested to be important for T-cell suppression, as it inhibits the phosphorylation of Y^{319} and the binding of Lck to ZAP-70 [72], emphasizing the necessity of such a regulatory tyrosin within the interdomain B of halibut ZAP-70. Also, other tyrosine residues found to be important for down-regulation of antigen receptor function, Y^{292} and Y^{492} [27, 28], are conserved within the halibut ZAP-70 sequence. Phosphorylated Y^{292} have been found to interact with the negative regulator protein Cbl [73], hence the identification and binding analysis of halibut Cbl may support the importance of this tyrosine residue in fish ZAP-70 regulation.

Another identified regulatory phosphotyrosine, Y^{474} within the SH1 kinase domain of human ZAP-70, is not seen within halibut or green spotted puffer ZAP-70 (Figure 5), a tyrosine residue found to bind to a phosphotyrosine binding domain within the Shc adaptor protein [32]. Evidence was found that the interaction between ZAP-70 and Shc lead to the phosphorylation of the Gbr2 binding site of Shc, which thereby couples the activated TCR to the Ras activation pathway. The functional consequence of the substitution of Y^{474} with phenylalanine within halibut and green spotted puffer ZAP-70 can be interesting for further study, as Y^{474} is preserved in zebrafish ZAP-70, and in zebrafish [15], salmon (accession no.: BT059215), common carp (*Cyprinus carpio*; accession no.: AF253045) and green spotted puffer (accession no.: CAF96564.1) Syk sequences. Other conserved tyrosine residues as Y^{69}, Y^{126}, and Y^{178} (Figure 5) might function as a binding sites for Ras activators in halibut, predicted to be sites for phosphorylation in halibut as well as in mammals [74]. However, the

functionality of these predicted phosphotyrosines should be carefully evaluated before any conclusions can be drawn.

Real time RT-PCR analysis revealed a highly similar expression pattern of the halibut CD4-2, Lck, and ZAP-70 genes (Figure 6). Compared to the expression levels in other organs, the halibut thymus was found to have a relatively high mRNA level of the three genes, which is expected as it is believed to be the site of thymocyte maturation [75]. Next to thymus, some expression was seen in spleen, anterior and posterior kidney, gills, and fin; similar to what is seen for rainbow trout Lck [13] and different CD4-like molecules of several fish species [4-10]. Also, a highly similar expression patterns is seen in halibut transcription of the CD4-like molecule containing four Ig-like domains [3], as well as CD8α, CD8β [2], TCRα, TCRβ, CD3γδ, CD3ε, and CD3ζ [1], confirming the functional relationship of the genes as T-cell markers and signaling molecules. The evaluation of the ZAP-70 relative expression pattern is very important in the characterization of this kinase, as structural features defining ZAP-70 are likely to be shared by both ZAP-70 and Syk, as seen in other vertebrates [15, 76]. In mammals, ZAP-70 is found to be exclusively expressed in all major thymocyte populations and NK cells [55, 77]. Syk on the other hand, is found to be most abundantly expressed in B-cells [77, 78], and the expression pattern of a halibut Syk gene would therefore be expected to be more similar to the IgM expression pattern of halibut [79]. The expression pattern of halibut ZAP-70 is thereby a confirmation of its homology, in addition to its high sequence conservation and the phylogenetic analysis grouping the halibut ZAP-70 sequence together with other known ZAP-70 sequences with high bootstrap support (Figure 1C).

5. Conclusion

Here we report the cloning and characterization of the genes encoding CD4-2, Lck, and ZAP-70 from Atlantic halibut, believed to be essential in T-cell signaling. The sequence data suggest that the three genes are involved in T-cell immunity in teleost fish, as structural features important for T-cell signaling are preserved. The tissue expression pattern of the three genes was highly similar, also to other T-cell markers previously reported in Atlantic halibut, likely to be related to functional assets of T-cells. This study facilitates further research on halibut T-cell signaling and activation, and eases further studies involving halibut immune responses against pathogenic stimuli and vaccines.

Acknowledgements

Reference list

[1] Øvergård AC, Hordvik I, Nerland AH, Eikeland G, Patel S. Cloning and expression analysis of Atlantic halibut (*Hippoglossus hippoglossus*) CD3 genes. Fish Shellfish Immunol 2009; 27:707-13.

[2] Patel S, Øvergård AC, Nerland AH. CD8α and CD8β in Atlantic halibut, *Hippoglossus hippoglossus*: Cloning, characterization and gene expression during viral and bacterial infection. Fish Shellfish Immunol 2008; 25:570-80.

[3] Patel S, Øvergård AC, Nerland AH. A CD4 homologue in Atlantic halibut (*Hippoglossus hippoglossus*): Molecular cloning and characterisation. Fish Shellfish Immunol 2009; 26:377-84.

[4] Dijkstra JM, Somamoto T, Moore L, Hordvik I, Ototake M, Fischer U. Identification and characterization of a second CD4-like gene in teleost fish. Mol Immunol 2006; 43:410-9.

[5] Edholm ES, Stafford JL, Quiniou SM, Waldbieser G, Miller NW, Bengten E, et al. Channel catfish, *Ictalurus punctatus*, CD4-like molecules. Dev Comp Immunol 2007; 31:172-87.

[6] Laing KJ, Zou JJ, Purcell MK, Phillips R, Secombes CJ, Hansen JD. Evolution of the CD4 family: teleost fish possess two divergent forms of CD4 in addition to lymphocyte activation gene-3. J Immunol 2006; 177:3939-51.

[7] Moore LJ, Dijkstra JM, Koppang EO, Hordvik I. CD4 homologues in Atlantic salmon. Fish Shellfish Immunol 2009; 26:10-8.

[8] Suetake H, Araki K, Suzuki Y. Cloning, expression, and characterization of fugu CD4, the first ectothermic animal CD4. Immunogenetics 2004; 56:368-74.

[9] Buonocore F, Randelli E, Casani D, Guerra L, Picchietti S, Costantini S, et al. A CD4 homologue in sea bass (*Dicentrarchus labrax*): Molecular characterization and structural analysis. Mol Immunol 2008; 45:3168-77.

[10] Sun XF, Shang N, Hu W, Wang YP, Guo QL. Molecular cloning and characterization of carp (*Cyprinus carpio* L.) CD8β and CD4-like genes. Fish Shellfish Immunol 2007; 23:1242-55.

[11] Triebel F, Jitsukawa S, Baixeras E, Roman-Roman S, Genevee C, Viegas-Pequignot E, et al. LAG-3, a novel lymphocyte activation gene closely related to CD4. J Exp Med 1990; 171:1393-405.

[12] Brenner S, Venkatesh B, Yap WH, Chou CF, Tay A, Ponniah S, et al. Conserved regulation of the lymphocyte-specific expression of lck in the Fugu and mammals. Proc Natl Acad Sci 2002; 99:2936-41.

[13] Laing KJ, Dutton S, Hansen JD. Molecular and biochemical analysis of rainbow trout LCK suggests a conserved mechanism for T-cell signaling in gnathostomes. Mol Immunol 2007; 44:2737-48.

[14] Langenau DM, Ferrando AA, Traver D, Kutok JL, Hezel JP, Kanki JP, et al. *In vivo* tracking of T cell development, ablation, and engraftment in transgenic zebrafish. Proc Natl Acad Sci 2004; 101:7369-74.

[15] Christie TL, Carter A, Rollins EL, Childs SJ. Syk and Zap-70 function redundantly to promote angioblast migration. Dev Biol 2010; 340:22-9.

[16] Patel S, Malde K, Lanzen A, Olsen RH, Nerland AH. Identification of immune related genes in Atlantic halibut (*Hippoglossus hippoglossus* L.) following *in vivo* antigenic and *in vitro* mitogenic stimulation. Fish Shellfish Immunol 2009; 27:729-38.

[17] Radonic A, Thulke S, Mackay IM, Landt O, Siegert W, Nitsche A. Guideline to reference gene selection for quantitative real-time PCR. Biochem Biophys Res Commun 2004; 313:856-62.

[18] Øvergård AC, Nerland AH, Patel S. Evaluation of potential reference genes for real time RT-PCR studies in Atlantic halibut (*Hippoglossus Hippoglossus* L.); during development, in tissues of healthy and NNV-injected fish, and in anterior kidney leucocytes. BMC Mol Biol 2010; 11:36.

[19] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-DDCT} Method. Methods 2001; 25:402-8.

[20] Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001; 29:e45.

[21] Kim PW, Sun ZY, Blacklow SC, Wagner G, Eck MJ. A zinc clasp structure tethers Lck to T cell coreceptors CD4 and CD8. Science 2003; 301:1725-8.

