ORIGINAL PAPER

Hiroaki Suetake · Kyosuke Araki · Yuzuru Suzuki

Cloning, expression, and characterization of fugu *CD4*, the first ectothermic animal *CD4*

Abstract We have cloned and sequenced the first ectothermic animal *CD4* gene from fugu, *Takifugu rubripes*, using a public database of the third draft sequence of the fugu genome. The fugu *CD4* gene encodes a predicted protein of 463 amino acids containing four extracellular immunoglobulin (Ig)-like domains, a transmembrane region, and a cytoplasmic tail. Fugu CD4 shares low identity of about 15–20% with avian and mammalian CD4 proteins. Unlike avian and mammalian CD4, fugu CD4 lacks the Cys pair of the first Ig-like domain, but has a unique possible disulfide bond in the third domain. These differences suggest that fugu CD4 may have a different structure that could affect binding of major histocompatibility complex class II molecules and subsequent T-cell activation. In the putative fugu cytoplasmic region, the protein tyrosine kinase p56[lck] binding motif is conserved. The predicted fugu *CD4* gene is composed of 12 exons, differing from other *CD4* genes, but showing conserved synteny and many conserved sequence motifs in the promoter region. RT-PCR analysis demonstrated that the fugu *CD4* gene is expressed predominantly in lymphoid tissues. We also show that fugu CD4 can be expressed on the surface of cells via transfection. Molecular characterization of CD4 in fish provides insights into the evolution of both the CD4 molecule and the immune system.

Keywords CD4 · Fugu · Cell surface molecules · T lymphocytes

H. Suetake (✉) · K. Araki · Y. Suzuki
Fisheries Laboratory, Graduate School of Agricultural and Life Sciences, The University of Tokyo,
2971-4, Maisaka, Hamana,
Shizuoka, 431-0211, Japan
e-mail: asuetake@mail.ecc.u-tokyo.ac.jp
Tel.: +81-53-5922821
Fax: +81-53-5922822

Introduction

A variety of specific immune functions have been observed in fish. Of these, allograft rejection, mixed leukocyte reaction, and delayed hypersensitivity reaction indicate that T cells are involved in these specific immune activities in a similar manner to that of mammals (Nakanishi et al. 2002). In addition, major histocompatibility complex (MHC) and T-cell receptor (TCR) genes have been identified in fish and amphibians (Hansen and Zapata 1998). It is expected that T-cell subsets exist in fish and amphibians. Nevertheless, T-cell subsets are yet to be demonstrated due to the lack of reagents for fish surface markers such as CD3, CD4 and CD8. Recently, the genes encoding polypeptides homologous to the known T-cell surface markers, CD3 and CD8, were reported in fish (Hansen and Strassburger 2000; Alabyev et al. 2000; Park et al. 2001) and amphibians (Dzialo and Cooper 1997; Gobel et al. 2000; Ropars et al. 2002). However, CD4 has not been identified in an ectothermic animal.

In avian and mammalian species, cytotoxic T cells and helper T cells are major T-cell subsets that carry CD8 and CD4, respectively. These molecules are involved in recognition of their specific antigens in association with MHC class I and class II molecules. Expression of CD8 and CD4 is critical for cell-mediated immune defense and T-cell development in the thymus. CD4 is a membrane-bound glycoprotein found on thymocytes and helper T cells (Parnes 1989), and the primary structure has been identified in several avian and mammalian species (Koskinen et al. 1999; Clark et al. 1988; Classon et al. 1986; Dumont-Drieux et al. 1992; Fomsgaard et al. 1992; Gustafsson et al. 1993; Hague et al. 1992; Maddon et al. 1985; Milde et al. 1993; Norimine et al. 1992; Romano et al. 1999). The CD4 protein is composed of four extracellular immunoglobulin (Ig)-like domains, a transmembrane region, and a cytoplasmic tail that associates with a tyrosine protein kinase, p56[lck], involved in T-cell activation. *CD4* genes have also been identified in human, mouse and chicken, and all are composed of ten exons (Gorman et al. 1987; Maddon et al. 1987; Koskinen et al.

2002). In these species, the promoter region of the *CD4* genes does not have a TATA sequence (Koskinen et al. 2002; Salmon et al. 1993). The activation of the *CD4* promoter also requires several factors, and many sequence motifs have been conserved during evolution (Koskinen et al. 2002). From an evolutionary standpoint, it is important to determine if these elements are conserved in lower vertebrates as well.

Fugu, *Takifugu rubripes*, has the smallest genome among vertebrates: 365 Mb, one-eighth the size of the human genome (Brenner et al. 1993). It has recently been the subject of a whole genome shotgun sequencing program (Aparicio et al. 2002) and the draft sequence data are publicly available. The availability of these data has given us a great advantage in studying the fish immune system. In this paper, we describe the cloning of the first ectothemic animal *CD4* cDNA. We also report the expression and genomic organization of fugu *CD4*.

Materials and methods

Animals

Fugu (*Takifugu rubripes*) specimens were purchased from an aquaculture farm. These fish were kept in a circulating seawater tank at 20°C and were subjected to collection of various tissues and extraction of total RNA.

