Comparison of CD45 Extracellular Domain Sequences from Divergent Vertebrate Species Suggests the Conservation of Three Fibronectin Type III Domains[1,2]

Meinoshin Okumura,* R. James Matthews,[4]* Bruce Robb,* Gary W. Litman,† Peer Bork,‡ and Matthew L. Thomas[3,5]*

Mammalian CD45 is a transmembrane protein tyrosine phosphatase expressed by all nucleated cells of hematopoietic origin. In lymphocytes, CD45 is required for Ag-induced signal transduction due to its ability to positively regulate Src family members. The mechanisms by which CD45 function is regulated are unknown. Indeed, the interactions of CD45 extracellular domains are largely undefined. To gain insight into potentially important regions of the extracellular domain, we sought to identify conserved features from divergent species. cDNAs encoding the putative CD45 homologue from *Heterodontus francisci* (horned shark) were isolated. The cDNA sequence predicts a protein of 1200 amino acids that contains a 452-amino acid extracellular domain, a 22-amino acid transmembrane region, and a 703-amino acid cytoplasmic domain. Alignment searches revealed that the *Heterodontus* cytoplasmic domain sequence was most identical to mammalian CD45 and a transmembrane protein tyrosine phosphatase sequence identified from chickens, ChPTPλ. A dendrogram with other transmembrane protein tyrosine phosphatase sequences suggest that the *Heterodontus* and chicken sequences represents CD45 orthologues for their respective species. Analysis of vertebrate CD45 extracellular domain sequences indicates the conservation of three structural regions: a region containing potential *O*-linked carbohydrate sites, a cysteine-containing region, and a region containing three fibronectin type III domains. For each vertebrate species, multiple isoforms are generated by alternative splicing of three exons that encode a portion of the region containing potential *O*-linked glycosylation sites. These studies provide evidence for a conservation in CD45 extracellular domain structure between divergent species and provide a basis for understanding CD45 extracellular domain interactions. *The Journal of Immunology,* **1996, 157: 1569–1575.**

Transmembrane protein tyrosine phosphatases (PTPases)[6] are a large group of enzymes conserved throughout metazoan evolution. Mammalian transmembrane PTPases can be grouped into at least six subfamilies based on the identity of PTPase domain sequences and general structural features (1, 2). Ligands for most transmembrane PTPases have not been identified. However, members of one subfamily of transmembrane PTPases, PTPμ and PTPκ, have been shown to interact in a homotypic fashion (3–5), and another transmembrane PTPase, PTPζ, has been shown to interact with neural adhesion molecule N-CAM via carbohydrate groups and with contactin (6, 7). It is possible that other transmembrane PTPases may also bind to proteins or carbohydrates on opposing cells and be involved in cellular recognition. Understanding the interactions of transmembrane PTPase extracellular domains is important for determining how these enzymes regulate cellular physiology.

CD45 is a PTPase expressed by all nucleated cells of hematopoietic origin. The generation of CD45-deficient lymphocyte lines demonstrates that CD45 expression is required for efficient signaling through the Ag receptors (8–12). Further support for this idea has come from the development of mice essentially ablated in CD45 expression due to gene targeting (13). CD45-deficient mice are blocked in thymocyte development and are unable to proceed with positive selection. B cells from these mice respond poorly to BCR cross-linking, supporting the notion that CD45 is required for Ag-receptor signal transduction.

The block in Ag-induced signal transduction cascade in CD45-deficient cells is apparently due to the dysregulation of Src family members (14–16). In vitro, CD45 will dephosphorylate the negative regulatory site of $p56^{lck}$ and $p59^{fyn}$ (17, 18). In CD45-deficient cells, both $p56^{lck}$ and $p59^{fyn}$ demonstrate increased phosphorylation of the negative regulatory site and accordingly, decreased kinase activity. Since both of these Src family members have been implicated in TCR signaling, it is likely that the inability to sufficiently activate $p56^{lck}$ and $p59^{fyn}$ results in the inability to signal through the Ag receptor. In T cell clones, CD45 functions to increase $p56^{lck}$ kinase activity in resting cells (14, 19), suggesting that both the phosphatase and kinase are active before stimulation.