[22] Lin RS, Rodriguez C, Veillette A, Lodish HF. Zinc is essential for binding of p56[lck] to CD4 and CD8alpha. J Biol Chem 1998; 273:32878-82.

[23] Yurchak LK, Sefton BM. Palmitoylation of either Cys-3 or Cys-5 is required for the biological activity of the Lck tyrosine protein kinase. Mol Cell Biol 1995; 15:6914-22.

[24] Yasuda K, Kosugi A, Hayashi F, Saitoh S, Nagafuku M, Mori Y, et al. Serine 6 of Lck tyrosine kinase: a critical site for Lck myristoylation, membrane localization, and function in T lymphocytes. J Immunol 2000; 165:3226-31.

[25] Kabouridis PS, Magee AI, Ley SC. S-acylation of LCK protein tyrosine kinase is essential for its signalling function in T lymphocytes. Embo J 1997; 16:4983-98.

[26] Chan AC, Dalton M, Johnson R, Kong GH, Wang T, Thoma R, et al. Activation of ZAP-70 kinase activity by phosphorylation of tyrosine 493 is required for lymphocyte antigen receptor function. EMBO J 1995; 14:2499-508.

[27] Wange RL, Guitian R, Isakov N, Watts JD, Aebersold R, Samelson LE. Activating and inhibitory mutations in adjacent tyrosines in the kinase domain of ZAP-70. J Biol Chem 1995; 270:18730-3.

[28] Kong G, Dalton M, Bubeck Wardenburg J, Straus D, Kurosaki T, Chan AC. Distinct tyrosine phosphorylation sites in ZAP-70 mediate activation and negative regulation of antigen receptor function. Mol Cell Biol 1996; 16:5026-35.

[29] Zeitlmann L, Knorr T, Knoll M, Romeo C, Sirim P, Kolanus W. T cell activation induced by novel gain-of-function mutants of Syk and ZAP-70. J Biol Chem 1998; 273:15445-52.

[30] Di Bartolo V, Mege D, Germain V, Pelosi M, Dufour E, Michel F, et al. Tyrosine 319, a newly identified phosphorylation site of ZAP-70, plays a critical role in T cell antigen receptor signaling. J Biol Chem 1999; 274:6285-94.

[31] Di Bartolo V, Malissen M, Dufour E, Sechet E, Malissen B, Acuto O. Tyrosine 315 determines optimal recruitment of ZAP-70 to the T cell antigen receptor. Eur J Immunol 2002; 32:568-75.

[32] Pacini S, Ulivieri C, Di Somma MM, Isacchi A, Lanfrancone L, Pelicci PG, et al. Tyrosine 474 of ZAP-70 is required for association with the Shc adaptor and for T-cell antigen receptor-dependent gene activation. J Biol Chem 1998; 273:20487-93.

[33] Pancer Z, Mayer WE, Klein J, Cooper MD. Prototypic T cell receptor and CD4-like coreceptor are expressed by lymphocytes in the agnathan sea lamprey. Proc Natl Acad Sci 2004; 101:13273-8.

[34] Wang JH, Meijers R, Xiong Y, Liu JH, Sakihama T, Zhang R, et al. Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule. Proc Natl Acad Sci 2001; 98:10799-804.

[35] Ioerger TR, Du C, Linthicum DS. Conservation of cys-cys-trp structural triads and their geometry in the protein domains of immunoglobulin superfamily members. Mol Immunol 1999; 36:373-86.

[36] Maddon PJ, Molineaux SM, Maddon DE, Zimmerman KA, Godfrey M, Alt FW, et al. Structure and expression of the human and mouse T4 genes. Proc Natl Acad Sci 1987; 84:9155-9.

[37] Maekawa A, Schmidt B, Fazekas de St Groth B, Sanejouand YH, Hogg PJ. Evidence for a domain-swapped CD4 dimer as the coreceptor for binding to class II MHC. J Immunol 2006; 176:6873-8.

[38] Moldovan MC, Yachou A, Levesque K, Wu H, Hendrickson WA, Cohen EA, et al. CD4 dimers constitute the functional component required for T cell activation. J Immunol 2002; 169:6261-8.

[39] Moldovan MC, Sabbagh L, Breton G, Sekaly RP, Krummel MF. Triggering of T cell activation via CD4 dimers. J Immunol 2006; 176:5438-45.

[40] Crise B, Rose JK. Identification of Palmitoylation Sites on Cd4, the Human-Immunodeficiency-Virus Receptor. J Biol Chem 1992; 267:13593-7.

[41] Fragoso R, Ren DJ, Zhang XP, Su MWC, Burakoff SJ, Jin YJ. Lipid raft distribution of CD4 depends on its palmitoylation and association with Lck, and evidence for CD4-induced lipid raft aggregation as an additional mechanism to enhance CD3 signaling. J Immunol 2003; 170:913-21.

[42] Pitcher C, Honing S, Fingerhut A, Bowers K, Marsh M. Cluster of differentiation antigen 4 (CD4) endocytosis and adaptor complex binding require activation of the CD4 endocytosis signal by serine phosphorylation. Mol Biol Cell 1999; 10:677-91.

[43] Koskinen R, Lamminmaki U, Tregaskes CA, Salomonsen J, Young JR, Vainio O. Cloning and modeling of the first nonmammalian CD4. J Immunol 1999; 162:4115-21.

[44] Moore LJ, Somamoto T, Lie KK, Dijkstra JM, Hordvik I. Characterisation of salmon and trout CD8α and CD8β. Mol Immunol 2005; 42:1225-34.

[45] Bergman M, Mustelin T, Oetken C, Partanen J, Flint NA, Amrein KE, et al. The human p50csk tyrosine kinase phosphorylates p56lck at Tyr-505 and down regulates its catalytic activity. EMBO J 1992; 11:2919-24.

[46] Mustelin T, Coggeshall KM, Altman A. Rapid activation of the T-cell tyrosine protein kinase pp56lck by the CD45 phosphotyrosine phosphatase. Proc Natl Acad Sci 1989; 86:6302-6.

[47] Ostergaard HL, Shackelford DA, Hurley TR, Johnson P, Hyman R, Sefton BM, et al. Expression of CD45 alters phosphorylation of the lck-encoded tyrosine protein kinase in murine lymphoma T-cell lines. Proc Natl Acad Sci 1989; 86:8959-63.

[48] Mustelin T, Altman A. Dephosphorylation and activation of the T cell tyrosine kinase pp56lck by the leukocyte common antigen (CD45). Oncogene 1990; 5:809-13.

[49] Xu W, Harrison SC, Eck MJ. Three-dimensional structure of the tyrosine kinase c-Src. Nature 1997; 385:595-602.

[50] Veillette A, Fournel M. The CD4 associated tyrosine protein kinase p56lck is positively regulated through its site of autophosphorylation. Oncogene 1990; 5:1455-62.

[51] Watts JD, Sanghera JS, Pelech SL, Aebersold R. Phosphorylation of serine 59 of p56lck in activated T cells. J Biol Chem 1993; 268:23275-82.

[52] Winkler DG, Park I, Kim T, Payne NS, Walsh CT, Strominger JL, et al. Phosphorylation of Ser-42 and Ser-59 in the N-terminal region of the tyrosine kinase p56lck. Proc Natl Acad Sci 1993; 90:5176-80.

[53] Schroder AJ, Quehl P, Muller J, Samstag Y. Conversion of p56^{lck} to p60^{lck} in human peripheral blood T lymphocytes is dependent on co- stimulation through accessory receptors: involvement of phospholipase C, protein kinase C and MAP-kinases in vivo. Eur J Immunol 2000; 30:635-43.

[54] Park I, Chung J, Walsh CT, Yun Y, Strominger JL, Shin J. Phosphotyrosine-independent binding of a 62-kDa protein to the src homology 2 (SH2) domain of p56lck and its regulation by phosphorylation of Ser-59 in the lck unique N-terminal region. Proc Natl Acad Sci 1995; 92:12338-42.

[55] Chan AC, Iwashima M, Turck CW, Weiss A. ZAP-70: a 70 kd protein-tyrosine kinase that associates with the TCR zeta chain. Cell 1992; 71:649-62.

[56] Iwashima M, Irving BA, van Oers NS, Chan AC, Weiss A. Sequential interactions of the TCR with two distinct cytoplasmic tyrosine kinases. Science 1994; 263:1136-9.

[57] Alabyev BY, Guselnikov SV, Najakshin AM, Mechetina LV, Taranin AV. CD3ε homologues in the chondrostean fish *Acipenser ruthenus*. Immunogenetics 2000; 51:1012-20.

[58] Araki K, Suetake H, Kikuchi K, Suzuki Y. Characterization and expression analysis of CD3ε and CD3γ/δ in fugu, *Takifugu rubripes*. Immunogenetics 2005; 57:158-63.

