

ORIGINAL PAPER

Tracy A. Romano · Sam H. Ridgway
David L. Felten · Vito Quaranta

Molecular cloning and characterization of CD4 in an aquatic mammal, the white whale *Delphinapterus leucas*

Received: 1 July 1998 / Revised: 6 October 1998

Abstract Given the importance of the cell surface recognition protein, CD4, in immune function, the cloning and characterization of CD4 at the molecular level from an odontocete cetacean, the white whale (*Delphinapterus leucas*), was carried out. Whale CD4 cDNA contains 2662 base pairs and translates into a protein containing 455 amino acids. Whale CD4 shares 64% and 51% identity with the human and mouse CD4 protein, respectively, and is organized in a similar manner. Unlike human and mouse, however, the cytoplasmic domain, which is highly conserved, contains amino acid substitutions unique to whale. Moreover, only one of the seven potential N-linked glycosylation sites present in whale is shared with human and mouse. Evolutionarily, the whale CD4 sequence is most similar to pig and structurally similar to dog and cat, in that all lack the cysteine pair in the V2 domain. These differences suggest that CD4 may have a different secondary structure in these species, which may affect binding of class II and subsequent T-cell activation, as well as binding of viral pathogens. Interestingly, as a group, species with these CD4 characteristics all have high constitutive expression of class II molecules on T lymphocytes, suggesting potential uniqueness in the interaction of CD4, class II molecules, and the immune response. Mo-

lecular characterization of CD4 in an aquatic mammal provides information on the CD4 molecule itself and may provide insight into adaptive evolutionary changes of the immune system.

Key words CD4 · T-helper lymphocyte · Class II molecule · HIV · Whale

Introduction

CD4 is a 55000 M_r cell recognition protein on thymocytes and T-helper lymphocytes that interacts with major histocompatibility (MHC) class II molecules to initiate the immune response (Swain et al. 1984), and is the receptor for the human immunodeficiency virus (HIV) (Klatzmann et al. 1990). The structure of CD4 is comprised of four extracellular immunoglobulin-like (Ig-like) domains, a membrane-spanning segment, and a cytoplasmic tail that associates with the protein kinase, p56^{lck}, involved in T-cell activation. V1 is the primary domain involved in binding of class II molecules and HIV (Bowman et al. 1990). In addition to human and mouse (Maddon et al. 1985, 1987), CD4 was identified in several other terrestrial mammals (Clark et al. 1987; Classon et al. 1986; Fomsgaard et al. 1992; Gustafsson et al. 1990; Hague et al. 1992; Milde et al. 1993; Norimine et al. 1992). Given its importance in the immune system, CD4 was cloned and characterized in the white whale (or beluga) to gain insight into the evolutionary and/or environmental adaptations of the immune system.

Materials and methods

Animals

Whales and dolphins were maintained in accordance with federal regulations at SPAWARS Systems Center in San Diego, Calif., (accredited by the American Association for Laboratory Animal

T.A. Romano
Texas A&M University, Department of Veterinary Anatomy
and Public Health, VMA Building, Room 107A,
College Station, TX 99743, USA

T.A. Romano · S.H. Ridgway
SPAWARS Systems Center, D3503 (PLI-BS) Room 204,
49620 Beluga Road, San Diego, CA 92152, USA

T.A. Romano (✉) · V. Quaranta
The Scripps Research Institute, Department of Cell Biology,
SBR-12, 10550 N. Torrey Pines Road, La Jolla, CA 92037,
USA, E-mail: tromano@scripps.edu, Tel.: +1-619-7849871,
Fax: +1-619-7842246

D.L. Felten
Loma Linda University, 11021 Campus, Room AH321,
Loma Linda, CA 92354, USA

Care). Blood samples were collected from white whales and bottlenose dolphins by venipuncture on the ventral aspect of the tail-stock, and mixed with sodium heparin and placed at 4°C until lymphocyte isolation with Histopaque (1077) (Sigma, St. Louis, Mo.) (adaptation of methodology from Romano et al. 1992). Thymus and kidney were harvested from whales from native sanctioned hunts, or from bottlenose dolphins that died of natural causes. Organs were collected approximately 1 h postmortem, cut into small pieces, and snap-frozen in liquid nitrogen. Tissues were stored in liquid nitrogen until RNA isolation and Northern blot analysis according to the methods of Sambrook and co-workers (1989).