*Howard Hughes Medical Institute and the Departments of Pathology and Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110; †Department of Pediatrics, University of South Florida, FL 33701; and ‡EMBL, Heidelberg, Germany

Received for publication November 22, 1995. Accepted for publication June 7, 1996.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by United States Public Health Service Grant AI-26363 (M.L.T.) and R37 AI23338 (G.W.L.).

[2] The sequence reported in this paper has been deposited in the GenBank data base with the accession number of U34750.

[3] Address correspondence and reprint requests to Dr. Matthew L. Thomas, Washington University School of Medicine, 660 South Euclid Avenue, Box 8118, St. Louis, MO 63110.

[4] Present address: University of Wales College of Medicine, Heath Park, Cardiff CF4 4XX, Wales, U.K.

[5] M.L.T. is an investigator of the Howard Hughes Medical Institute.

[6] Abbreviation used in this paper: PTPase, transmembrane protein tyrosine phosphatase.

0022-1767/96/$02.00

How CD45 extracellular domain affects cellular functions is unknown. The extracellular domain contains numerous sites for *N*-linked glycosylation, and the analysis of carbohydrate content suggests that most sites are occupied (20). The immediate amino-terminal region is enriched in *O*-linked carbohydrate sites. This region is encoded by six exons and is approximately 200 amino acids in length (21–23). Three of the exons are alternatively spliced to yield as many as eight isoforms (24–27). The pattern of isoform expression is highly regulated in lymphocyte activation and differentiation, suggesting that the particular isoforms expressed are important to cellular responses. While the interactions of the extracellular domain are largely undefined, it is possible that CD45 carbohydrate groups may mediate binding to an unknown ligand (28). Indeed, the B cell-specific transmembrane lectin, CD22, will bind to CD45 if properly sialylated (29–31). Importantly, different cell types differentially glycosylate CD45 (28, 32), suggesting that there may be specific interactions for other lectins that have not been defined. Although the interactions for different CD45 isoforms have not been defined, it has been suggested based on differential expression of distinct CD45 isoforms in either CD45-deficient cells or in transgenic mice, that the expression of different isoforms affects signal transduction through the TCR (33–36).

Human, rat, and mouse CD45 extracellular domains share only approximately 25% identical residues (28). The domain structure for this region is unknown although it has been previously suggested to contain at least one fibronectin type III domain (37). To gain further insights into this region, we sought to identify features that have been conserved in divergent vertebrate species. Here we describe the isolation of *Heterodontus francisci* (horned shark) cDNAs that encode the putative CD45 orthologue.

Materials and Methods

Isolation of cDNA clones

A 1.2-kb *Dra II* restriction fragment from mouse CD45 cDNA (2271–3518 bp) (25) was used to screen 5×10^5 plaques from a cDNA library derived from *Heterodontus francisci* spleen (38). Filters were hybridized as described, except that 35% formamide was used in the hybridization mixture (39). Filters were washed at a final stringency of 7.5 mM NaCl/.75 mM trisodium citrate/pH 7.7/00.1% SDS at 52°C. Seventeen individual clones were isolated, of which the largest contained a 2.9-kb cDNA insert. The 2.9-kb cDNA was incomplete and by comparison with mammalian CD45 cDNAs, an estimated approximately 1 kb of additional sequence would be required to contain the entire reading frame. To obtain the entire open reading frame, reverse transcriptase reaction followed by rapid amplification of cDNA ends PCR was performed. Specific oligonucleotides complementing *Heterodontus* CD45 cDNA were used to prime *Heterodontus* spleen RNA for reverse transcription according to the manufacturer's instructions (Invitrogen, San Diego, CA). The reaction was initiated by adding 50 U of avian myoblastosis virus reversed transcriptase (Life Science, St. Petersburg, FL). Following cDNA synthesis, the RNA template was degraded by the addition of 2 U RNase H for 20 min at 37°C. cDNA was purified on a glass matrix (BIO 101, La Jolla, CA). An oligodeoxyadenosine extension was added to the cDNA by the addition of 1 U of terminal deoxynucleotidyl transferase in 5 mM Tris-HCl, pH 9.0/25 mM KCl/0.75 mM $MgCl_2$/1% Triton X-100/0.5 mM dATP for 5 min at 37°C, followed by 10 min at 65°C. The cDNA was amplified by PCR using oligo(dT) and a nested 3′-derived oligo to *Heterodontus* CD45 cDNA. Restriction endonuclease sites were incorporated into both the 5′ and 3′ oligonucleotides to facilitate cloning. Thirty cycles were performed using the conditions of 1 min at 94°C, 1 min at 50°C, and 2 min at 72°C. Amplification with further nested 3′ oligonucleotides was required for specificity. Three rounds of reverse transcription-extension PCR were required to complete the isolations of cDNAs containing the entire coding region.