Screening of *CD4*

The fugu genomic sequence databases ver. 3.0 at the Fugu Genomics Group (http://fugu.hgmp.mrc.ac.uk) and DOE Joint Genome Institute (http://genome.jgi-psf.org/fugu6/fugu6.home.html) were screened in silico with sequences of duck and chicken *CD4* (accession number AF378701 and Y12012) as a probe on a BLAST platform to obtain overlapping sequences (scaffolds) including a putative *CD4*.

cDNA cloning

Specific primers for RACE were designed from the putative coding region of *CD4* (CD4-F1, CD4-F2, CD4-R1, and CD4-R2; Table 1). Total RNA was isolated from the thymus of fugu using RNA extraction solution (ISOGEN, Nippon gene, Toyama, Japan) following the manufacturer's instructions. First-strand cDNA was synthesized from 1 μg of the poly(A)$^+$ RNA with the SMART RACE cDNA amplification kit (Clontech, Palo Alto, Calif.) for 3′RACE and 5′RACE. First-round PCR amplification of 3′RACE was performed with CD4-F1 and universal primer mix-A (UPM-A, Clontech: Table 1). Nested PCR amplification of 3′RACE was performed with CD4-F2 and nested universal primer-A (NUP-A, Clontech; Table 1). Both reactions were carried out for 30 cycles of 94°C for 10 s, 60°C for 5 s, and 72°C for 1 min

with a final extension of 72°C for 7 min. For 5′RACE, first-round PCR amplification was performed with CD4-R1 and UPM-A. Nested PCR amplification of 5′RACE was performed with CD4-R2 and NUP-A. Both reactions were carried out for 30 cycles of 94°C for 10 s, 60°C for 5 s, and 72°C for 30 s with a final extension of 72°C for 7 min. Products were subcloned into pCR 2.1 vector (Invitrogen, San Diego, Calif.) and sequenced using an ABI 310 genetic analyzer (ABI, Foster City, Calif.).

Structural and phylogenetic analysis

Putative leader sequence, transmembrane and cytoplasmic regions were predicted based on SignalP (http://www.cbs.dtu.dk/services/SignalP/) and SOSUI (http://sosui.proteome.bio.tuat.ac.jp/sosuimenu0.html) programs. Immunoglobulin-like domains were predicted using InterPro (http://www.ebi.ac.uk/interpro/). Aligning the full-length amino-acid sequences of CD4 and construction of a phylogenetic tree were conducted using Clustal W (Thompson et al. 1994) and TreeView ver. 1.6.6 (Page 1996).

Genomic analysis

Genomic organization was determined by comparing the cDNA sequence and two scaffold sequences in fugu genomic database version 3.0. Transcription factor binding sites were predicted by TFSEARCH ver. 1.3 (http://www.cbrc.jp/research/db/TFSEARCHJ.html).

Reverse transcriptase PCR

Total RNA was isolated from fugu peripheral blood leukocytes (PBL), thymus, head kidney, kidney, spleen, heart, liver, gonad, intestine, skin, gill, brain, and muscle. First-strand cDNA was synthesized from 1 μg of total RNA of each organ in a 20-μl reaction. PCR amplification was performed with CD4-F2 and CD4-R3 for 30 cycles of 96°C for 10 s, 60°C for 5 s, and 72°C for 30 s (Table 1). PCR products were subjected to agarose-gel electrophoresis and stained with ethidium bromide.

Table 1 Oligonucleotide primers used to amplify fugu *CD4* cDNA

Primers	Sequences
CD4-F1	CCTCCGAGAGGGACACTGGCAGTTCT
CD4-F2	CGAAGTCAAACCTGGTTTCTGCTGACGC
CD4-R1	AACACGTTCACCCGAACGGTCGTGCT
CD4-R2	ACTTCAGGGTGCACACGTAGTCACCGC
CD4-R3	GGATGGTGAGCTGGCCGGATCGTATTG

Transient expression

*Bam*HI linker-linked insert cDNA with Kozak sequence was amplified by PCR, digested, and ligated into the *Bam*HI site of mammalian expression vector pFlag-CMV-1 (Sigma, St Louis, Mo.). COS-7 cells were transfected with 12 μg of this construct in 100-mm plates using TransFast transfection reagent (Promega, Tokyo, Japan). Transfected COS-7 cells were grown for 3 days and then harvested with PBS containing 5 mM EDTA. The cells were washed twice in PBS containing 0.5% BSA, 2 mM EDTA, and 0.01% sodium azide and adjusted to 10^8 cells/ml. To each 100-μl aliquot, anti-Flag antibody was added to give a final dilution of 1:1,000, and cells were incubated on ice for 30 min. The cells were subsequently washed twice and resuspended in PBS/BSA/EDTA/azide and FITC-labeled goat anti-mouse IgG (Jackson Immunoresearch Laboratories, West Grove, Pa.) added to give a final dilution of 1:100. Cells were incubated at 4°C for 30 min, and then washed as above and resuspended in PBS/BSA/EDTA/azide. After suspension, the stained cells were analyzed with a flow cytometer (Partec, Münster, Germany).

Results and discussion

Isolation of fugu *CD4* cDNA

We cloned a *CD4* homologue gene from a teleost species, fugu. We chose amino acid sequences of duck and chicken CD4s to query the fugu genomic sequence database by BLAST search. A sequence similar to the duck and

Table 2 Amino acid identity (%) between the CD4 proteins of fugu and other vertebrates