Generation of whale CD4

Five prime (ACCACCAGGTTCACTTCT) and 3' (GGGCA-GAGCCTGACCCTGACCTTGG) oligonucleotide primers complemented conserved regions of mouse and human CD4, and were used to amplify white whale (*Delphinapterus leucas*) cDNA by polymerase chain reaction (PCR). Amplification conditions consisted of a 2 min denaturation at 94°C, followed by 30 cycles (94°C, 1 min; 50°C, 2 min; 72°C, 3 min), and a final 72°C incubation for 7 min. The predicted 500 base pair (bp) fragment was cloned into pCR-Script SK(+) (Stratagene, San Diego, Calif.), and subsequently used to probe a whale Lambda Zap II, cDNA library custom-made with beluga thymus by Clontech Laboratories, Inc. (Palo Alto, Calif.). It was then plated on *Escherichia coli* strain XL1-Blue and screened by standard procedures (Sambrook et al. 1989). Two of the lambda ZAP clones, pBELT4.2 and pBELT4.6, showing the approximate kb of human CD4, were sequenced. Full-length CD4 was obtained by using primers to the 5' (AGAAGCAGAGGGGAAGACAG) and 3' (CAGAGG-GTTCACTTCTCTGGCC) ends of pBELT 4.6 in PCR.

DNA analysis

Whale CD4 was analyzed using NCBI BLAST and the Genetics Computer Group, Inc. (GCG) program. Phylogenetic analysis using parsimony (PAUP 3.1) (Swofford, 1993) was used to infer phylogenetic relationships among nucleotide and amino acid sequences of the coding region of CD4 genes. Bootstrap analysis (Felsenstein 1985) was used to place confidence values on individual tree nodes.

Whale CD4 antibody generation and western blot analysis

Rabbit antisera to two CD4 peptides (Trp-Lys-Gly-Pro-Gly-Asn-Lys-Arg-Lys-Asn-Glu-Ala-Lys-Ser-Leu-Ser) and (Gln-Ser-Met-Arg-Val-Ser-Asp-Gln-Gln-Lys-Leu) were produced following

standard protocol, and affinity purified (Harlow and Lane 1988). Whale and dolphin lymphocytes were lysed in 20 mM Tris-HCl, 150 mM NaCl, 0.5% NP40, 5 mM ethylene-diaminetetra-acetate, and 0.5% Tween 20, pH=7.0, in the presence of protease inhibitors for 30 min and centrifuged for 15 min at 4°C. Lysates were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis using 12% polyacrylamide gels, followed by western blot analysis (Harlow and Lane 1988).

Results

A 534 base pair (bp) fragment of CD4 amplified by PCR from whale thymus cDNA was 76% and 66%, respectively, identical to the human and mouse CD4 cDNAs. It was then used to probe a whale thymus cDNA library (see Fig. 1). The first clone, pBELT4.2, contained 1399 bp that aligned with 830 bp of the most 3' portion of human and mouse CD4 open reading frame (ORF), followed by 569 bp of 3' untranslated region (UTR). The second clone pBELT4.6 (2,017 bp) spanned 96 bp of 5' UTR, 852 bp of CD4 ORF including the ATG initiation codon, and 1069 bp of 3'UTR. Clone alignment showed pBELT4.6 identity with the first 317 bp and the last 437 bp of pBELT4.2, but the absence of 645 base pairs in the middle. Primers were generated to the beginning (containing the initiation codon), and the end sequence (3' UTR) of pBELT4.6, and used to amplify whale thymus cDNA by PCR. Sequence analysis of the resultant 2662 bp confirmed full-length whale CD4.

The nucleotide sequence and the predicted protein sequence of whale CD4 are shown in Fig. 2. Whale CD4 translates into a protein containing 455 amino acid residues. The initial methionine is followed by a characteristic leader sequence of 25 hydrophobic amino acid residues that after cleavage would leave a mature whale CD4 protein consisting of 430 amino acid residues. A region of high hydrophobicity (367–392) likely represents a transmembrane domain, followed by highly charged amino acids of the cytoplasmic domain. There are 7 potential N-linked glycosylation sites in whale CD4, two Asn-X-Thr and five Asn-X-Ser consensus sites. Cysteines are present extracellularly, and may form disulphide bonds as in other mammals.

Fig. 1 Schematic diagram showing whale CD4 cDNA clones, pBELT 4.6, and pBELT4.2, aligned with full-length beluga CD4 (DI CD4). Rectangles represent 5' and 3' untranslated regions. Start (ATG) and stop codons (TGA, TAG) are shown for each. Restriction sites and base pair numbers are indicated. The space shown in pBELT4.6 represents base pairs that are absent. Full-length CD4 was amplified by PCR using primers in the 5' and 3' regions of pBELT4.6