To examine the presence of alternatively spliced isoforms, spleen RNA was used for reverse transcriptase PCR using oligonucleotides derived to the immediate 5′ end of the cDNA and position 430–447. The reaction was performed according to the manufacturer's instructions (Invitrogen).

Sequence analysis of the 2.9-kb cDNA clone resulted in two regions of potential sequence ambiguity. One of these regions contained a potential frame shift mutation. To confirm the sequence of these two regions, oligonucleotides were derived to sequences that corresponded to the 5′ and 3′ boundaries of mouse exons 14 and 15. *Heterodontus* genomic liver DNA (500 ng) was used as a template to hybridize to 80 ng of the oligonucleotide primers in a buffer of 20 mM Tris-HCl, pH 8.4/50 mM KCl/1.5 mM $MgCl_2$/0.2 mM dNTPs/0.1 mg of BSA per ml/0.1% Triton X-100. The reaction was initiated with 1.25 U of *Taq* polymerase with 30 temperature cycles of 94°C for 30 s, 48°C for 1 min, and 72°C for 30 s. The genomic sequence resolved an apparent insertion of a single base that occurred during the PCR amplification of the cDNA.

Northern blot analysis

Total mRNA from horned shark tissues was isolated using guanidinium isothiocyanate as previously described (38, 40). Five micrograms was separated on a 1% agarose gel containing 2.2 M formaldehyde and 20 mM 3-[*N*-morpholine]propanesulfonic acid, pH 6.8. After extensive washing in H_2O, the gel was briefly stained with ethidium bromide, extensively washed again, photographed, and transferred to either Nitroplus 2000 (MCI, Westboro, MA) or nitrocellulose (Bio-Rad, Richmond, CA). Filters were prehybridized overnight at 42°C in 50% formamide containing 0.75 M NaCl, 0.075 M sodium citrate, 10× Denhardt's solution, and 50 mM sodium phosphate, pH 5.7. Hybridization was in the same solution with the addition of 10% dextran sulfate, 0.1% SDS, and 0.1 mg herring sperm DNA. A 2.1-kb *Eco*RI cDNA fragment encoding the *Heterodontus* PTPase domains was ^{32}P-labeled by random priming and 1×10^6 cpm/ml added to the hybridization solution. After hybridizing overnight at 42°C, the filter was washed in a final stringency of 30 mM NaCl, 3 mM sodium citrate, and 0.1% SDS at 55°C. Bands were visualized by autoradiography.

DNA sequencing

DNA fragments were subcloned into Bluescript (Stratagene, La Jolla, CA) and sequenced on both strands by the oligonucleotide-directed priming as previously described (25). The shark sequence was subjected to a variety of computer-aided analysis methods (41). CLUSTALW was used for multiple sequence alignment, construction of phylogenetic trees, and bootstrapping (42). For displaying the trees, treetool (written by Dr. M. Maciukenas, EMBL, Heidelberg, Germany) was used. The alignments of the divergent CD45 extracellular domains were further thread by hand through the fibronectin type III domain consensus to adjust for maximal homology and compatibility with structural data.