	Fugu
Chicken	15.4
Mouse	16.7
Human	19.6
Rabbit	17.1
W. whale	18.9
Cat	16.6

```
              Leader                                                           D1
Fugu     MTFVSRHIPDMEPLPSGLLLLTALLSASRAEELIYAQVGQTVTLKPPENYKTPTYYLSWHFGELELAWTNHMS------------GNKVIKHENWDTALSDNSLVVKEIRQNQFGIYKCN
Chicken  --MERCGAVVSCVFAVI.V.QLG.TPIMAQQ.QQIGIA.KE.I.SCKAINNQKDGTCT.KYKYK.VS-STII.FSKAQVFKGKAPMTHRSELNSN-----SKK.K.SDLSLDDA...T.A
Mouse    ---MC.A.SLRRL.-LL..Q.SQ..AVTQGKT.VLGKE.ESAE.PCESSQ-KKITVFT.K.SDQRKILGQ.GKG--VLIRGGSPSQFDRFDSKKGAWEKGSFP.IINKLKMEDSQT.I.E
Human    ----MNRGVPFRH.LLV.Q.ALLPAATQ-GKKVVLGKK.D..E.TCTASQ-KKSIQFH.KNSNQIKILG.QG.----FLTKGPSKL.DRADSRRSLWDQGNFP.II.NLKIEDSDT.I.E
Rabbit   ----MNRRIYPQC.LLV.P.ALLPAATW-GKTVVRGKA.AI.E.PCQSSQ-KRNSVFN.KHANQVKILG.QG.SSSSFWLKGNSPLSNRVESKKNMWDQGSFP..I.DL.MDDS.T.I.E
W.whale  ----MDPRTSLRH.FLV.Q.VMLPAGTQ-GKKVVLGKA.ELAE.PCKASQ-NKSLFF..KNSYQTKILGR.GY----FWHKGASNLHSRVESKINLWDQGSFP..I.DLEVPDS.T.I.E
Cat      ----MNQGAVFRH.LLV.Q.VMLKAAVPQGK.VVLGKA.G.AE.PCQASQ-KKYMTFT.RLSSQVKILESQH.S---LCLTGSSKLKTRFESKKILWDQGSFP..I.SLQVADS...T.E
                       *                        *     *              *                                                   *           *   * *
                                                                               D2
Fugu     VNEKIWTYKVLRLKVSAEPPS----------------LVLSGRTVTLVCDAEPPNSLQKPGIHWLNPQ-----------GEKITQATHSVQVSSRHSGRWTCVVTLDRKEATAQISVTV
Chicken  CYSPVVSISLHVF.LTISSNG----------------HF.TNEDLE.TLMQNSSH.QPHLS.KLF.INNDIVTTEILQEEAPQKYILKLK.LKAID..T.M.H.YSNSPSINQN..FD.
Mouse    LENRKEEVELWVF..TFS.GT----------------SL.Q.QSL..TL.SNSKV.NPLTECKHKKGK----------VVSGSKVLSMSNLRVQD.DF.N.T....--QKKNWFGM.L
Human    .EDQKEEVQL.VFGLT.NSDT----------------HL.Q.QSL..TLES-..G.SPSVQCRSPRGK-----------NIQGGKTLSVS.LELQD..T...T.LQN--QKKVEFKIDI
Rabbit   .GD.KMEVEL.VFRLT.N.NT----------------RL.H.QSL..TLEG-.SVGSPSVQWKSPENK-----------IIETGPTCSMPKLRLQD..T.S.HLSFQD-QNKLELDIKI
W.whale  .ED.KIEVELQVFRLT.SSDT----------------RL.L.QSL..TLEG-.SG.NPSVQWKGPGNK-----------RKNEAKSLSLP..GLQD..T...T.SQA--QQ.LVFNKHI
Cat      .EN.KREVEL.VFGLT.KVDPSGSGGSSSSSTSTSTSIYL.Q.QSL..TLES-..SS.NPSVQWKGPGNK-----------SKSGVHSLSLS.LELQE..TC.LT.SQS--QK.LVFNTNI
                      *          *                              *         *                                           *   *         *   *
                              D3
Fugu     VDLYSPPMAYTSTSSPLAVPCSVPKVSWEQIKSLGLREGHWQFFPRSKSNLVSADAQRLFTLSLEEPVSWKANQTRGLTPVSDFKTPNLSLGRTLGRADDRGDYVCTLKFESGPPLSTTV
Chicken  KV.GFEKERLEIIYTTVGN-TAILSWRLNFR.IKWKEGFTGKLNWEPQG.TAIHELLNFSVTTHQ.LHKT.KSNHIWFEISEGKTDGTMDVKIPKVQLNHS.Q.K.Q.EING-RRTESVR
Mouse    SV.GFQST.I.AYK.EGE----SAEF.FPLNFAEENGW.--ELMWKAEKDSFF-QPWIS.SIKNK.VSVQ.STKDLK.QLKETLP-LT.KIPQVSLQFAGS.NLTL..DK---GT.HQE.
Human    .V.AFQKASSIVYKKEGE----QVEF.FPLAFTVEKLT.SGELWWQAERASS.-KSWIT.D.KNK.VSVKRVT.DPK.QMGKKLP-LH.T.PQA.PQYAGS.NLTLA.EAKT-GK.HQE.
Rabbit   IV.GF.KASA.VYKKEGE----QVEF.FPLNFEDESLS.--ELMWQVDGASSA-QSWVS.S.EDRKVSVQ.ILPDLKIQMSKGLP-LS.T.PQA.H.YAGS.NLSL..DK---GK.HQQ.
W.whale  LV.AFQOEVSS.VYAKEGE----QMNF.FPLTFGDENLS.--ELSWLQAKGNS.PESWIT.K.NNGKVTVG..RKDLK.RMSKALP-LH.T.PQA.PQYAGS.NLTLN.TK---GK.YQE.