[59] Liu Y, Moore L, Olaf Koppang E, Hordvik I. Characterization of the CD3ζ, CD3γδ and CD3ε subunits of the T cell receptor complex in Atlantic salmon. Dev Comp Immunol 2008; 32:26-35.

[60] Park CI, Hirono I, Aoki T. Molecular characterization of the Japanese flounder, *Paralichthys olivaceus*, CD3epsilon and evolution of the CD3 cluster. Dev Comp Immunol 2005; 29:123-33.

[61] Park CI, Hirono I, Enomoto J, Nam BH, Aoki T. Cloning of Japanese flounder *Paralichthys olivaceus* CD3 cDNA and gene, and analysis of its expression. Immunogenetics 2001; 53:130-5.

[62] Shang N, Sun XF, Hu W, Wang YP, Guo QL. Molecular cloning and characterization of common carp (*Cyprinus carpio* L.) TCRγ and CD3γ/δ chains. Fish Shellfish Immunol 2008; 24:412-25.

[63] Guselnikov SV, Najakshin AM, Taranin AV. Fugu rubripes possesses genes for the entire set of the ITAM-bearing transmembrane signal subunits. Immunogenetics 2003; 55:472-9.

[64] Neumeister EN, Zhu Y, Richard S, Terhorst C, Chan AC, Shaw AS. Binding of ZAP-70 to phosphorylated T-cell receptor zeta and eta enhances its autophosphorylation and generates specific binding sites for SH2 domain-containing proteins. Mol Cell Biol 1995; 15:3171-8.

[65] Bubeck Wardenburg J, Fu C, Jackman JK, Flotow H, Wilkinson SE, Williams DH, et al. Phosphorylation of SLP-76 by the ZAP-70 protein-tyrosine kinase is required for T-cell receptor function. J Biol Chem 1996; 271:19641-4.

[66] Zhang W, Sloan-Lancaster J, Kitchen J, Trible RP, Samelson LE. LAT: the ZAP-70 tyrosine kinase substrate that links T cell receptor to cellular activation. Cell 1998; 92:83-92.

[67] Caron L, Abraham N, Pawson T, Veillette A. Structural requirements for enhancement of T-cell responsiveness by the lymphocyte-specific tyrosine protein kinase p56lck. Mol Cell Biol 1992; 12:2720-9.

[68] Duplay P, Thome M, Herve F, Acuto O. p56lck interacts via its src homology 2 domain with the ZAP-70 kinase. J Exp Med 1994; 179:1163-72.

[69] Pelosi M, Di Bartolo V, Mounier V, Mege D, Pascussi JM, Dufour E, et al. Tyrosine 319 in the interdomain B of ZAP-70 is a binding site for the Src homology 2 domain of Lck. J Biol Chem 1999; 274:14229-37.

[70] Williams BL, Irvin BJ, Sutor SL, Chini CC, Yacyshyn E, Bubeck Wardenburg J, et al. Phosphorylation of Tyr319 in ZAP-70 is required for T-cell antigen receptor-dependent phospholipase C-gamma1 and Ras activation. EMBO J 1999; 18:1832-44.

[71] Wu J, Zhao Q, Kurosaki T, Weiss A. The Vav binding site (Y315) in ZAP-70 is critical for antigen receptor-mediated signal transduction. J Exp Med 1997; 185:1877-82.

[72] Guntermann C, Alexander DR. CTLA-4 suppresses proximal TCR signaling in resting human CD4$^+$ T cells by inhibiting ZAP-70 Tyr^{319} phosphorylation: a potential role for tyrosine phosphatases. J Immunol 2002; 168:4420-9.

[73] Lupher ML, Jr., Songyang Z, Shoelson SE, Cantley LC, Band H. The Cbl phosphotyrosine-binding domain selects a D(N/D)XpY motif and binds to the Tyr292 negative regulatory phosphorylation site of ZAP-70. J Biol Chem 1997; 272:33140-4.

[74] Watts JD, Affolter M, Krebs DL, Wange RL, Samelson LE, Aebersold R. Identification by electrospray ionization mass spectrometry of the sites of tyrosine phosphorylation induced in activated Jurkat T cells on the protein tyrosine kinase ZAP-70. J Biol Chem 1994; 269:29520-9.

[75] Hansen JD, Zapata AG. Lymphocyte development in fish and amphibians. Immunol Rev 1998; 166:199-220.

[76] van Oers NS, Weiss A. The Syk/ZAP-70 protein tyrosine kinase connection to antigen receptor signalling processes. Semin Immunol 1995; 7:227-36.

[77] Chan AC, van Oers NS, Tran A, Turka L, Law CL, Ryan JC, et al. Differential expression of ZAP-70 and Syk protein tyrosine kinases, and the role of this family of protein tyrosine kinases in TCR signaling. J Immunol 1994; 152:4758-66.

[78] Law CL, Sidorenko SP, Chandran KA, Draves KE, Chan AC, Weiss A, et al. Molecular cloning of human Syk. A B cell protein-tyrosine kinase associated with the surface immunoglobulin M-B cell receptor complex. J Biol Chem 1994; 269:12310-9.

[79] Patel S, Sørhus E, Fiksdal IU, Espedal PG, Bergh Ø, Rødseth OM, et al. Ontogeny of lymphoid organs and development of IgM-bearing cells in Atlantic halibut (*Hippoglossus hippoglossus* L.). Fish Shellfish Immunol 2009; 26:385-95.

Tables

Table 1

Primers and probes for real time RT-PCR and *in situ* hybridisation. Abbreviations: F – forward primer, R – reverse primer, P – probe, is – *in situ* hybridisation, 6FAM – 6-carboxyfluorescein, BHQ – black hole quencher, MGBNFQ – Minor groove binder non-fluorescent quencher.

Name	Sequence 5' → 3'	Accesion nr	E%
EF1A1 F	CCATGGTTGTGGAGTCCTTCTC	EU561357	96
EF1A1 R	GATGACACCGACAGCCACTGT		
EF1A1 P	6FAM – CTCCCCTCGGTCGTTTCGCTGTG - BHQ		
CD4-2 F	CGAGGTCCAAGACGTACTGTCA	GU985449	90
CD4-2 R	CGTCCTGGCTGTGGTTTCTC		
CD4-2 P	VIC - TGTAACTGCCGTCCAGC - MGBNFQ		
Lck F	TGCCACTTCCCTGTCATTATTG	FJ769822	100
Lck R	CTTGGAGCAGGACAGTTTGGA		
Lck P	TATGGATGGGAGAATACAAT		
Lck is F	AACTGCAGTTCGGATTATTCAGACA		
Lck is R	CCCAAAGGACCACACATCTGATTT		
ZAP-70 F	TTCCTCAGCAGCAACAAGGATAC	GU985452	98
ZAP-70 R	CCAGATACTTCATCCCCATCGA		
ZAP-70 P	6FAM - CACTGCAGAGAACATT - MGBNFQ		