Results and Discussion

Isolation and analysis of Heterodontus *cDNAs sequences*

To define conserved features of CD45 extracellular domain, we sought to compare the extracellular domain of mammalian CD45 with other vertebrate species. A *Heterodontus* cDNA library was screened with a probe derived from murine CD45. Multiple cDNA clones were isolated, of which a single clone contained the largest inserts of 2.9 kb. By limited sequence analysis and restriction endonuclease digestion, the smaller cDNA clones were contained within the 2.9-kb cDNA (data not shown). The 2.9-kb cDNA insert was subcloned and sequenced in its entirety. The largest open reading frame did not contain an initiator methionine or a leader sequence. By comparison with mammalian CD45, it appeared that the *Heterodontus* cDNA clone contained a partial fragment encoding a portion of the extracellular domain, the transmembrane region, the entire cytoplasmic domain, and a portion of the 3′-untranslated region (Fig. 1).

Northern blot analysis showed that the *Heterodontus* cDNA hybridized predominately to a 5.0-kb mRNA present in spleen but not liver or pancreas (Fig. 2). This supports the idea that the 2.9-kb cDNA represents a partial cDNA fragment. The *Heterodontus* cDNA library was repeatedly screened with probes derived from the 5′ region of the cDNA. However, no cDNAs that extended further 5′ were obtained. Furthermore, no other related cDNAs were obtained, suggesting that the cDNAs isolated represent the only cross-hybridizing cDNAs in the library under the conditions used.

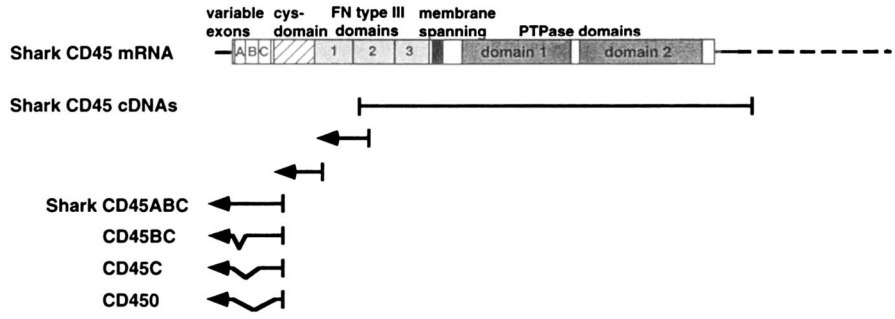

FIGURE 1. Schematic diagram of *Heterodontus* mRNA and corresponding cDNAs. The arrows indicate the direction of extension for the PCR cloning. cDNAs encompassing alternative spliced forms are indicated by the variation in exons A, B, and C. The dashed line indicates the putative region of mRNA not isolated as cDNA.

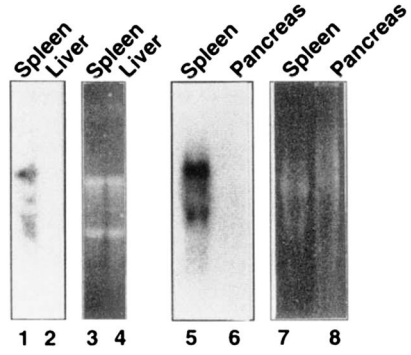

FIGURE 2. Hematopoietic expression of *Heterodontus* CD45 mRNA. *Heterodontus* total mRNA from spleen and liver (*lanes 1–4*) or spleen and pancreas (*lanes 5–8*) were subjected to electrophoresis on formaldehyde-denaturing agarose gels, stained with ethidium bromide, photographed (*lanes 3, 4, 7,* and *8*), and transferred to nitrocellulose. Filters were hybridized with a ^{32}P-labeled 2.1-kb cDNA probe encompassing the 3′ end and encoding the PTPase domains. Bands were visualized by autoradiography (*lanes 1, 2, 5,* and *6*). Spleen mRNA shown in *lanes 1* and *3* is from a separate preparation from that shown in *lanes 5* and *7*.