Cat      LV.AFRKVSN.VYAKEGE----QVEF.FPLNFEDENLM.--NLRWKAEGAPS.-LLWIS...KNKQLSVKEVDPYSK.QMMDSLP-LRFT.PNV.S.YAGS.NLTLV.DK---GQ.QQE.
                *                                                                                              *      *
                                      D4                                                                TM
Fugu     RVNVLEIAASPGTVLISGQQLNLTCGLGVPLTSDLHLKWIS----PERATIRSGQLTIPAVGAGNSGKWRCELWRNDTRLTSAVITLKIEPKLSVWM--LVIICSVAVIVLLLLLGFIL
Chicken  ALV.MQVT.I.AGP.SR.GKMT.L.QVSG..P.NA..L.ERVNGTQMEMKKSKQHEAKVE.NVSAP.L.N.H.VEDNNKKI.LNY.VEEAHVWNSYA-VIG..IGAS.L.IG.ACMCI.T
Mouse    NLV.MKV.QLN---------NT...EVMG.TSPKMR.TLK----QENQEARV.EEQKVVQ.V.PET.L.Q.L.SEG.KVKMDSR.QVLS--RGVNQTVF.ACVLGGSFGF.GF.G.CILC
Human    NLV.MRATQLQ---------K....EVMG.TSPK.M.SLK----LENKEAKV.KREKAVW.LNPEA.M.Q.L.SDSGQV.LESN.KVLPTWSTP.QP-MAL.VLGGVAGL..FIG..IFF
Rabbit   SLVM.KVTQVK---------NK...EVLG.IDPKMK.SLK----LEDQEAKV.-TQKMVQ.LDPKA.T.Q.L.SSG.QV.LESKADVLATGLSHQQPTL.AGALGGTAGLV.FAG.CIYC
W.whale  NLV.MRVTK..---------NS...EVLG.TSPR.I.SLK----KENQSMRV.D.QKLVT.LGPEA.M.Q.L.SDKGKV.LESKVKILPPVLAHA.PKL.AVVLGGITSL...AGFCIFS
Cat      KLV.MRVTQ.G---------N....EVLG.TSPE.T.SLK----LKGQ.AKV.K.QKMVR.EDAEA.T.Q.L.SHK.KV.LASKAEVLPPVLTRT.TNL.T.VLGGVLGLV.YIG.WVYC
              *   *      *         *                                              *            *   *   *  * *

Fugu     CRRRRARVRHVR----------HQLCDCKNPKPKGFYRT
Chicken  GM.WQR.RKRA.RMAQAKQYLLEKKT.....QRRMY.-----
Mouse    .V.C.HQQ.QAARMSQIKRLLSEKKT.....PHRMQ.SHNLI
Human    .V.C.H.R.QAERMSQIKRLLSEKKT.....PHRFQ.TCSPI
Rabbit   .VKC.H.RHQAQRMSQIKKLLSEKKT.....PHRLQ.TYNLL
W.whale  -AKCWH.R.RAERTSQIKRLLSEKKT.H..SHRLQ.TCSL.
Cat      .VKCWH.R.QAARMSHIKRLLSEKKT.....SHRLQ.TCNPI
              *   *
```
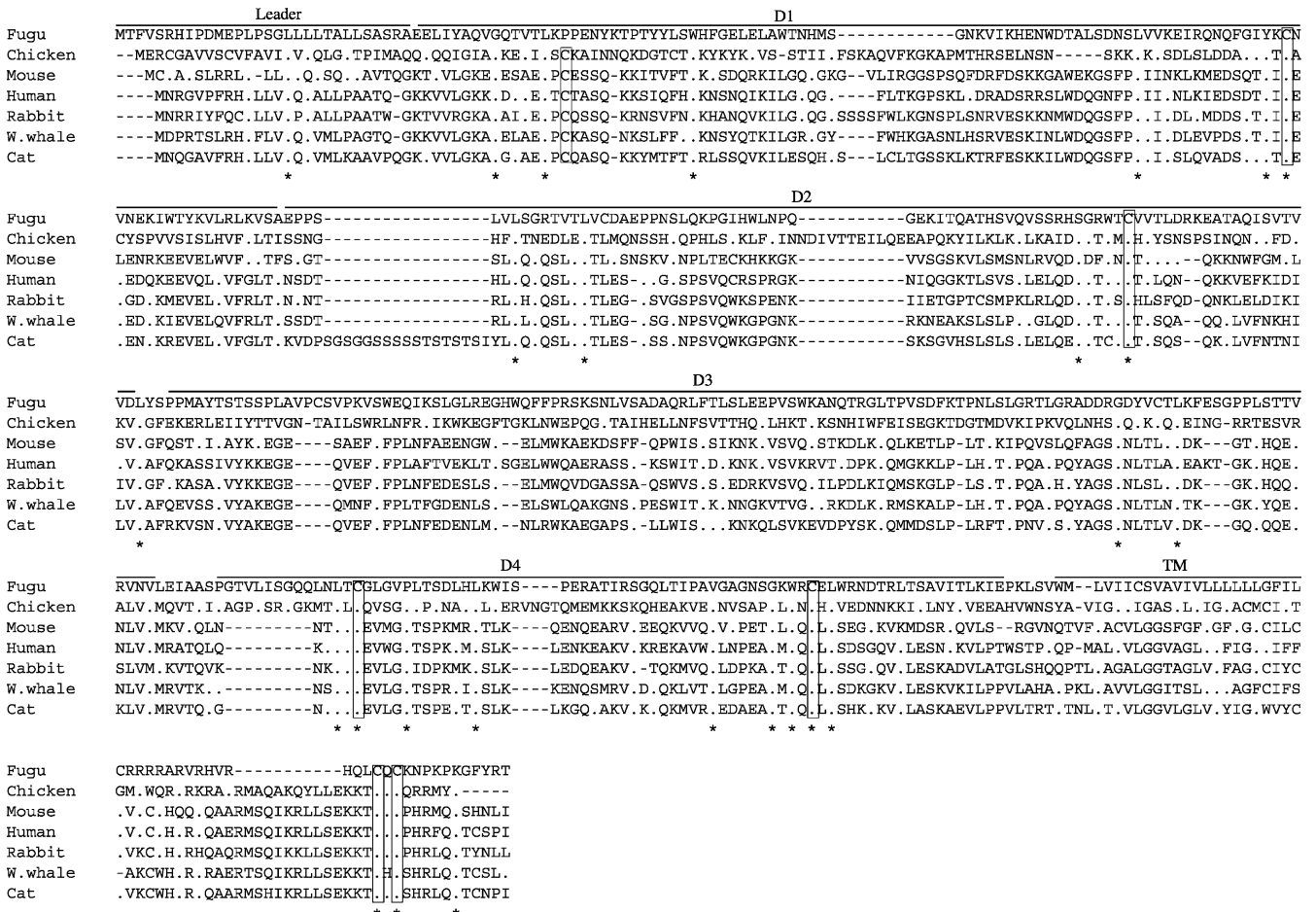
Fig. 1 Alignment of CD4 amino-acid sequences. Gaps, shown as *dashes*, have been introduced to maximize the alignment of the sequences. Residues identical with the fugu sequence are denoted by *dots*. Amino acids conserved in all sequences are indicated by *asterisks*. The conserved Cys residues are *boxed*. The leader sequence, the four Ig-like domains (D1–D4), and the transmembrane segment, and the cytoplasmic tail boundaries are indicated above the sequences. CD4 accession numbers are noted in the text. Other accession numbers are as follows: CD4, mouse (NM_013488); human (M12807); rabbit (M92840); white whale (AF071799); cat (AB000483)
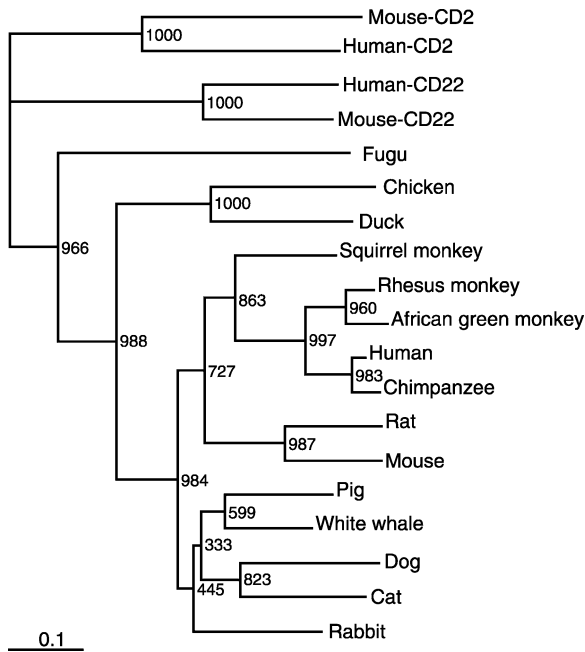
Fig. 2 Phylogenetic tree of CD4, CD2, and CD22. This unrooted tree was constructed by the neighbor-joining method, based on the amino acid alignment (Clustal W) of full-length protein sequences. Node values represent bootstrap analysis of 1,000 replicants. CD4 accession numbers are noted in the text and the legend of Fig. 1. Other accession numbers are as follows: CD2, mouse (NM_013486); human (M16445); CD22, mouse (NM_009845); human (NM_001771); CD4, squirrel monkey (Q29037); rhesus monkey (D63349); African green monkey (D86589); chimpanzee (X73323); rat (NM_012705); pig (X65629); dog (X68565)