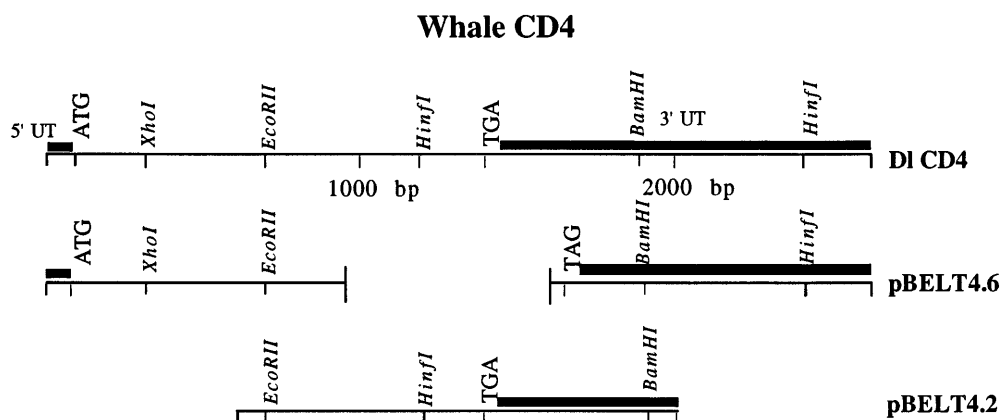


Fig. 2 Nucleotide sequence of whale CD4 (accession number AF071799) and the predicted amino acid sequence. The numbers on the right next to the nucleotide sequence indicate nucleotide position. Amino acid residues are shown by the one-letter abbreviation code below the nucleotide sequence and are numbered on the left. Nucleotides in the 5' and 3' untranslated regions are shown preceding the ATG (methionine codon) and following the TGA or termination codon, indicated with a star (3). An arrow (↓) delineates the predicted end of the leader sequence (amino acids -25 to -1) and the beginning of the mature protein (amino acid 1). Cysteines are indicated with an asterisk (*), and potential N-linked glycosylation sites are underlined