Figures
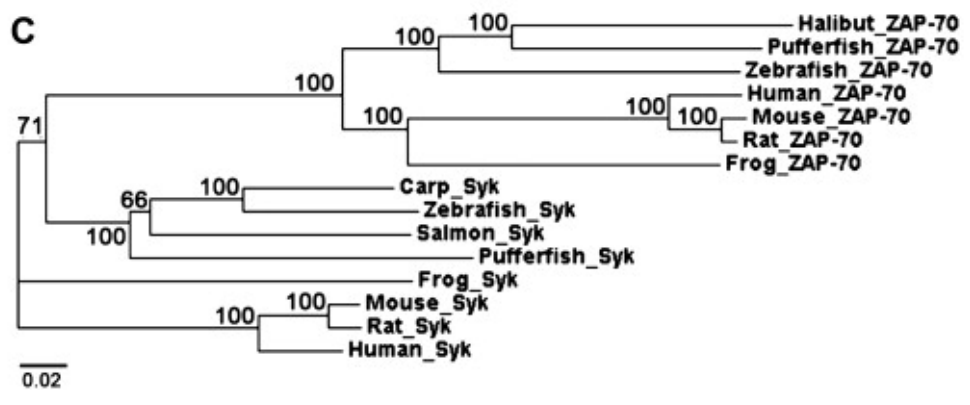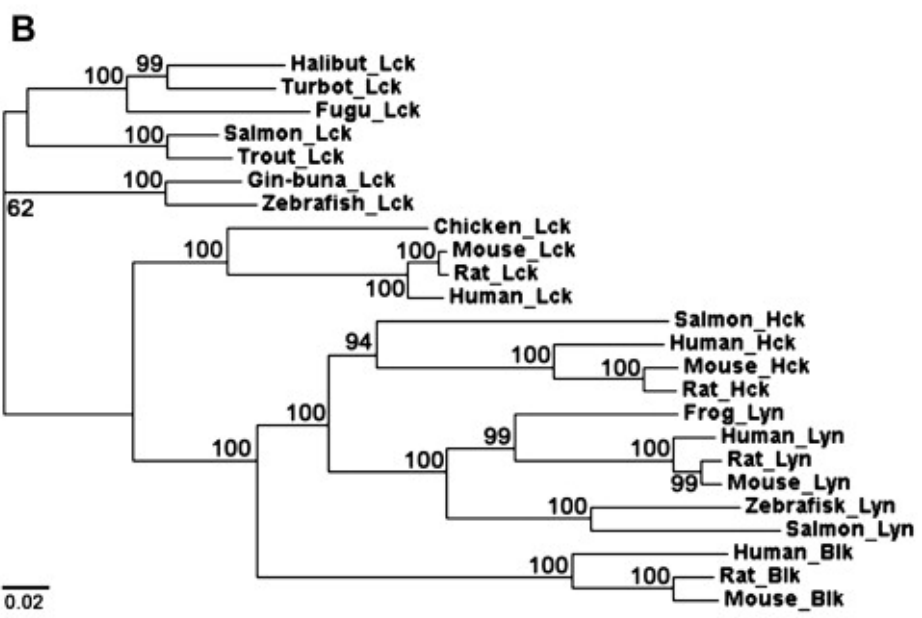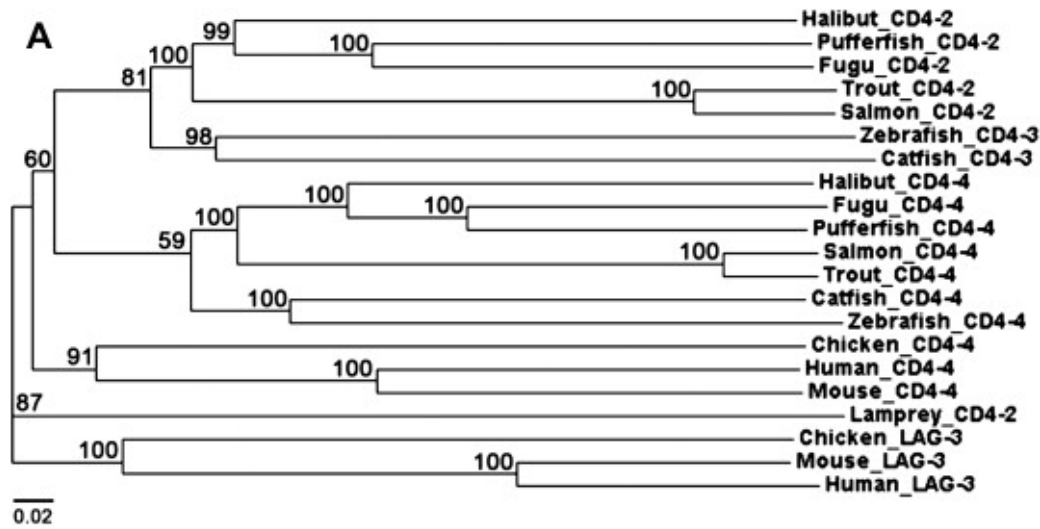
Figure 1