chicken *CD4* genes was found in a scaffold of fugu genome. We named the sequence as a putative fugu *CD4* and designed primers for cDNA cloning. We cloned the full-length cDNA encoding a putative fugu *CD4* by RACE. This cDNA sequence was subjected to a BLASTX search. The clone was most similar to the complete African green monkey *CD4* (2e-5), followed by various avian and mammalian CD4 sequences.

Fugu *CD4* cDNA (accession no. AB164054) consists of 2,664 nucleotide residues, encoding a single open-reading frame of 463 amino acid residues. After cleavage of the putative 30-amino-acid hydrophobic leader sequence, a predicted mature protein would be generated. The size of
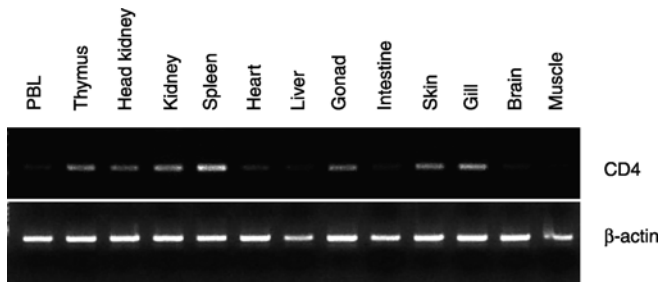

Fig. 3 Expression of fugu *CD4*. Tissue distribution of fugu *CD4* transcripts by RT-PCR is shown in the *upper panel*. Primers are located in exons 8 and 9 of the fugu *CD4* gene. *Lower panel* shows *β-actin* transcripts
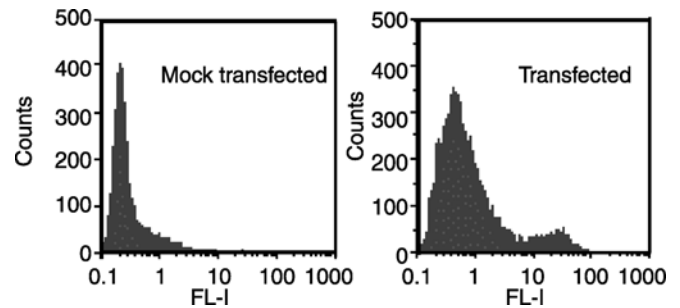

Fig. 4 Cell surface expression of fugu CD4. Fugu CD4 is expressed on the surface of COS-7 cells tranfected with pFlagCD4. Mock transfected cells were negative for Flag surface expression

the putative fugu CD4 is similar to those of other CD4 molecules. The alignment of CD4 sequences and domain analysis indicates that the fugu sequence contains four Ig-like domains, a transmembrane region, and a cytoplasmic tail like avian and mammalian CD4 molecules (Fig. 1). While cloning the fugu *CD4* cDNA, we found an allotypic variant (accession no. AB164055) with 6-bp difference resulting in one amino acid substitution ($N^{299} \rightarrow D^{299}$).