	CGGCAAGAAGCAGAGGGGAAGACAGATCCAGGCCAGAGGCCCTGCCTTTCTTTGGGGG	60
-25	AAGGCCAGGTCCTGCCTGCCTCAGCAAAGCCACAATGGACCCAGAACCTCTCTTAGG	120
	M D P R T S L R	
-17	CACCTGTTCCTGGTGTGCAACTGGTGATGCTCCAGCCGGCACTCAGGGGAAGAAAGTG	180
	H L F L V L Q L V M L P A G T Q G K K V	
4	GTGCTGGGTAAGGCAGGAGAATCGCAGAGCTGCCTGCAAAGCTTCCCAGAATAAGAGC	240
	V L G K A G E L A E L P C K A S Q <u>N K S</u>	
24	CTGTTCTTCAGCTGGAAAAATCTTACCAGACCAAGATTCTGGGGCGTCAATGGCTACTTC	300
	L F F S W K N S Y Q T K I L G R H G Y F	
44	TGGCACAAGGTGCCCTCCAACTGCACTCTCGAGTGAATCAAAAATAAACCTGTGGGAC	360
	W H K G A S N L H S R V E S K I N L W D	
64	CAAGGATCCTTTCTCTGTCATCAAGGATCTTGAAGTACCTGACTCGGGGACTTACATC	420
	Q G S F P L V I K D L E V P D S G T Y I	
84	TGCGAAGTGGAGGACAAGAAGATCGAGGTGGAATTCAGAGTTCAGATTGACTGCCAGC	480
	C E V E D K K I E V E L Q V F R L T A S	
	*	
104	TCGGACACCCGCTGTGTGCTGGGACAGAGCTGACCTGACCTTGAAGGCCCTCTGGT	540
	S D T R L L L G Q S L T L T L G E G P S G	
124	AGTAACCCCTTCAGTGAATGGAAAGGTCCAGGCAATAAAAAGGAATGAGGCCAAGAGT	600
	S N P S V Q W K G P G N K R K N E A K S	
144	CTGTCACTGCCCCAGTGGGGCTGCAGGACAGTGGCACCTGGACATGTACCCTCTCCAG	660
	L S L P Q V G L Q D S G T W T C T V S Q	
	*	
164	GCCCAGCAGACGCTGGTGTCAACAAGCACATCTGGTGTGGCCTTCCAGGAGGTCTCC	720
	A Q Q T L V F N K H I L V L A F Q E V S	
184	AGCACGTCTATGCAAGGAGGGGAGCAGATGAACCTTTCTCTCCACTCACATTCCGA	780
	S T V Y A K E G E Q M <u>N F S</u> F P L T F G	
204	GATGAAAATCTGAGCGGGAGCTGAGTTGGCTGCAGGCAAAGGGGAATTCCTCCCCGAG	840
	D E <u>N L S</u> G E L S W L Q A K G <u>N S S</u> P E	
224	TCCTGGATCACCTTCAAAATGAACAATGGGAAGGTGACTGTGGGAAGGCTCGCAAGGAC	900
	S W I T F K L N N G K V T V G A A R K D	
244	CTCAAGCTCCGCATGAGTAAGGCGCTCCCCCTCCACTGACTCTGCCCCAGGCTTTGCCT	960
	L K L R M S K A L P L H L T L P Q A L P	
264	CAGTACGCAGGTTCTGGAAACCTGACCTGAAATCTCACAAGGGGAAGTTGTATCAGGAA	1020
	Q Y A G S G <u>N L T L N L T</u> K G K L Y Q E	
284	GTGAACCTCGTGGTGTGAGAGTGACTAAGTCCCAACAGTTTACCTGTGAGGTGCTG	1080
	V N L V V M R V T K S P N S L T C E V L	
	*	
304	GGACCCAGTCCCCAGGCTGATCTGAGCTTGAAGAAGGAGAACCAGAGTATGAGGGTC	1140
	G P T S P R L I L S L K K E <u>N O S</u> M R V	
324	TCAGATCAGCAGAAGCTGGTGCCTGCTGGGCCCTGAAGCAGGATGTGGCAGTGTCTA	1200
	S D Q Q A K L V T V L G P E A A G M W Q C L	
	*	
344	CTGAGCGCAAGGGCAAAGTCTGCTGGAATCCAAGGTCAAGATTTTGCCCGTGTCTC	1260
	L S D K G K V L L E S K V K I L P P V L	
364	GCCCAGCCTGGCCGAAGCTTCTGGCTGTGGTGTGGGGGGCATACCAGTCTTCTGCTT	1320
	A H A W P K L L A V V L G G I T S L L L	
384	CTCGCTGGATTTTGCATCTCTCTGCTAAATGCTGGCACCAGCGCCGGGAGAGCGG	1380
	L A G F C I F S A K C W H R R R R A E R	
404	ACATCTCAGATCAAGAGACTCCTCAGTGAAGAAGACTGCCACTGCTCGCACCGGCTC	1440
	T S Q I K R L L S E K K T C H C S H R L	
424	CAGAAGACATGTAGTCTCACCTGAGGCCAGCCAGGAGGAGCTTCCCATCTGCAGCCTC	1500
	Q K T C S L T ★	
	CCCAGCTGGCTCCCTGCATTTCTGTGGTCTCCTGGGCTGCGGACCAGATGAATGT	1560
	AGCCGGCATAGCGCCTCTGTCCACCTCTGCCCTCCCGTCCAGTGGGCCCTGGGTCCGC	1620
	TCCTCAGAGGCTCACTCACACCCCTCTTCCCAATTTCTCTTTCACTCAAACCTTAGCCC	1680
	TTCTTTGATGATTTCTTCTCCACCCCTTCCCTCACTGCTCACTTGCATTCAGGGGACA	1740
	CGTGGGACCGGCCCTGCCTGCCCTGGAGGGTGGGCTGGGTGTCGAAAGCAGGAAGACA	1800
	GGCTGTCACTGTTCGGGAGAGGACCTTGGGACCAGAGAAGGAGGACTAGCCAGAGGT	1860
	CACACAGCCATCAAGGACGGATCCAGATCCAAAGCTCCTTCTGACTGCCAGCTGTG	1920
	CTGCTCCATTCGCTGCATCTTACCTTGTGTTAGCAGAACCACAGACTTACATCTTGAC	1980
	CTGAGCACAGACCAGCACCCCTGGACACATTCGTGCGCAGATACAGTTCCCTTACCACAG	2040
	GCAGCGTGGTGTGTCTACAAGTCAAGTTTACATCTCTAGTCCATTTTGAACCTCATGAAG	2100
	ATCCATCAAGACTGGTGTGACAGGACCAAGATTACCATCTCCAGCTTCTAGCCAGAAA	2160
	TGAAATGCAGAGGCTAGGTGATGGCTAGGTTTCCACCGGCTAGTACGTGGGCTAGCCC	2220
	AGGTCCTTTGACCTCTAGTCCACTGCTATCTTTGAACACAGGGACAATATCCACACAG	2280
	ATCTCTGCAGTTGGCTGGTCAATAGATGCCCGTACTACCTTGGATTCCTTACCACAGG	2340
	GAGGTTGCCCTCAGAGAGTTTCAGGGCCCTCCCTGCCGAGGATCCCTCCTCATCAGCAGA	2400
	ATCCATTAAGAGGACAAAACCTGTAGTTCTCCTAATCAGGGCACAGGCTGCCAGCAC	2460
	AGTGACCCGCACTCACCTGTGCAACCTTCTAGAAGGTGAACTTTCCAGGGGGACAG	2520
	GGTGGAGAGCTGGCCAGAGAGTGAACCTCTGAGGTTACCTTGGGCTCAGAGATACCC	2580
	ACTGGGCTCAGTCTCCCTGCCCATACCCCTTCACTCCCTCCCAAAATGATTCAGGA	2640
	GACAAGAAAATGGGTTTACAA	2662