Phylogenetic tree, constructed by the neighbor-joining algorithm using ClustalW Multiple Alignment. All CD4 homologues with two Ig-like domains are designated CD4-2, with three Ig-like domains are designated CD4-3, and with four Ig-like domains are designated CD4-4. Bootstrap values above 50 % are shown at the branches, and the branch length scale in terms of genetic distance is indicated below the trees. Accession numbers are written in parenthesis. The three is showing the relationship between **A)** CD4 and LAG-3 molecules of halibut (*Hippoglossus hippoglossus*) CD4-2 (GU985449) and CD4-4 (FJ185042); green spotted puffer (*Tetraodon nigroviridis*) CD4-2 (EF601918) and CD4-4 (EF601919): rainbow trout (*Oncorhynchus mykiss*) CD4-2 (AY772711) and CD4-4 (AY973028); salmon (*Salmo salar*) CD4-2 (BT056594) and CD4-4 (EU409794); fugu (*Takifugu rubripes*) CD4-2 [4] and CD4-4 (AB164055); sea lamprey (*Petromyzon marinus*) CD4-2 (AY686862); zebrafish (*Danio rerio*) CD4-3 (EF601915) and CD4-4 (EF601917); channel catfish (*Ictalurus punctatus*) CD4-3 (DQ435304) and CD4-4 (DQ435301); chicken (*Gallus gallus*) CD4-4 (AY560013) and LAG-3 (XP_416510); mouse (*Mus musculus*) CD4-4 (M36850) and LAG-3 (X98113); and human (*Homo sapiens*) CD4-4 (M12807) and LAG-3 (X51985), **B)** Halibut Lck (FJ769822); turbot (*Psetta maxima*) Lck (DQ848967); fugu Lck (AF411956); salmon Lck (BT044859), Hck (AF321110), and Lyn (BT045857); rainbow trout Lck (AY973032); gin-buna (*Carassius auratus langsdorfii*) Lck (AB279594); zebrafish Lck (BC115230), and Lyn (BC081601); chicken Lck (BC115230); mouse Lck (M12056), Hck (Y00487), Lyn (M64608), and Blk (BC030668); human Lck (X05027), Hck (M16592), Lyn (M16038), and Blk (Z33998); rat (*Rattus norvegicus*) Lck (BC160881), Hck (BC078890), Lyn (L14951), and Blk (BC098683); and frog (*Xenopus laevis*) Lyn (BC170242), and **C)** Halibut ZAP-70 (GU985452); green spotted puffer ZAP-70 (CAG00734) and Syk (CAAE01014482); zebrafish ZAP-70 (NP_001018425) and Syk (AF253046); frog ZAP-70 (BC077883) and Syk (BC077278); mouse ZAP-70 (U04379) and Syk (U25685); human ZAP-70 (L05148) and Syk (Z29630); rat ZAP-70 (BC089855) and Syk (U21684); and common carp (*Cyprinus carpio*) Syk (AF253045).

```
                     SIGNAL PEPTIDE              D1
HAL CD4-2  MNVI-----VCLGFVLGALSAAA---TVFHTKPRQRVTLECG-VDSYSRSVAWSHR---DELIIQIDKKSGFPRK---G  64
PUF CD4-2  MKKIGRLEFAFLLAVLSCAPPADG--DVIVGKPGGRVAFRCG-LPRGGGPLGWYHQ---ENVVYREDARTGIPQR---G  69
TRO CD4-2  MKTL-----SWFVFALCILHVVG---EVIYKRIGLPVNIDCG-VKTSNKDMEWSHKAVGGSKSVIVDYFGKNGKERKG  70
HAL CD4    MERF----VLILFPALISTSGA----KSVYAQVGGRVVLKPPAINT-YKYVYWRFAKE-DGLQLAWRNHLGGTGINK--  68
PUF CD4    MEKFRSALLLLLLSALVSASRAQ---EVTYAEVGQPVSLTCPKTQS-WGHLEWLFG----GRELAWRNHMGGQKVNT--  69
TRO CD4    MKCVSG-FLSIIIALFISSTGA--EDVVVYGQVGETVTLPRSKWGSERVLVQWFFG-I-DTQPLISRNSHGRETI----  70
HUM CD4    MNRGVPFRHLLLVLQLALLPAATQGKKVVLGKKGDTVELTCTASQKKSIQFHWKNS-----NQIKILGNQGSFLTK--G  72
             *:              :      :       .       . * :       . *            *

                                                                          D2
HAL CD4-2  PAAIKERTKLKRDKN------LEISDVRREDAGAFTCEAD-------GRSHAHTLLVVS--VSASPSEELQPGSKVTLQ  128
PUF CD4-2  KGDICLRATMRQMT-------LEIRDLKETDAGTFTCKKN-------LQSEQHLIVIT--VSASPSASLQLGGTATLQ  132
TRO CD4-2  NAPMVERAKVRRDR-------LEISALNDGDAGLYICKVD-------GKDMDHRLDIVT--VKVHPSNELNEGNNAILE  133
HAL CD4    DEPWTNKLTLPDES-------LVITNLQPENFGTYFCEITS----GSQTLPIHSIELIKLNVSVSPAEPLLPGESLSLS  136
PUF CD4    DGNWKELALSPDGS-------LTVNYIRPEQFGIYVCKVSESGSSGTGTEKAVSYKVLRV--SAEPS-LVLSGQTVTLV  140
TRO CD4    DPEWKDRLSLSKTDFS-----LIINNIRLEDFKSFKCELKDF---MPQTSTSVTFRLFRLTVSVQPVSPLLAGKNLNLK  139
HUM CD4    PSKLNDRADSRRSLWDQGNFPLIIKNLKIEDSDTYICEVE-------DQKEEVQLLVFGL--TANSDTHLLQGQSLTLT  142
                          * : :.  :   : *: *:              :.    ..  : * . *

                                                      D3
HAL CD4-2  CEVAG--LNPDSTVQWMRPDGSAH--------EGSDTVHLNPVAASDSGTWECAFSHGGETYREKLDVRVK--------  188
PUF CD4-2  CRVKG--PNLEPEVRWKKPDGSLY--------SGSKDADLTEVARSDEGTWNCTFDYQGRQYGETLDIHVI--------  192
TRO CD4-2  CQVTG--VDPLPSVEWVSPGGKVEGAPGR---PGSRNVSFSSVALSDTGEWTCQITQDEKTHKETQTINVR--------  198
HAL CD4    CTADY--RKK-PEIYWLNPRGERIKN------NQGTVTVR-VTSQDDGMWICVVAEEKQ---VKMPVKVVGLSPAPLV  201
PUF CD4    CEAPHPSSRQTPEIHWLDPQGEEVK------RGNGEVKVS---GRHSGQWTCMVTLGQRTHRAHVSVTV--VDLDSPP  207
TRO CD4    CDIEE--IFKGTQRRWLSPQKQDLNEDKRAQIRNDGSLTVMSVTDQDHGEWTCVVTYQGREAYANTHVTVIDLSPAHPQ  216
HUM CD4    LESPP---GSSPSVQCRSPRGKNIQ-------GGKTLSVSQLELQDSGTWTCTVLQNQKKVEFKIDIVVLA-FQKASS  209
                .         * .   .         .      .  . * * * .  .       : *

HAL CD4-2  ------------------------------------------------------------------------------  188
PUF CD4-2  ------------------------------------------------------------------------------  192
TRO CD4-2  ------------------------------------------------------------------------------  198
HAL CD4    PHYTSTSSP--ITVPC--SIPPLITWEQIKAKGIHEVHWQFFPETSSGLISQDAQRLFSLSLAGPLSWKRDQPRGLSPA  276
PUF CD4    LQYTSKSSP--LSIPC--SIPAHVSWEQIKALGLQEGHWHFFPRSASDLLSSEAQRLFSLSLD-PVAWQSDQTRGLSSS  281
TRO CD4    PIYTSVSSLSLLHLPCFFSIPPPLSWSDSQEKSIQGGRWTFTPSPAAGSLTGVVQTLANLSLGPPLAWVVNQKRELDVS  295
HUM CD4    IVYKKEGEQ--------VEFSFPLAFTVEKLTGSGELWWQAERASSS-------KSWITFDLKNKEVSVKRVTQDPKLQ  273

                                              D4
HAL CD4-2  ------------------------------------------------------------------------------  188
PUF CD4-2  ------------------------------------------------------------------------------  192
TRO CD4-2  ------------------------------------------------------------------------------  198
HAL CD4    RDPKT-GNLDLTRKLARVEDRGDYMCTMKFKNGRTLNRTVHVEVLQIISSPGPHLLSGQQLNLTCSVGQPLPSDLHLQW  354
PUF CD4    PDLKN-RNLSLGRRKGDDGDRGDYVCSLKFDNGLTLSRTVRVDVLEIVSAPGTDLISGQQLNLSCSLGVPLTSDLRPRW  359
TRO CD4    ALQRTNLNLSLSKKGVTEGDRGEYTCAVEFQRGDTLKRSMRVEVLQVFSSPAPVAFVGQEVNLTCTLGHPLTSDLKVKW  374
HUM CD4    MGKKLPLHLTLPQALPQYAGSGNLTLALEAKTG-KLHQEVNLVVMRAT--------QLQKNLTCEVWGPTSPKLMLSL  342

                                                      CONNECTING PEPTIDE
HAL CD4-2  ----------------------------------------------------------DPVSPTSPAVRTSEF-N  205
PUF CD4-2  ----------------------------------------------------------APPPAPVTSRGP-G  206
TRO CD4-2  ----------------------------------------------------------SLLPNDGQYDGQGHSG  215
HAL CD4    FRPKQSAQPDL----KSARLTIPEVSTDDGGQWECGLL-QGETRLTSAAITLTIEPK-LSVW----------------  409
PUF CD4    IPPEGSSLQRP----LSGRLAIPAVSAGDGGKWRCELR-RNDTLLLTSAVITLKIESR-LTVW---------------  414
TRO CD4    IPPRQSSLLALGSAPDSAHLTIPEARDINGGRWRCELW-RNKTKLTSVEITLKIERVPMDVW---------------  434
HUM CD4    KLENKEAKVSK-------REKAVWVLNPEAGMWQCLLSDSGQVLLESNIKVLPTWSTPVQP---------------  395

                                      TRANSMEMBRANE          CYTOPLASMIC DOMAIN
HAL CD4-2  NG------TPCPSC--GAD--PLG-----WWGPLGWWLWAAAGAGGLVVLLLMVLVIVLCRRIRRRRKRKWQNKMKNGQR  269
PUF CD4-2  PN------PPCKDC--GANGRPVGPSPSPLLG-LSWWMWVVGVGVCLIVLVLVVFIICLYKRIQRKKRKLR-RMENSRQ  276
TRO CD4-2  PNSDVHTVTTCHHCTKGSQ-QPVEWVP--MLG-LSLWVWVAVGAGCLVGVLLLVTIVLLHRRNKIMKRRDR-KMKNIRV  289
HAL CD4    -------------------------------MLVISCSAAAIVSLLLLILGLIIHR-----RRRR-KMRHLR-  444
PUF CD4    -------------------------------MLVIICSAAVIV-FLLLLLGFLCRH-----RRAQ-KMRHMR-  448
TRO CD4    -------------------------------LLVTICGAAVIF-VLLLILTVILNR-----RHRQ-RVTMPRR  469
HUM CD4    -------------------------------MALIVLGGVAGLLLFIGLGIFFCVRCRHRRRQAERMSQIKR  437
                                     .   . .  .*.:: :  :       *:   ::  :

HAL CD4-2  SARSKTYCQCNCRPAAEKPQPGRQKARPSTLAQKPLLRE  308
PUF CD4-2  LLMNKQYCQCDRPAAAPKPQ---QDACERALGPP---RH  309
TRO CD4-2  PLKSNDYCQCNRTLEGPPRR--TQREKPSAGPRQQ--R-  323
HAL CD4    ----PQLCRCK------KPK-------PKGFYRT-----  462
PUF CD4    ----HRLCQCK------NPK-------PKGFYRT-----  466
TRO CD4    --GKRRICRCK------DPQ-------PKGFYRN-----  489
HUM CD4    LLSEKKTCQCP-------HR------FQKTCSPI-----  458
                 *:*            :

Figure 2
```

Figure 2

Alignment of the predicted halibut CD4-2 amino acid sequence with other known CD4 sequences. Gaps, shown as dashes, maximized the alignment of sequences. Identical residues are indicated by an asterisk below the alignment, while double dots and single dots indicate chemical similarity between the amino acids. The positions of domains are shown above the alignment in upper case. Conserved cysteines believed to form inter- and intra-chain disulfide bridges, as well as conserved tryptophan residues that may form a structural triad (C-CW) thought to stabilize the intra-chain disulfide bridges within the Ig-domains are on a grey background. In human CD4, the $F^{43}xxK$ interacting with MHC, K^{318} and Q^{344} important in non-covalent dimer formation, the palmiotylation sites (CxxC), and the dileucin motif essential for endocytosis of CD4 with the phosphorylation sites S^{408} and S^{415}, are in bold. So are the $F^{43}xxK$ motif and the predicted palmitoylation site within halibut CD4-2. The conserved CxC motif within the cytoplasmic tail believed to interact with Lck through a zinc clasp structure, are indicated with white letters on a black background. N-linked and O-linked glycosylation sites are shown in bold letters and underlined. Abbreviations: HAL CD4-2 – Atlantic halibut CD4-2 (*Hippoglossus hippoglossus*; GU985449), PUF CD4-2 - Green puffer CD4-2 protein (*Tetraodon nigroviridis*; EF601918), TRO CD4-2 - Rainbow trout CD4-like 2a*01 protein (*Oncorhynchus mykiss*; AY772711), HAL CD4 – Atlantic halibut CD4 (*Hippoglossus hippoglossus*; FJ185042), PUF CD4 – Green spotted puffer CD4-4a protein (*Tetraodon nigroviridis*; EF601919), TRO CD4 - Rainbow trout CD4 (*Oncorhynchus mykiss*; AY973028), and HUM CD4 – Human T-cell surface glycoprotein CD4 (*Homo sapiens*; M12807).