A multiple sequence alignment of CD4 amino acid sequences is shown in Fig. 1. Fugu CD4 has a percentage of identity with those of avian and mammalian species ranging from 15.4% to 19.6% (Table 2). Between chicken and human CD4 molecules, the percentage of identity is also very low (22.8%), with conserved overall structure (Koskinen et al. 1999). These values are much lower than those among mammalian species. These results indicate that, in spite of the low level of sequence identity in CD4 molecules, the basic structure of CD4 has been conserved during evolution.

A phylogenetic analysis was performed using CD4 full sequences (Fig. 2). As depicted by neighbor-joining tree, the putative fugu CD4 falls into the cluster of CD4 sequences.

Structural features of fugu CD4

The putative fugu CD4 structure is conserved with some differences from avian and mammalian CD4 molecules. The fugu CD4 has four Ig-like domains (D1–D4), like other known CD4s. The first and second Ig-like domains (D1 and D2) are known to be important for binding to MHC class II in mammals (Clayton et al. 1989; Bowman et al. 1990; Moebius et al. 1992, 1993; Huang et al. 1997). A disulfide bond within the D1 domain is conserved in the avian and mammalian CD4 molecules. However, in fugu D1 lacks a Cys residue capable of forming a disulfide bond within this domain, differing from the D1 domain of other known CD4s (Fig. 1). The second domain of fugu CD4 has a pair of Cys residues, like mouse and human CD4. However, in chicken, dog, cat, and whale the first Cys is replaced, such that no disulfide bond can be formed in D2 (Koskinen et al. 1999; Milde et al. 1993; Norimine et al. 1992; Romano et al. 1999). Thus, lack of a disulfide bond in the Ig-like domain is observed in some species. A

disulfide bond is not likely to be necessary to form Ig-like domains and a similar type of Ig-like domain is found in the sequence of CD2 (Williams et al. 1987) and D3 of CD4 (Parnes 1989) in mammals. CD4 is expressed on helper T cells that recognize antigen with MHC class II on antigen-presenting cells (Parnes 1989). In cat, dog, and cetacean, highly constitutive expression of MHC class II molecules on T cells was observed (Deeg et al. 1982; Neefjes et al. 1986; Romano et al. 1992). In these species, different functions and interactions of T cells and antigen-presenting cells are suggested (Romano et al. 1999). Like these species, the absence of a disulfide bond in D1 of fugu CD4 may give a different conformation to this molecule and affect the binding and interaction of CD4 and MHC class II. In D3, fugu CD4 has a pair of Cys residues that might form an additional disulfide bond lacking in other animal CD4s. This disulfide bond also may give the fugu CD4 molecule a different conformation. The conserved Cys residues, which might form a disulfide bond, are found in D4 of fugu CD4. D4 is involved in the dimerization of human CD4 during antigen recognition by the TCR (Wu et al. 1997; Moldovan et al. 2002). The D4 of fugu CD4 is conserved, suggesting that dimerization of fugu CD4 may occur. The conformation of the fugu extracellular domain may be different from that of other CD4 molecules due to different disulfide bonding patterns in the first three domains. These differences may affect the function of CD4-expressing cells in fugu.

Fugu CD4 has four potential N-linked glycosylation sites, and only one is conserved and located in the D4 domain (Fig. 1). The consensus p56lck tyrosine kinase motif (lck motif, Cys-X-Cys) is conserved in the cyto-

plasmic domain of fugu CD4. Lck is critical for CD4 association in the cytoplasmic region and has been implicated in T-cell maturation and signaling, and the lck gene shows a tightly controlled lymphocyte-specific expression in mammals (Marth et al. 1985; Perlmutter 1989; Veillette et al. 1989). The fugu lck gene also shows lymphoid-specific expression (Brenner et al. 2002). This result is consistent with the expression pattern of the fugu CD4 gene (Fig. 3). These observations suggest that the lck may associate with the consensus lck motif in the cytoplasmic tail of fugu CD4.

Expression of the fugu CD4 gene

Fugu CD4 was expressed in the thymus, head kidney, kidney, spleen, gonad, skin and gill (Fig. 3). Slight background expression of CD4 was found in the PBL, heart, liver, and brain. No expression was seen in the muscle. Fugu CD4 is expressed predominantly in the tissues of the immune system in agreement with the lymphocyte-specific expression of CD4 in mammals.

COS-7 cells were transfected with pFlagCD4 (Fig. 4). The expression of fugu CD4 was observed on the surface of transfected COS-7 cells. We found that fugu CD4 was efficiently expressed on 11% of COS-7 cells.

Genomic analysis of fugu CD4

Fugu CD4 and putative GAPDH genes were found in the same scaffold. A syntenic relationship between the CD4

Fig. 5 A Genomic organization of the CD4 gene. The genomic organization of the fugu CD4 gene was predicted by comparison with fugu genome draft sequence. The exons are numbered and shown in boxes separated by introns. Black and white boxes indicate coding exons and non-coding exons, respectively. B The putative promoter region of the fugu CD4 gene. The Ihr and Ihr-like sequences are underlined. The potential transcription factor binding sites are also underlined and named above the sequence. The sequence after an arrow is the first intron. An asterisk indicates transcriptional initiation site

and *GAPDH* genes in fugu was observed, as in human and chicken (Koskinen et al. 2002). This shows conserved synteny within these regions during evolution.