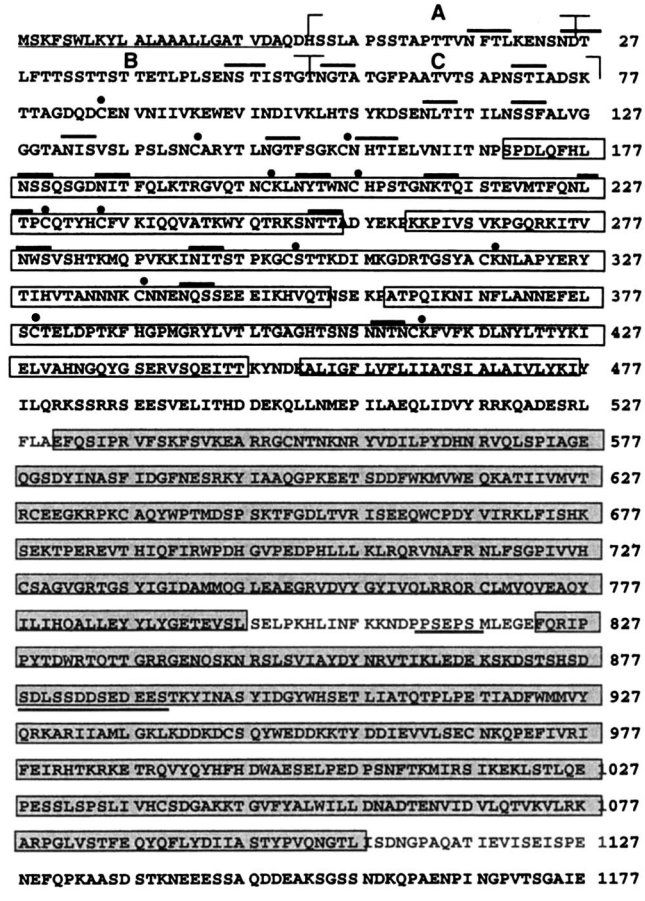

FIGURE 3. Predicted protein sequence of *Heterodontus* CD45. The leader sequence is underlined. Sequences encoded by alternative spliced exons are bracketed. Cysteine residues are indicated by a dot. Potential *N*-linked carbohydrate sites are overlined. The fibronectin type III domains are boxed and lightly shaded. The transmembrane region is boxed. The PTPase domains are boxed and darkly shaded. Potential glycogen synthase kinase 3 and casein kinase II sites are underlined with a bold line. Numbering is based on the predicted amino acid of the mature protein. The single letter amino acid code is used.

To obtain cDNAs encompassing the entire coding region, rapid amplification of cDNA ends PCR was performed. Three consecutive rounds of extension resulted in the isolation of cDNAs that contained an initiator methionine and predicted a long open reading frame (Fig. 1). The sequence encoding the predicted initiator methionine is flanked by a sequence that conforms to the consensus sequence for initiation of translation sites (data not shown) (43). An in-frame stop codon is located six nucleotides upstream, further suggesting that the entire coding region was obtained. The composite sequence of the four cDNA clones contains 3914 bp. The length of the composite cDNA sequence differs from the mRNA size by approximately 1 kb. By comparison to mammalian CD45 mRNA structure, it is likely that we have not isolated cDNAs corresponding to the 3′-untranslated region.

The predicted initiator methionine is located at position 91. The open reading frame encodes a protein of 1200 amino acids. A potential 23-amino acid leader sequence is present. The mature protein is predicted to be 1177 amino acids and contain an extracellular domain sequence of 452 amino acids, a 22-amino acid transmembrane region, and a 703-amino acid cytoplasmic domain (Fig. 3). Similar to other transmembrane PTPases, the cytoplasmic domain contains two protein tyrosine phosphatase subdomains.

Heterodontus cytoplasmic domain sequence was compared with other transmembrane PTPases (Fig. 4). The *Heterodontus* sequence is most closely related to mammalian CD45 cDNA sequences and a chicken transmembrane PTPase sequence, ChPTPλ (Fig. 4) (44). Human and *Heterodontus* share 76% identity in the first PTPase domain, and 56% sequence identity in the second PTPase domain (Fig. 5). The *Heterodontus* and ChPTPλ PTPase sequences are similarly related. The chicken PTPase sequence is more closely related to the human than to the shark PTPase sequence, 84 and 76% identity for each domain, respectively.