To detect expression of CD4 messenger RNA in whale, total RNA was isolated from whale and dolphin, thymus and kidney, and used in northern blot analysis using pBELT4.2 as a probe. Northern blot analysis confirms CD4 message in the thymus of both species and its absence in the kidney (Fig. 3).

Antibodies to beluga CD4 were used for western blot analysis on whale and dolphin peripheral blood lymphocyte lysates. Analysis with one of the affinity purified whale CD4 peptide-specific antibodies demonstrates an approximate 65 000 M_r protein in beluga and an approximate 60 000 M_r protein in the bottlenose dolphin, which is absent when labeling cell lysates with purified pre-immune serum alone (Fig. 4).

Whale CD4 is similar to human and mouse, but with differences (Figs. 5A, B). Alignment of whale CD4 with human shows 76% identity at the nucleotide level and 64% at the amino acid level, and with mouse 68% and 51%, respectively. As with human and mouse CD4, whale CD4 has four extracellular immunoglobulin-like domains consisting of variable (V-like), and joining (J-like) regions, a membrane-spanning segment, and a cytoplasmic domain. The V1 domain of beluga shares similar identity with immunoglobulin κ light-chain V regions as the human and mouse sequences (30–35%), and contains the characteristic pair of cysteine residues separated by 67 amino acids that form disulfide links. Domains V2 - V4, however, appear to be truncated im-

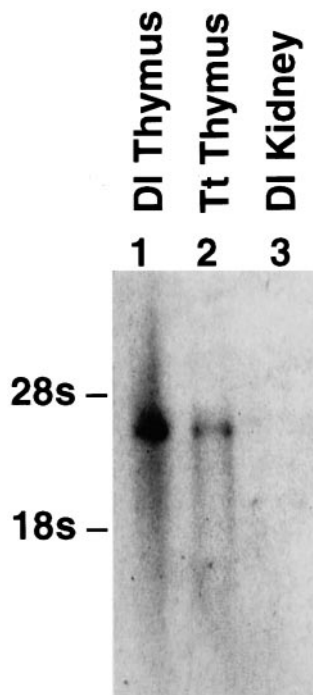


Fig. 3 Northern blot analysis of RNA isolated from beluga (DI) and bottlenose dolphin (Tt) thymus, and kidney, probed with pBELT4.2. The positions of the 28 s and 18 s ribosomal RNA are indicated. Lanes 1 and 2 show a band at approximately 3.0 kb, corresponding to CD4 message in beluga and bottlenose dolphin thymus. No band is present in Lane 3, which contains kidney RNA