Figure 3

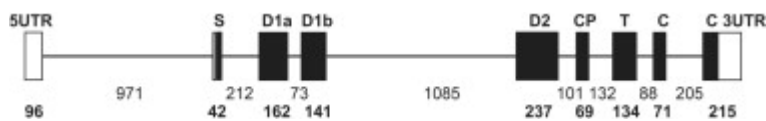Schematic representation of the halibut CD4-2 gene. Open boxes represents untranslated regions (UTRs), shaded boxes represent coding regions, and introns are indicated with a line. Numbers refer to the size of the corresponding exon or intron, where bold numbers refer to the size of exons. Abbreviations: S – signal peptide, D – Ig-like domains, CP – connecting peptide, T – transmembrane region, C – cytoplasmic tail.

```
    SH4 (UNIQUE DOMAIN)                                                    SH3
HAL MGCNCSSDYSDTDWIENLDEICEQCNCPIPPQS-------CNPYTDQLIPCHS-QHSPPMSPLPDNLVEAIYSYEPNHDGDLGF -  76
FUG MGCNCSSDYSDGEWIENLDEMCEHCNCPIAPQS-------CNPYTDQLIPIHSPQLSPPSSPLPDNLVVAIYSYEPKHDGDLGF -  77
TRO MGCNCSSDYDD-DWVENLDEVCDNCNCPIPTQS-------AKPYTDQLIPYPS-HLTPPSSPLPDSLVIAIYSYEPNHNDDLGF -  75
ZEB MG-NCGSFDPDDEFSYHPDEWCDQCNCPKPPVS-------QN--FDPLIPYPSQYTPPPSSPLPENLVVAIYKYDPAHSDDLGF -  74
MOU MGCVCSSNPEDDWMEN--IDVCENCHYPIVPLDSKISLPIRNGSEVRDPLVTYEGSLPPASPLQDNLVIALHSYEPSHDGDLGF -  82
HUM MGCGCSSHPEDDWMEN--IDVCENCHYPIVPLDGKGTLLIRNGSEVRDPLVTYEGSNPPASPLQDNLVIALHSYEPSHDGDLGF -  82
    **  *.*   *          : *::*: *  . .      :              ** *** :.** *::.*:* *..****

                                                           SH2
HAL EKGDKLKIINKDDPEWYLAESLTTGQQGFIPYNFIAMT-TVETEPWFFKSISRNEAMRLLLAPGNTQGSFLIRESETIQGSYSL - 159
FUG EKGDKLKIISKEDPEWYLAESLTTGQRGYVPYNFVAMS-TMEIEPWFFKNISRNEANRRLLAPGNTQGSFLIRESETTPGSYSL - 160
TRO EKGDKLKILNKDDPEWFMAESLITGQKGFIPYNFVAPLNSMEMETWFFKNLSRNDAMRLLLAPGNTQGSFMVRESETAQGSFSL - 159
ZEB EKGEKMKILDCDDPEWYMAESLFTGQRGYIPKNFVAKLNSMETEPWFYKNLSRNDAMRQLLAPGNTQGSFLIRESETQPGSFSI - 158
MOU EKGEQLRILEQSG-EWWKAQSLTTGQEGFIPFNFVAKANSLEPEPWFFKNLSRKDAERQLLAPGNTHGSFLIRESESTAGSFSL - 165
HUM EKGEQLRILEQSG-EWWKAQSLTTGQEGFIPFNFVAKANSLEPEPWFFKNLSRKDAERQLLAPGNTHGSFLIRESESTAGSFSL - 165
    ***::::*. .. **: *:** ***.*::* **:*   ::* *.**:*.:**::* * *******:***:****:   **:*:

                                                         LINKER       SH1
HAL SIRDLDHNTGEGVKHYRIRNMDNGGFYITAKISFSSLKELVQHHVRETDGLCTKLGKPCQSRAPQKPWWQDEWEIPRESLKLLR - 243
FUG SIRDLDSNVGDEVKHYRIRNMDNGGFYITAKISFNALKELVQHYSRDSDGLCTKLVKPCQSKAPQKPWWQDEWEIPRESLKLER - 244
TRO SVRDLDPNTGDTVKHYRIRNLDTGGFYITAKISFNSLKELVQHHSREADGLCTRLMKPCQSRMPQKPWWQDEWEIPRESLKMER - 243
ZEB SVRDLDPMQGDIIKHYRIRNMDAGGFYITNKISFNSLSELVKHYSREADGLCTRLVKPCQTRAPQKPWWQDEWEVPRESLKLER - 242
MOU SVRDFDQNQGEVVKHYKIRNLDNGGFYISPRITFPGLHDLVRHYTNASDGLCTKLSRPCQTQKPQKPWWEDEWEVPRETLKLVE - 249
HUM SVRDFDQNQGEVVKHYKIRNLDNGGFYISPRITFPGLHELVRHYTNASDGLCTRLSRPCQTQKPQKPWWEDEWEVPRETLKLVE - 249
    *:**:*    *: :***:***:* *****: :*:* .* :**:*:  . :*****:* :***:: ******:****:***:**: .

    Catalytic core
HAL RLGAGQFGEVWMGEYNNDREVAIKKLKMGTMSVEAFLAEANMMKNLQHPRLVRLFAVVTQEPILIVTEYMENGSLVDFLKTTEG - 327
FUG KLGAGQFGEVWMGIHNNERRVAIKCLKIGTMSVEAFLAEANMMKSLQHMHLVRLFAVVTQEPIFIVTEYMENGSLVDYLKTTEG - 328
TRO RLGAGQFGEVWMGLYNNHRRVAIKNLKVGTMSMAAFLAEANLMKELQHPRLVRLFAVVTQEPIYIITEFMENGALVDFLKSSEG - 327
ZEB RLGAGQFGEVWMGLYNNNRQVAIKSLKPGTMSISAFLAEANLMKSLQHPRLVRLFAVVTQEPIYIITEYMENGSLVDFLKTPEG - 326
MOU RLGAGQFGEVWMGYYNGHTKVAVKSLKQGSMSPDAFLAEANLMKQLQHQRLVRLYAVVTQEPIYIITEYMENGSLVDFLKTPSG - 333
HUM RLGAGQFGEVWMGYYNGHTKVAVKSLKQGSMSPDAFLAEANLMKQLQHQRLVRLYAVVTQEPIYIITEYMENGSLVDFLKTPSG - 333
    :** ********* :*.. .*:* ** *:** *******:**.*** :****:******* *:**:****:***:**:..*

                      Catalytic loop        Activation loop
HAL SSLPMNTLIDMASQVADGMAFIEQENYIHRDLRAANILVSHEHICKIADFGLARLIEDNEYTAREGAKFPIKWTAPEAINYGTF - 411
FUG SSLSINTLIDMASQARHGMAFIEARNYIHRDLRAANILVSHELICKIADFGLARLIENNEYTAREGAKFPIKWTAPEAINYGTF - 412
TRO SNIPINTLIDMASQVAEGMAYIEAKNYIHRDLRAANILVSDELICKIADFGLARLIKNNEYTAREGAKFPIKWTAPEAINYGTF - 411
ZEB SDIPINTLIDMAAQVAEGMAYVEQKNYIHRDLRAANILVSHELTCKIADFGLARLIKNNEYTAREGAKFPIKWTAPEAINYGTF - 410
MOU IKLNVNKLLDMAAQIAEGMAFIEEQNYIHRDLRAANILVSDTLSCKIADFGLARLIEDNEYTAREGAKFPIKWTAPEAINYGTF - 417
HUM IKLTINKLLDMAAQIAEGMAFIEERNYIHRDLRAANILVSDTLSCKIADFGLARLIEDNEYTAREGAKFPIKWTAPEAINYGTF - 417
     .: :*.*:***:*  .***::*  .**************.    ***********::*************************

HAL SIKSDVWSFGILLTEIVTYGRIPYPGMSNPEVIQNLERAYRMPKPDNCPEGLYNVMGMCWRETPDDRPTFEYLRSVLEDFLTAT - 495
FUG SIKSDVWSFGILLTEIVTYGRIPYPGMSNPEVIHQLEQNYRMPKPENCPDGLYNFMLLCWREKPEDRPTFDYLRSVLEDFFTAT - 496
TRO SIKSDVWSFGILLTEIVTYGRIPYPGMSNPEVIQNLEKGYRMPRPENCPEDLYNIMDLCWKESPENRPTFEYLRSVLEDFFTAT - 495
ZEB SIKSDVWSFGVLLTEIVTYGRIPYPGMTNPEVIANLERGYRMPCPDNCPEALYNVMKHCWTENPDNRPTFEFLRSVLEDFFTAT - 494
MOU TIKSDVWSFGILLTEIVTHGRIPYPGMTNPEVIQNLERGYRMVRPDNCPEELYHLMMLCWKERPEDRPTFDYLRSVLDDFFTAT - 501
HUM TIKSDVWSFGILLTEIVTHGRIPYPGMTNPEVIQNLERGYRMVRPDNCPEELYQLMRLCWKERPEDRPTFDYLRSVLEDFFTAT - 501
    :*********:*******:*******:**** :**: ***  *:***: **:.*  ** * * *::****::*****:**.***

    COOH-domain
HAL ERQYQEDPCMGRRT - 509
FUG ERQYQE-------- - 502
TRO ERQYQEQP------ - 503
ZEB EGQYQEQPC----- - 503
MOU EGQYQPQP------ - 509
HUM EGQYQPQP------ - 509
    * ***
```