Hallmark sequences found in mammalian CD45 but not found in other transmembrane PTPases are contained in both the *Heterodontus* and chicken cytoplasmic domains. For example, all contain an approximately 20-amino acid insertion in the second phosphatase domain, which is enriched in acidic residues and is a

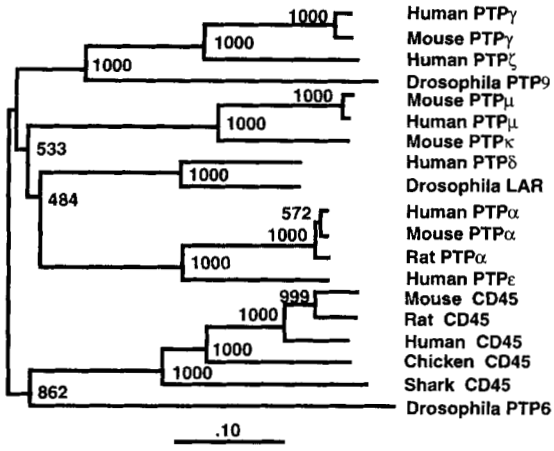

FIGURE 4. Dendrogram of transmembrane PTPases. Comparisons of PTPase domains were constructed using CLUSTALW (42). Both domains of *Heterodontus* CD45 cluster together with mammalian and avian CD45 with high reliability (1000 is the number of bootstrap trials). The horizontal bar indicates the fractional divergence (0.1 = 10% divergence).

potential casein kinase II phosphorylation site. Furthermore, all contain a glycogen synthase kinase 3 site immediately preceding the second phosphatase domain, and an extended carboxyl-terminal tail of approximately 70 amino acids. The sequence relatedness and conservation of unique features supports the idea that the chicken and *Heterodontus* proteins are orthologues of mammalian CD45.

Alternative splicing of Heterodontus *CD45*

Alternative splicing of exons encoding a region containing potential *O*-linked glycosylation sites is a feature common to mammalian CD45 mRNA. For mammals, there are three exons: 4, 5, and 6 (also termed exons A, B, and C), which are alternatively spliced. To determine whether the corresponding region in *Heterodontus* CD45 is also subject to alternative splicing, sequences encoding this region were amplified by reverse transcription reaction, followed by PCR. Analysis of the derived cDNA clones identified three additional isoforms (Fig. 1). From the analysis of the four isoforms, it can be inferred that there are three exons that are alternatively spliced. The four isoforms contain either exons ABC, BC, C, or a deletion of all three exons. These data demonstrate that splicing of the region encoding the *O*-linked carbohydrate sites is a conserved feature of CD45 from elasmobranchs to mammals. This strongly suggests that variation in the length of the region containing *O*-linked carbohydrate sites is important to CD45 function.

Comparison of the elasmobranch, avian, and mammalian CD45 extracellular domain sequences indicates a conservation of domain structure. To identify features of CD45 extracellular domain conserved through evolution, the human, chicken, and horned shark sequences were compared. Three conserved structural features were revealed: 1) an amino-terminal region enriched in potential *O*-linked carbohydrate sites that is subject to a change in length due to alternative splicing of three exons; 2) a short cysteine-containing region; and 3) a region containing three fibronectin type III domains (Figs. 5 and 6).

For all three species, the amino-terminal region is enriched in serines and threonines, contains multiple α-helix breaking residues, and does not contain any cysteine residues. However, there is low sequence identity between the vertebrate sequences.

The absolute size for the amino-terminal region varies between mammals, chicken, and horned shark (Fig. 5). The length of each of the variable exons is markedly different between different vertebrate species. Human exon 4 (A) is larger than that of rat and mouse (28). This finding has been attributed to a change in splice sites (45). Interestingly, in chicken CD45, there is a small exon between exons A and B that is not apparently alternatively spliced (44). Taken together, the data suggest that the length of the amino-terminal regions is dynamic and changes between species. Importantly, the low sequence identity between mammals, chicken, and *Heterodontus* supports the idea that the putative *O*-linked carbohydrate structures are important for functional interactions with potential ligands.

The putative *O*-linked carbohydrate region is followed by a cysteine-containing region. Comparisons between vertebrate CD45 sequences indicate that it is variable in length (Fig. 6*A*) but, significantly, there is a conserved cysteine residue at the beginning of this region. In addition, there is an odd number of cysteine residues in this region and the absolute number varies between one for chicken CD45 and five for human CD45. It is possible, therefore, that this region is linked by at least one disulfide bridge to regions further carboxyl-terminal.