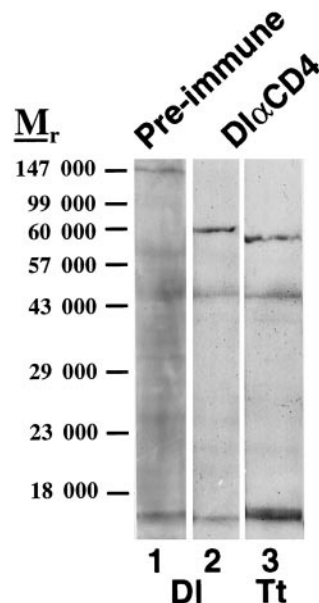


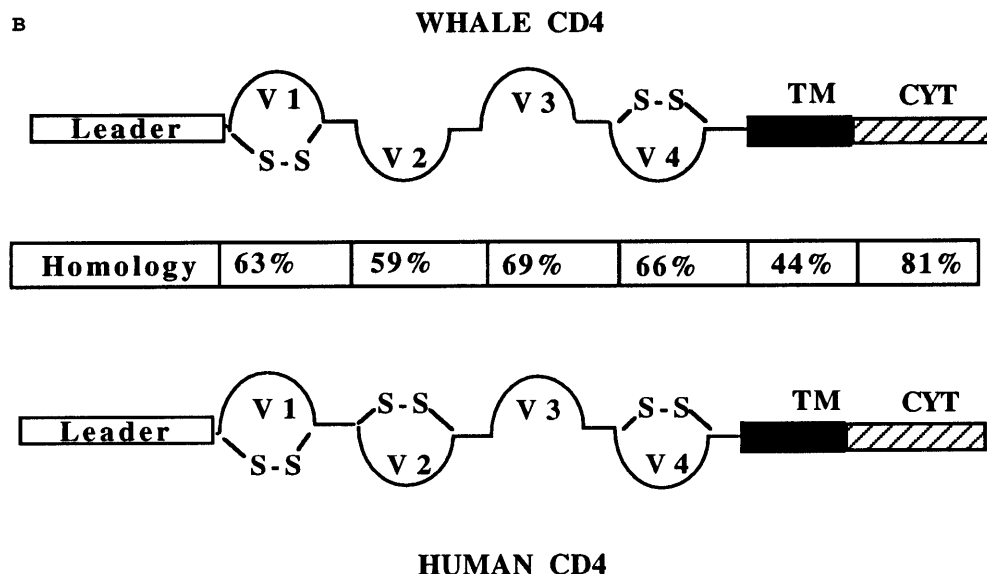
Fig. 4 Western blot analysis of beluga (DI) and bottlenose dolphin (Tt) peripheral blood lymphocyte cell lysates labeled with anti-beluga CD4 (DI α CD4). Lane 1 contains beluga lymphocyte lysates labeled with pre-immune rabbit IgG as a control. Lanes 2 and 3 contain beluga and bottlenose dolphin lymphocyte lysates labeled with anti-beluga CD4. The antibody recognizes an approximate 65 000 M_r band in the beluga, and an approximate 60 000 M_r band in the bottlenose dolphin

munoglobulin domains as in human and mouse. Overall identity of beluga CD4 domains compared with human CD4 domains are shown in Fig. 5B. Domain 3 lacks cysteines as does human and mouse V3, while domain 4 (V4) in beluga also contains a pair of cysteines in the relative positions as the human and mouse sequence.

The cytoplasmic domain has the highest level of identity among whale, human, and mouse (81% whale and human; 63% whale and mouse). The whale cytoplasmic domain contains the cysteine motif that has been shown to be essential for association with p56^{lck}, a src-like tyrosine kinase involved in lymphocyte signal transduction. The serine residues thought to be involved in dissociation of p56^{lck}, are also present.

Differences between whale, human, and mouse include structural changes in V2, unique amino acid substitutions in the cytoplasmic domain despite the high level of conservation, and additional potential glycosylation sites. V2 in the beluga lacks a cysteine pair. A tryptophan is substituted for the first cysteine of the human/mouse pair. Although the cytoplasmic domain is highly conserved among species, whale CD4 has differences. Whale substitutes a histidine for glutamine in the cysteine motif (Fig. 6). An arginine (amino acid 400) replaces the conserved glutamine in all other species, and a threonine (amino acid 404) replaces the highly conserved methionine in other species. Only one of seven potential N-linked glycosylation sites in whale CD4 is shared with human and mouse. Whale CD4 was

Fig. 5 B Schematic diagram of whale CD4 vs human CD4. Domains are shown as (Leader) leader sequence, (V1-V4) V-like, (TM) transmembrane, (CYT) cytoplasmic. The (J1-J4) or joining regions, immediately follow each corresponding V region. Domains that contain cysteine pairs are indicated by disulfide links (S-S). Amino acid identity between whale and human for V1-V4, TM, and CYT is shown



		+	+	•	•	*+*	430																																					
WHALE	KC	W	H	R	R	R	R	R	A	E	R	T	S	Q	I	K	R	L	L	S	E	K	T	C	H	C	S	H	R	L	Q	K	T	C	S	L	T							
MOUSE	R	-	R	-	Q	-	Q	-	Q	-	A	-	M	-	-	-	-	-	-	-	-	-	-	-	Q	-	P	-	M	-	-	S	H	N	-	I								
RAT	R	-	R	-	Q	-	Q	-	A	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	P	-	M	-	-	S	H	N	-	I								
CHIMP	R	-	R	-	-	-	Q	-	Q	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	P	-	F	-	-	-	-	-	-	P	I							
HUMAN	R	-	R	-	-	-	Q	-	-	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	P	-	F	-	-	-	-	-	-	-	P	I						
RHEMONK	R	-	R	-	-	-	Q	-	-	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	P	-	F	-	-	-	-	-	-	-	-	P	I					
CAT	-	-	-	-	-	-	Q	-	A	-	M	-	H	-	-	-	-	-	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	N	P	I			
DOG	-	-	-	-	R	-	-	Q	-	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	-	I				
RABBIT	-	-	R	-	-	H	Q	-	Q	-	M	-	-	-	-	-	-	-	-	-	-	-	-	-	Q	-	P	-	-	-	-	-	-	-	-	-	-	-	-	-	Y	N	-	L