Figure 4

Alignment of the predicted halibut Lck amino acid sequence with other known Lck sequences. Gaps, shown as dashes, maximized the alignment of sequences. Identical residues are indicated by an asterisk below the alignment, while double dots and single dots indicate chemical similarity between the amino acids. The positions of domains are shown above the alignment in upper case. The N-terminal domain associated with myristoylation and palmitoylation (GCxCS), the conserved CD4/CD8 interaction motif (CxxC), the nucleotide binding region $(GxGxxG(x)_{15}K)$, the protein tyrosine kinase motif (IHRDLRAANI), and the activation loop are in bold letters on a grey background. Conserved tyrosines, Y^{394} and Y^{505}, that are phosphorylated and dephosphorylated respectively upon Lck activation, and S^{59} that is believed to be a part of a MAPK target sequence (PxPS) are indicated with white letters on a black background. Abbreviations: HAL – Atlantic halibut (*Hippoglossus hippoglossus*; FJ769822), FUG – Fugu (*Takifugu rubripes*; AF411956), TRO – Rainbow trout (*Oncorhynchus mykiss*; AY973033), ZEB – Zebrafish (*Danio rerio*; AY390224), MOU – Mouse (*Mus musculus*; M12056), and HUM – Human (*Homo sapiens*; X05027).

```
                    SH2 DOMAIN 1
HAL MSTDPAAELPFYYGSISRTDAEQHLKLAGMVDGLFLLRQCLRSLGGYVLSIACNVEFNHYTIEKLLNGTYCIVGGKPHCGPAEL - 84
PUF MSTDPAAELPFFYGSISRAEAEQHLKLAGMDDGLFLLRQCIRSLGGYVLSVVCSLEFHHYSIEKQLNGTYCIAGGKPHCGPAEL - 84
ZEB -MTEPAADLPFFYGSISRSEAEEHLKLAGMGGGLFLLRQCLRSLGGYVLSMIWNLDFYHYSIEKQLNGTYCIAGGKPHCGPAEL - 83
FRO -MPDVAGHLPFYYGSISRADAEEYLKLGGMMDGLFLLRQCLRTLGGYVLSMVYNVHFHHYPVERQLNGTYAIAGGKAHCGPAEL - 83
MOU -MPDPAAHLPFFYGSISRAEAEEHLKLAGMADGLFLLRQCLRSLGGYVLSLVHDVRFHHFPIERQLNGTYAIAGGKAHCGPAEL - 83
HUM -MPDPAAHLPFFYGSISRAEAEEHLKLAGMADGLFLLRQCLRSLGGYVLSLVHDVRFHHFPIERQLNGTYAIAGGKAHCGPAEL - 83
        .: *..***:******::**:***.** .********:*:*******:  .: * *:.:*: ****:.*.***.*******

HAL CEFYSKDADGLVSNLRKPCLRAPDTPIQPGVFDSLRENMLREYVKQTWALEGEAMEQAIISQAPQLEKLIATTAHERMPWYHGK - 168
PUF CEFYSKDPDGLVCLLRKPCLRPADMPVRPGTFENLKEKMVRDYVKQTWNLEGEAMEQAIVSQASQLEKLIAITAHEQMPWYHGN - 168
ZEB CEYYSKDPDGLVCTLRKPCLRSPETEIRKGVFDNLRDNMLREYVRHTWKLEGDAMEQAIISQAPQLEKLIATTAHEKMPWFHGK - 167
FRO CEYYSKDADGLSCTLRRPCNRPVGVECQAGVFDNMRDNMMRRYVRQTWKLEGDALEQAIISQAPQVEKLIATTAHERMAWYHGS - 167
MOU CQFYSQDPDGLPCNLRKPCNRPPGLEPQPGVFDCLRDAMVCDYVRQTWKLEGDALEQAIISQAPQVEKLIATTAHERMPWYHSS - 167
HUM CEFYSRDPDGLPCNLRKPCNRPSGLEPQPGVFDCLRDAMVRDYVRQTWKLEGEALEQAIISQAPQVEKLIATTAHERMPWYHSS - 167
        *:.**:*.***  .**:** *.     : *.*: :::  *: :**::** ***:*:****:***.*:***** ****:*.*:*..

    SH2 DOMAIN 2
HAL ITRQEGERRLYSGAQPDGKFLVRDRDKSGTFALSMIYGKTVYHYQILQEKSGKYCMPEGTKFDTIWQLVEYLKMKPDGLVTVLV - 252
PUF IPRQEGERRLYCGAQPDGKFLVRERDEPGSFALSVMYGKTVYHYQILQAKSGKYSMPGGTNFDTIWQLVEYLKMKPDGLVTVLG - 252
ZEB IPRQEGERRLYSGSQPDGKFLVRERDEMGTFALSVTYGKTVYHYQILRDKSGKIAMPEGTKFDTVWQLVEYLKMKPDGLVTVLR - 251
FRO ISRDEAERRLYSGAQPDGKFLMRERKENGTYALSVMYGKTVYHYKIDQDKSGKYSIPEGTKFDTLWQLVEYLKLKSDGILAVLK - 251
MOU LTREEAERRLYSGGQTDGKFLLRPRKEQGTYALSLVYGKTVYHYLISQDKAGKYCIPEGTKFDTLWQLVEYLKLKADGLIYRLK - 251
HUM LTREEAERRLYSGAQTDGKFLLRPRKEQGTYALSLIYGKTVYHYLISQDKAGKYCIPEGTKFDTLWQLVEYLKLKADGLIYCLK - 251
        :.*:*.**:**.* *.*****:* *.: *::***: ******** * : *:** .:* **:***:********:*.**:: *

            INTERDOMAIN B
HAL ESCMNGKAAA-KMPNLP-----------ASRR---VNVNGYTPPPRVVTEASEPAAERDVLPMDC-----TGFNPYHNPNEVK- - 315
PUF EVCINGRAAESTIPLFS-----------TLRP---QRQNGYTPPPGGHLAEARAAAPKTSVSMAADQDPGTDSNPYHNPNEIR- - 321
ZEB EPCVNQNNAAPAPANAP-----------RR---SRGNGYTPPPMVPKPMG-AEGSRPSLPMDHD----GFTSPYDDPNELKK - 314
FRO ESCANASTFSIAPAAAP-----PS---LPKVRPVASNSDGYTPEPF-------LGKSRI-LPMDTS----VYESPYSDPEELKE - 315
MOU EVCPNSSAS--AAVAAPTLPAHPSTFTQPQRRVDTLNSDGYTPEPARLASS--TDKPRP-MPMDTS----VYESPYSDPEELKD - 326
HUM EACPNSSASNASGAAAPTLPAHPSTLTHPQRRIDTLNSDGYTPEPARITS---PDKPRP-MPMDTS----VYESPYSDPEELKD - 327
        * * *          .                  . :**** *         : :.*       .** :*:*::

            SH1    Catalytic core
HAL -RFNIKRTQLLIDEVELGCGNFGSVKKGVLKTDAGQIDVAIKVLKNENEKLVKEEMMREAEIMHQLNNPFIVRMLGLCNAESLM - 398
PUF -KFSIQRSQLLIDEVELGSGNFGCVKKGVLRTDSGQVDVAIKVLKSDNESLVKDEMMREAEIMYQLSNRYIVRMLGLCNAENLM - 404
ZEB KTLFIKRDKLMIDEVELGSGNFGCVKKGVFKMESKQIDVAIKVLKNENEKSVRDEMMREAEIMHQLSDPFIVRMIGLCEAEALM - 398
FRO RKLFVKRELLLIDEVELGSGNFGCVKKGVYKLKKRQIDVAIKVLKVQEEKNVRDEMMKEAEFMHQLDNPYIVRMIGVCEAENLM - 399
MOU KKLFLKRENLLVADIELGCGNFGSVRQGVYRMRKKQIDVAIKVLKQGTEKADKDEMMREAQIMHQLDNPYIVRLIGVCQAEALM - 410
HUM KKLFLKRDNLLIADIELGCGNFGSVRQGVYRMRKKQIDVAIKVLKQGTEKADTEEMMREAQIMHQLDNPYIVRLIGVCQAEALM - 411
        : ::*  *:: ::***.****.*::** :    *:********* *.   :***:**::*:**.: :***:*:*:** **

                                Catalytic loop        Activation loop
HAL LVMEMAAAGPLNKFLSSNKDTVTAENIVNLMHQVSMGMKYLEEKNFVHRDLAARNVLVTQQFAKISDFGLSKALGADDSYYKA - 482
PUF LVMEMAPAGPLNKFLSSKRDTVTVENIVQLMHQVSMGMKYLEEKNFVHRDLAARNVLLVNQRFAKISDFGLSKALGADDSYYKA - 488
ZEB LVMEMAPAGPLNKFLSGKKDQITTENIVMLMHQVSMGMKYLEGRNFVHRDLAARNVLLVNQQYAKISDFGLSKALGADDNYYKA - 482
FRO LVMEMASGGPLNKFLGAKKDTITVSNVVELMHQVSMGMKYLEGKNFVHRDLAARNVMVNQHYAKISDFGLSKALAADDSYYKA - 483
MOU LVMEMAGGGPLHKFLLGKKEEIPVSNVAELLHQVAMGMKYLEEKNFVHRDLAARNVLLVNRHYAKISDFGLSKALGADDSYYTA - 494
HUM LVMEMAGGGPLHKFLVGKREEIPVSNVAELLHQVSMGMKYLEEKNFVHRDLAARNVLLVNRHYAKISDFGLSKALGADDSYYTA - 495
        ****** .***:*** .::: :...*:. *:***:****** :***********:*.::*.***********.***.**.*

HAL KSFGKWPLKWYAPECINFHKFSSKGDVWSFGITMWEAFSYGGRPYKKMKGPDITRFIESGNRMERPTACSERMYAVMNECWTYK - 566
PUF RTAGKWPLKWYAPECINFRKFSSKSDVWSFGVTMWEAFSYGGKPYKKMKGPEVNQFIEAGNRMERPEPCPEKMYALMKECWTYK - 572
ZEB RTGGKWPLKWYAPECIHFHKFSSKSDVWSFGITMWEAFSFGGKPYKKMKGPEVITYIEGGSRLDCPAACPEAMYELMKECWTYK - 566
FRO KSFGKWPLKWYAPECINYRKFSSRSDVWSYGITMWEAFSYGQKPYKKLKGTEVMSFIERNERLACPASCPPEMYQLMLDCWIFK - 567
MOU RSAGKWPLKWYAPECINFRKFSSRSDVWSYGVTMWEAFSYGQKPYKKMKGPEVDFIKQGKRMECPPECPPEMYALMSDCWIYK - 578
HUM RSAGKWPLKWYAPECINFRKFSSRSDVWSYGVTMWEALSYGQKPYKKMKGPEVMAFIEQGKRMECPPECPPELYALMSDCWIYK - 579
        :: *************::::****:.****:.*****:*:* :****:**.:: :*: ..*:  *  *. :*:* :** :*

HAL HDDRPDFKKVEESMRSYHYSISNKAKPEGAADGAAAAAAAAAEPVK---- - 612
PUF HEDRPDFKKVEESMKSYYTSISNKS------------------- - 597
ZEB HEERLNFAKVEEKMRTFYYSIAKKIPDYLKTENGTPPK----------- - 604
FRO MEDRPNFENVEYRMRMYYSIADKPDKESKEGKEGKEAEAGAAAEAPGKE - 617
MOU WEDRPDFLTVEQRMRNYYSLASRAEGPPQCEQVAEAACG---------- - 618
HUM WEDRPDFLTVEQRMRACYYSLASKVEGPPGSTQKAEAACA---------- - 619
        ::* :* .** *:  :  : *::.:
```