The identification of a single fibronectin type III domain in mammalian CD45 has been previously reported. However, the alignments were insufficient to identify additional fibronectin type III domains (37). By comparing mammalian, avian, and elasmobranch CD45, the presence of two additional fibronectin type III domains can be demonstrated (Fig. 6*B*). Each of the three domains contains signature amino acids found in most fibronectin type III domains. These domains are flanked by an amino-terminal proline residue and either a serine or threonine carboxyl-terminal residue. Fibronectin type III domains are composed of a series of β-strands that form two anti-parallel β sheets, similar to Ig domain structures (46, 47). For each of the fibronectin type III domains in CD45, six potential β-strands can be identified, each containing sequences similar to those found in other fibronectin type III domains. The identification of conserved amino acids for each of the predicted β-strands further supports the notion that each of these domains will fold into the fibronectin type III domain structure.

Previously, it has been reported for ChPTPλ that there is a region of similarity in the extracellular domain to a portion of spectrin (44). In making comparisons between vertebrate species, the similarity to fibronectin type III domains is much higher, suggesting that the potential for a spectrin-like domain in CD45 is not likely to be correct.

The first fibronectin type III domain of chicken CD45 is significantly longer than that found in CD45 from other species. This is due to a large insertion found between strands c and e (Fig. 6*B*). Interestingly, at a similar position, an insertion is also found in mouse CD45. This insertion is not found in either rat or human CD45 sequences, suggesting that insertion sequences are tolerated but not necessarily essential for function. For both chicken and mouse, the insertion sequence contains multiple cysteine residues, suggesting that this region is stabilized by a disulfide bond(s).

Variation in length between the fibronectin type III domains is often found in the region between the predicted β-strands. This is the region that is predicted to vary in length between the fibronectin type III domains of CD45 from various species. It is interesting to note that this intra-strand region contains the placement of many of the potential *N*-linked carbohydrate sites.

While the overall domain structure appears to be conserved between mammals, chicken, and horned shark CD45, sequence identity of the extracellular region is very low; only approximately

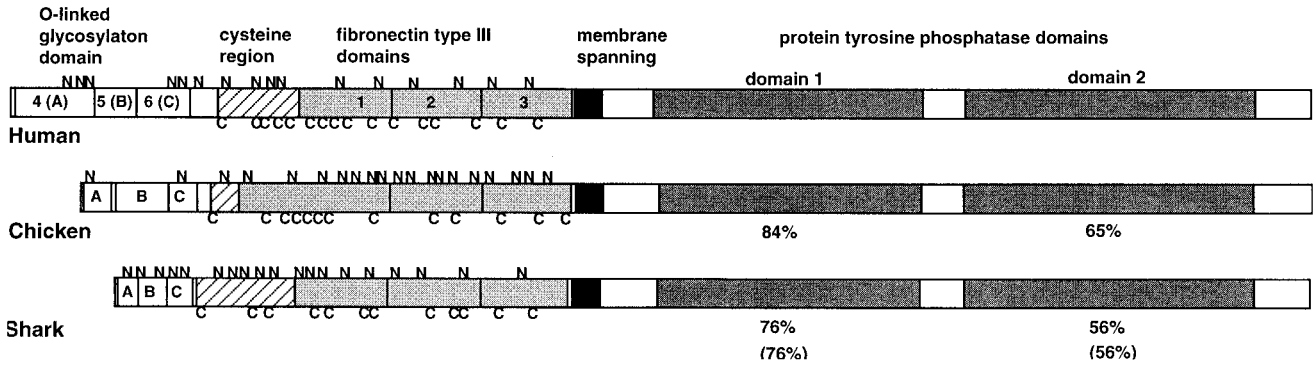

FIGURE 5. Schematic comparison of vertebrate CD45. Exons that are alternatively spliced are designated A, B, and C. Potential sites of *N*-linked glycosylation are shown by the N above the bar. Location of cysteine residues in the extracellular domain is shown by the C below the bar. The percent identity with human PTPase domains is shown below the bar. Percentage in parentheses is the percent identity between chicken and *Heterodontus* PTPase domains.