Fig. 6 Alignment of the cytoplasmic domains of whale CD4 with those of terrestrial mammals. Published CD4 sequences were aligned for mouse, rat, chimpanzee human, rhesus monkey, cat, dog, and rabbit. *Numbering* is based on the mature whale sequence. *Dashes* indicate amino acid identity with the whale sequence. The consensus motif (-C-X-C-(P)) found to be associated with p56^{lck} is boxed. Conserved cysteines in the consensus motif are labeled with an asterisk (*), and conserved serine residues implicated in phosphorylation are shown (●). Unique amino acids to the whale cytoplasmic domain are indicated (+)

whale, and the inability to form a disulfide bond in V2 may give the whale CD4 molecule a different conformation, which may affect binding and functional interactions of CD4. The same is true for dog (Milde et al. 1993), cat (Norimine et al. 1992), and rabbit (Hague et al. 1992). In rabbit however, there is a cysteine at position 144 that may form a disulfide bond with the conserved cysteine at 159 (Hague et al. 1992). The change in V2 may reflect evolutionary trends, given the close relatedness of whale CD4 to dog and cat CD4 as reflected in the single most parsimonious tree of CD4 and the proposed evolutionary relationships between the cetaceans and the carnivores, and/or ungulates (Slijper 1962). Unfortunately, the sequence of V2 in pig and sheep, both ungulates, are not available for comparison.

The differences in V2 may reflect functional changes. CD4 is primarily expressed on T-helper lym-

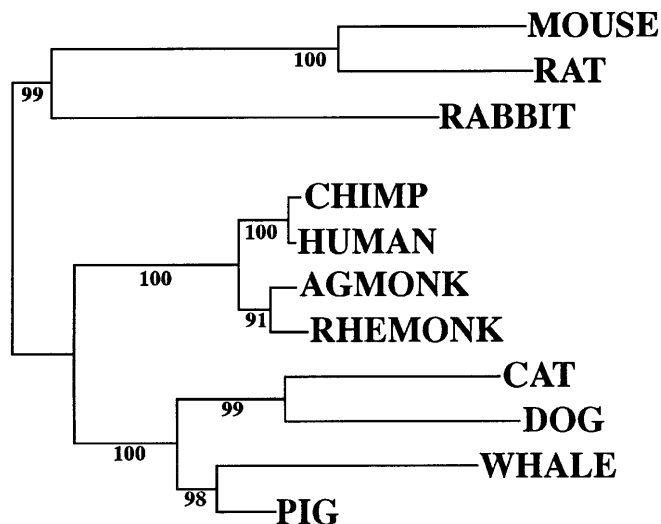


Fig. 7 Phylogram representing the single most parsimonious tree for CD4 amino acid sequences of Whale and the published sequences of mouse, rat, rabbit, chimpanzee, human, African green monkey (AGMONK), rhesus monkey (RHEMONK), cat, dog, and pig. The *horizontal* branch lengths are proportional to the number of nucleotide substitutions estimated to have occurred on those branches. *Numbers* indicate bootstrap confidence values (500 replications) for individual tree nodes

phocytes that recognize antigen when presented by cells bearing class II MHC molecules (Swain et al. 1984). The interaction between CD4 and class II molecules is critical during thymic development and for T-cell activation in an immune response (Zuniga-Pflucker et al. 1989). Dog, cat, and cetacean all have a high constitutive expression of class II molecules on T lymphocytes (Deeg et al. 1982; Neefjes et al. 1986; Romano et al. 1992). This may suggest different functions and interactions of T-cells and antigen presenting cells in the context of class II, CD4, and the T-cell receptor in the immune response of these animals.

The cytoplasmic domain of CD4 is highly conserved across species. The CD4 cytoplasmic tail is associated with a src-like tyrosine protein kinase, p56^{lck}, which is expressed in high levels in T-cells and is involved in signal transduction during T-cell activation (Glaichenhaus et al., 1991). The important region for p56^{lck} association has the motif $+-+X\text{-Cys-X-Cys-(Pro)}$ where + denotes a basic residue and X an amino acid residue other than cysteine (Turner et al. 1990), followed by a proline or serine. Alignment of this region in the whale with other species shows conservation of the lysine, the threonine preceding the first cysteine, and both cysteines. However, in the whale, a histidine (a charged amino acid) replaces the glutamine located between the cysteines. The two other unique amino acid substitutions are R400 and T404. These may affect T-cell signalling in the whale.

Whale CD4 has seven potential glycosylation sites, and shares only one in common with mouse (four sites) and human (two sites). In number, whale is most similar to dog (five sites). Differences in number and location of glycosylation sites could alter the shape of CD4 in these species and effect binding interactions.