Fugu *CD4* gene organization was analyzed by comparing the cDNA and the scaffold sequence of fugu genomic database ver. 3.0. Fugu *CD4* spans a total of 12.1 kb and is composed of 12 exons (Fig. 5A), differing from chicken and human *CD4* genes, which consist of ten exons, with lengths 11.5 kb and 31.3 kb, respectively (Fig. 5A). Exons 5–12 almost identically correspond to exons 3–10 of other *CD4* genes (Gorman et al. 1987; Maddon et al. 1987; Koskinen et al. 2002). The fugu *CD4* gene has two additional exons within the 5′-untranslated region (5′UTR) and signal peptide coding region, respectively. In the 5′ UTR, there is a long intron of 5.6 kb and 10.4 kb in chicken and human, respectively. The fugu *CD4* gene also has two long introns of about 4.6 kb and 1.5 kb in the region upstream of the predicted start codon. The long introns of human, mouse, and chicken *CD4* genes contain an important regulatory element (Gorman et al. 1987; Koskinen et al. 2002; Donda et al. 1996). The fugu *CD4* gene may also have regulatory elements in these long introns. The fugu *CD4* gene is longer than expected. Although human and mouse *CD4* genes are much longer than the fugu *CD4* gene, the chicken *CD4* gene spans only 11.5 kb, which is almost the same size as fugu *CD4*. Most of the introns of the fugu *CD4* gene are shorter than the corresponding human and chicken introns, but the fugu gene has additional introns, making the fugu sequence longer.

In the putative promoter region of the fugu *CD4* gene, no TATA box was found, but we detected some potential transcription binding motifs, such as an Ets site, an Ikaros binding site, and Myb and MAZ binding sites, which have been shown to play important roles in the expression of *CD4* genes (Fig. 5B, Salmon et al. 1993; Sarafova and Siu 2000). These features are also conserved in the chicken *CD4* promoter (Koskinen et al. 2002).

We have cloned the first ectothermic animal *CD4* gene from the fugu. The basic structure is similar to known CD4 molecules, but with a low level of overall sequence identity and some unique structural features. Our results suggest the possibility of producing specific molecules, such as antibody or nucleotide probes, to identify subtypes of lymphoid cells, which may lead to a deeper understanding of the basic structure and function of the immune system in fish.

References

Alabyev BY, Guselnikov SV, Najakshin AM, Mechetina LV, Taranin AV (2000) CD3ε homologues in the chondrostean fish *Acipenser ruthenus*. Immunogenetics 51:1012–1020

Aparicio S, Chapman J, Stupka E, Putnam N, Chia JM, Dehal P, Christoffels A, Rash S, Hoon S, Smit A, Gelpke MD, Roach J, Oh T, Ho IY, Wong M, Detter C, Verhoef F, Predki P, Tay A, Lucas S, Richardson P, Smith SF, Clark MS, Edwards YJ, Doggett N, Zharkikh A, Tavtigian SV, Pruss D, Barnstead M, Evans C, Baden H, Powell J, Glusman G, Rowen L, Hood L, Tan YH, Elgar G, Hawkins T, Venkatesh B, Rokhsar D, Brenner S (2002) Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. Science 297:1301–1310

Bowman MR, MacFerrin KD, Schreiber SL, Burakoff SJ (1990) Identification and structural analysis of residues in the V1 region of CD4 involved in interaction with human immunodeficiency virus envelope glycoprotein gp120 and class II major histocompatibility complex molecules. Proc Natl Acad Sci USA 87:9052–9056

Brenner S, Elgar G, Sandford R, Macrae A, Venkatesh B, Aparicio S (1993) Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. Nature 366:265–268

Brenner S, Venkatesh B, Yap WH, Chou CF, Tay A, Ponniah S, Wang Y, Tan YH (2002) Conserved regulation of the lymphocyte-specific expression of lck in the Fugu and mammals. Proc Natl Acad Sci USA 99:2936–2941

Clark GJ, Tobias GH, Pietersz GA, Classon BJ, Walker ID, McKenzie IF, Deacon NJ (1988) Isolation of a cDNA clone for the murine CD4 antigen. Transplant Proc 20:45–48

Classon BJ, Tsagaratos J, McKenzie IF, Walker ID (1986) Partial primary structure of the T4 antigens of mouse and sheep: assignment of intrachain disulfide bonds. Proc Natl Acad Sci USA 83:4499–4503

Clayton LK, Sieh M, Pious DA, Reinherz EL (1989) Identification of human CD4 residues affecting class II MHC versus HIV-1 gp120 binding. Nature 339:548–551

Deeg HJ, Wulff JC, DeRose S, Sale GE, Braun M, Brown MA, Springmeyer SC, Martin PJ, Storb R (1982) Unusual distribution of Ia-like antigens on canine lymphocytes. Immunogenetics 16:445–457

Donda A, Schulz M, Burki K, De Libero G, Uematsu Y (1996) Identification and characterization of a human CD4 silencer. Eur J Immunol 26:493–500

Dumont-Drieux AM, De Parseval A, Heiber M, Salmon P, Pancino G, Sonigo P, Klatzmann D (1992) Unusual amino acid sequence of the second Ig-like domain of the feline CD4 protein. AIDS Res Hum Retroviruses 8:1581–1591

Dzialo RC, Cooper MD (1997) An amphibian CD3 homologue of the mammalian CD3 gamma and delta genes. Eur J Immunol 27:1640–1647

Fomsgaard A, Hirsch VM, Johnson PR (1992) Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus. Eur J Immunol 22:2973–2981

Gobel TW, Meier EL, Du Pasquier L (2000) Biochemical analysis of the *Xenopus laevis* TCR/CD3 complex supports the "stepwise evolution" model. Eur J Immunol 30:2775–2781