Figure 5

Figure 5

Alignment of the predicted halibut ZAP-70 amino acid sequence with other known ZAP-70 sequences. Gaps, shown as dashes, maximized the alignment of sequences. Identical residues are indicated by an asterisk below the alignment, while double dots and single dots indicate chemical similarity between the amino acids. The positions of domains are shown above the alignment in upper case. The nucleotide binding region ($GxGxxG(x)_{18}K$), the protein tyrosine kinase motif (VHRDLAARNV), and the activation loop are indicated above the motif in lower case and are shown in the alignment with a grey background. Conserved tyrosines believed to be important in T-cell regulation have a black background. Abbreviations: HAL – Atlantic halibut (*Hippoglossus hippoglossus*; GU985452), PUF - Green spotted puffer (*Tetraodon nigroviridis*, CAG00734.1), ZEB – Zebrafish (*Danio rerio*, NP_001018425.1), FRO – African clawed frog (*Xenopus laevis*; BC077883), MOU – Mouse (*Mus musculus*; U04379), and HUM – Human (*Homo sapiens*; L05148).
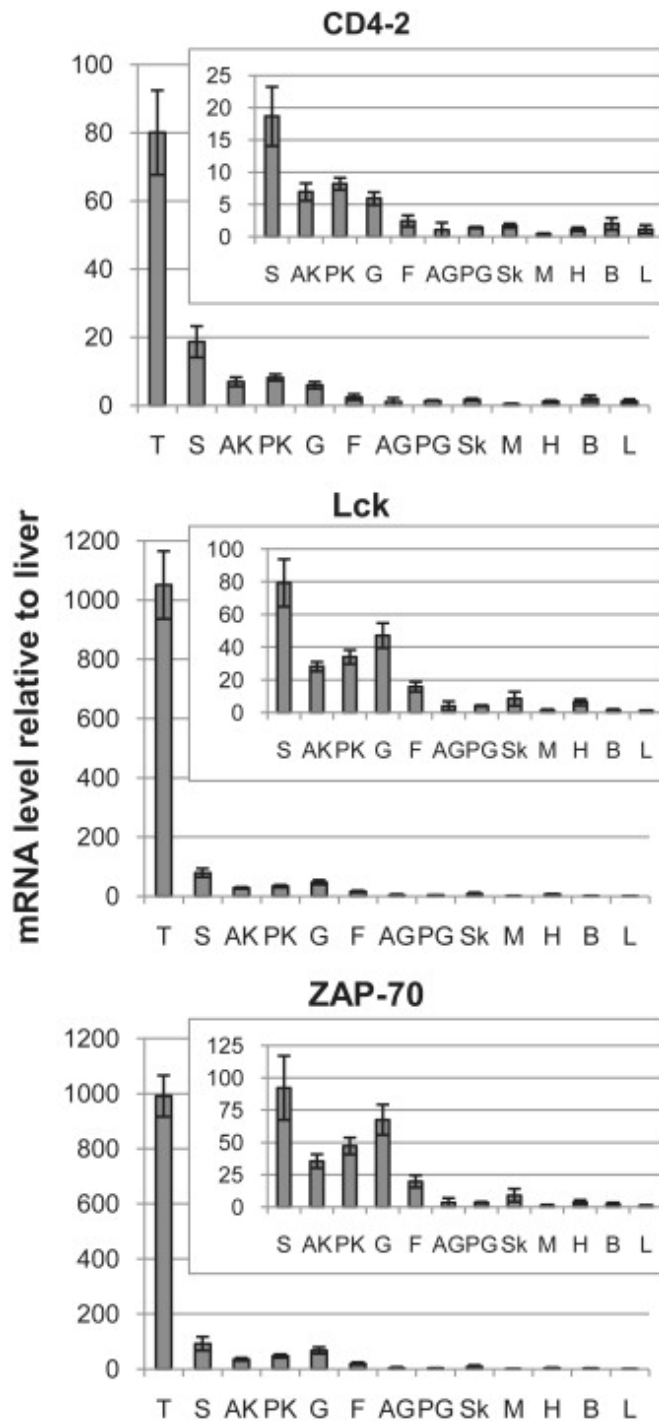
Figure 6
Relative mRNA level of halibut CD4-2, Lck, and ZAP-70 analyzed by real time RT-PCR. Elongation factor 1α (EF1A1) served as internal reference gene while liver was used as calibrator. Data represents mean values of $n = 4$ fish (±confidence interval). Abbreviations: T - thymus, S - spleen, AK - anterior kidney, PK - posterior kidney, G - gill, F – pectoral fin, AG - anterior gut, PG - posterior gut, Sk - skin, M - muscle, H - heart, B - brain, and L - liver.