Clone pBELT4.6 may prove to be a variant form of CD4 in the whale. This clone contains the 5' end of CD4 containing the methionine initiation codon and joins up with full-length CD4 in the 3' UTR. When lined up with full-length beluga CD4, the clone remains in open reading frame and contains its own termination codon. Variant forms of CD4 have not been found in the immune system of other mammals, but have been found in mouse and human brain (Klatzman et al. 1990; Longberg et al. 1988). Interestingly, when aligned with human CD4, this clone lacks the transmembrane domain of CD4, suggesting a soluble form of the molecule. Soluble CD4 is elevated in the serum of patients in several diseases, and recombinant soluble CD4 has been shown to inhibit HIV infection (Deen et al. 1988) and SIV infection (Watanabe et al. 1989). The presence and significance of this variant form of CD4 in the beluga is currently being investigated in regards to the immune system and brain.

Cloning of CD4 in whale is a further indication that the basic molecular components of the immune system are conserved across species; however, some of the unique structural features of whale CD4 may be relevant to adaptive functions. Further studies may provide insights into the adaptation of the mammalian immune system to the ocean environment.

Acknowledgments The authors thank Dr. Alan McLachlan for his guidance, Francine Frasier for technical assistance, Dr. Gregory O'Corry-Crowe for help with the phylogenetic analyses, Dr. Eric Jensen and Dr. Bill Van Bonn and the animal care staff at the SPAWARS Systems Center, Dr. Anthony Pelletier for review of the manuscript, Raymond Tarpley, Robert Suydam, and the people of Churchill, Manitoba, and Pt. Lay, Alaska, for their support in obtaining tissue samples. This work was supported by a grant from the National Science Foundation MCB-9317267, and a grant from the Office of Naval Research N00014-96-1-0619.

References

- 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
- Clark SJ, Jefferies WA, Barclay AN, Gagnon J, Williams AF (1987) Peptide and nucleotide sequences of rat CD4 (W3/25) antigen: Evidence for derivation from a structure with four immunoglobulin-related domains. *Proc Natl Acad Sci USA* 84:1649-1653
- Classon BJ, Tsagaratos J, McKenzie IFC, 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
- 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
- Deen KC, McDougal JS, Inacker R, Folea-Wasserman G, Arthros J, Rosenberg J, Maddon PJ, Axel R, Sweet RW (1988) A soluble form of CD4 (T4) protein inhibits AIDS virus infection. *Nature* 331:82-86
- Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791
- 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
- Glaichenhaus N, Shastri N, Littman DR, Turner JM (1991) Requirement for association of p56^{lck} with CD4 in antigen-specific signal transduction in T-cells. *Cell* 64:511-520
- Gustaffson K, Germana S, Sundt III TM, 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, Sawadikosol 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
- Harlow E, Lane D (1988) *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory, New York
- Klatzman DR, McDougal JS, Maddon PJ (1990) The CD4 molecule and HIV infection. *Immunodeficiency Rev* 2:43-66
- Longberg N, Gettner SN, Lacy E, Littman D (1988) Mouse brain CD4 transcripts encode only the COOH-terminal half of the protein. *Mol Cell Biol* 8:2224-2228
- 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
- Milde KF, Conner GE, Mintz DH, Alejandro R (1993) Primary structure of the canine CD4 antigen. *Biochim Biophys Acta* 1172:315-318
- Neefjes JJ, Hensen EJ, de Kroon TIP, Ploegh HL (1986) A biochemical 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 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
- 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

- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor
- Slijper EJ (1962) Evolution and external appearance. In: *Whales*. Basic Books, New York, pp 58–92
- Swain SL, Dialynas DP, Fitch FW, English M (1984) Monoclonal antibody to L3T4 blocks the function of T-cells specific for class 2 major histocompatibility complex antigens. *J Immunol* 132:118–1123
- Swofford DL (1993) PAUP: Phylogenetic analysis using parsimony, version 3.1. Computer program distributed by the Natural History Survey, Champaign, IL
- Turner JM, Brodsky MH, Irving BA, Levin SD, Perimutter RM, Littman DR Interaction of the unique N-terminal region of tyrosine kinase p56^{lck} with cytoplasmic domains of CD4 and CD8 is mediated by cysteine motifs. *Cell* 60:755–765
- Watanabe M, Reiman KA, DeLong PA, Liu T, Fisher RA, Letvin NL (1989) Effect of recombinant soluble CD4 in rhesus monkeys infected with simian immunodeficiency virus of macaques. *Nature* 337:267–270
- Zuniga-Pflucker JC, McCarthy SA, Weston M, Longo DL, Singer A, Kruisbeek AM (1989) Role of CD4 in thymocyte selection and maturation. *J Exp Med* 169:2085–2096