Gorman SD, Tourvieille B, Parnes JR (1987) Structure of the mouse gene encoding CD4 and an unusual transcript in brain. Proc Natl Acad Sci USA 84:7644–7648

Gustafsson K, Germana S, Sundt TM III, Sachs DH, LeGuern C (1993) Extensive allelic polymorphism in the CDR2-like region of the miniature swine CD4 molecule. J Immunol 151:1365–1370

Hague BF, Sawasdikosol S, Brown TJ, Lee K, Recker DP, Kindt TJ (1992) CD4 and its role in infection of rabbit cell lines by human immunodeficiency virus type 1. Proc Natl Acad Sci USA 89:7963–7967

Hansen JD, Strassburger P (2000) Description of an ectothermic TCR coreceptor, CD8α, in rainbow trout. J Immunol 164:3132–3139

Hansen JD, Zapata AG (1998) Lymphocyte development in fish and amphibians. Immunol Rev 166:199–220

374

Huang B, Yachou A, Fleury S, Hendrickson WA, Sekaly RP (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

Koskinen R, Lamminmaki U, Tregaskes CA, Salomonsen J, Young JR, Vainio O (1999) Cloning and modeling of the first nonmammalian CD4. J Immunol 162:4115–4121

Koskinen R, Salomonsen J, Tregaskes CA, Young JR, Goodchild M, Bumstead N, Vainio O (2002) The chicken *CD4* gene has remained conserved in evolution. Immunogenetics 54:520–525

Maddon PJ, Littman DR, Godfrey M, Maddon DE, Chess L, Axel R (1985) The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family. Cell 42:93–104

Maddon PJ, Molineaux SM, Maddon DE, Zimmerman KA, Godfrey M, Alt FW, Chess L, Axel R (1987) Structure and expression of the human and mouse T4 genes. Proc Natl Acad Sci USA 84:9155–9159

Marth JD, Peet R, Krebs EG, Perlmutter RM (1985) A lymphocyte-specific protein-tyrosine kinase gene is rearranged and over-expressed in the murine T cell lymphoma LSTRA. Cell 43:393–404

Milde KF, Conner GE, Mintz DH, Alejandro R (1993) Primary structure of the canine CD4 antigen. Biochim Biophys Acta 1172:315–318

Moebius U, Clayton LK, Abraham S, Diener A, Yunis JJ, Harrison SC, Reinherz EL (1992) Human immunodeficiency virus gp120 binding C′C″ ridge of CD4 domain 1 is also involved in interaction with class II major histocompatibility complex molecules. Proc Natl Acad Sci USA 89:12008–12012

Moebius U, Pallai P, Harrison SC, Reinherz EL (1993) Delineation of an extended surface contact area on human CD4 involved in class II major histocompatibility complex binding. Proc Natl Acad Sci USA 90:8259–8263

Moldovan MC, Yachou A, Levesque K, Wu H, Hendrickson WA, Cohen EA, Sekaly RP (2002) CD4 dimers constitute the functional component required for T cell activation. J Immunol 169:6261–6268

Nakanishi T, Fischer U, Dijkstra JM, Hasegawa S, Somamoto T, Okamoto N, Ototake M (2002) Cytotoxic T cell function in fish. Dev Comp Immunol 26:131–139

Neefjes JJ, Hensen EJ, de Kroon TI, Ploegh HL (1986) Abiochemical characterization of feline MHC products: unusually high expression of class II antigens on peripheral blood lymphocytes. Immunogenetics 23:341–347

Norimine J, Miyazawa T, Kawaguchi Y, Tohya Y, Kai C, Mikami T (1992) A cDNA encoding feline CD4 has a unique repeat sequence downstream of the V-like region. Immunology 75:74–79

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358

Park CI, Hirono I, Enomoto J, Nam BH, Aoki T (2001) Cloning of Japanese flounder *Paralichthys olivaceusCD3* cDNA and gene, and analysis of its expression. Immunogenetics 53:130–135

Parnes JR (1989) Molecular biology and function of CD4 and CD8. Adv Immunol 44:265–311

Perlmutter RM (1989) T cell signaling. Science 245:344

Romano TA, Ridgway SH, Quaranta V (1992) MHC class II molecules and immunoglobulins on peripheral blood lymphocytes of the bottlenosed dolphin, *Tursiops truncatus*. J Exp Zool 263:96–104

Romano TA, Ridgway SH, Felten DL, Quaranta V (1999) Molecular cloning and characterization of *CD4* in an aquatic mammal, the white whale *Delphinapterus leucas*. Immunogenetics 49:376–383

Ropars A, Bautz AM, Dournon C (2002) Sequencing and expression of the *CD3 γ/δ* mRNA in *Pleurodeles waltl* (urodele amphibian). Immunogenetics 54:130–138

Salmon P, Giovane A, Wasylyk B, Klatzmann D (1993) Characterization of the human *CD4* gene promoter: transcription from the *CD4* gene core promoter is tissue-specific and is activated by Ets proteins. Proc Natl Acad Sci USA 90:7739–7743

Sarafova S, Siu G (2000) Precise arrangement of factor-binding sites is required for murine *CD4* promoter function. Nucleic Acids Res 28:2664–2671

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Veillette A, Bookman MA, Horak EM, Samelson LE, Bolen JB (1989) Signal transduction through the CD4 receptor involves the activation of the internal membrane tyrosine-protein kinase p56[lck]. Nature 338:257–259

Williams AF, Barclay AN, Clark SJ, Paterson DJ, Willis AC (1987) Similarities in sequences and cellular expression between rat CD2 and CD4 antigens. J Exp Med 165:368–380

Wu H, Kwong PD, Hendrickson WA (1997) Dimeric association and segmental variability in the structure of human CD4. Nature 387:527–530