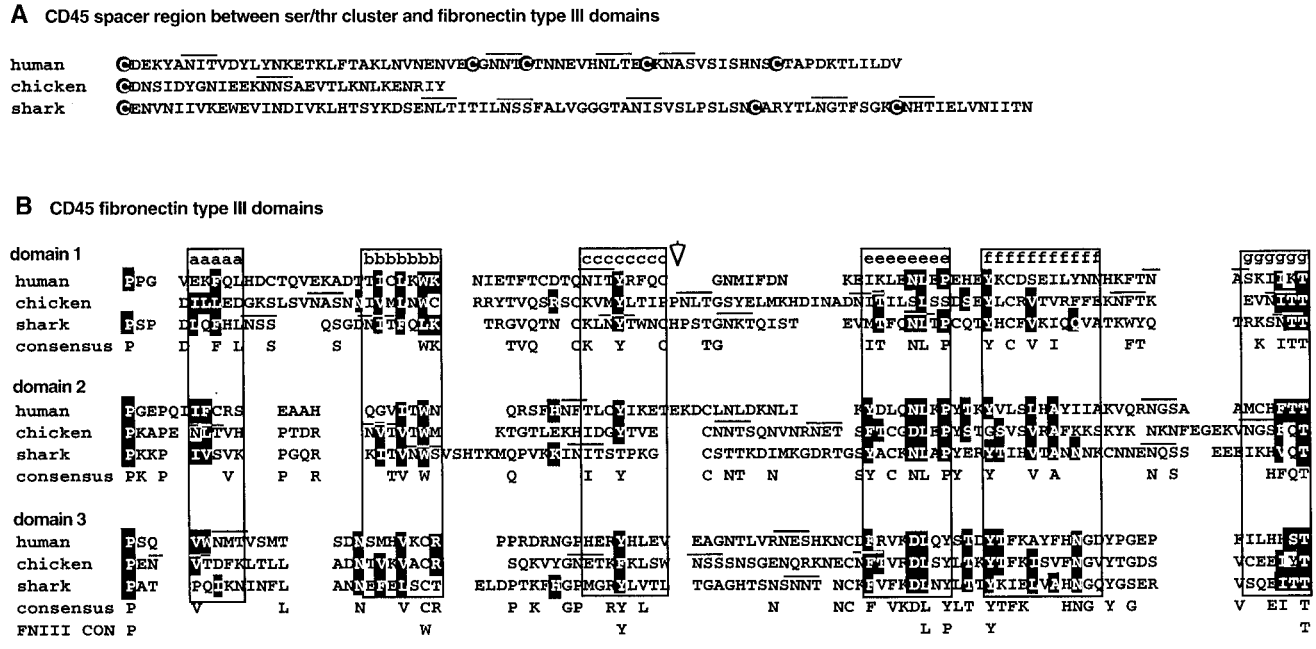

FIGURE 6. Comparison of vertebrate CD45 extracellular domain sequences. *Heterodontus* CD45 sequences are compared with chicken and human. *A*, Comparison of the cysteine-containing region. The sequence length and number of cysteines vary between vertebrate species. Additionally, there is poor sequence identity. Cysteine residues are highlighted and the potential *N*-linked carbohydrates sites are overlined. *B*, Comparison of vertebrate CD45 fibronectin type III domains. Amino acids in the majority of the domains that are either conserved or conservatively substituted are highlighted. Potential *N*-linked carbohydrate sites are overlined. The predicted β-strands found in fibronectin type III domains are boxed and correspond to strands a, b, c, e, f, and g, which are denoted above each predicted strand. The point of the insertion found in chicken and mouse CD45 is designated by the open arrow. The amino acids shown in the FNIII CON line are found in most fibronectin type III domains.

20% identical residues. However, there are a large number of *N*-linked carbohydrate sites, supporting the idea that the fibronectin type III domains serve as a platform for the carbohydrate groups.

The placement and number of cysteine residues is markedly different between vertebrate species. While the avian sequences have an additional cysteine residue near the transmembrane region, the mammalian and horned shark sequences contain an even number of cysteine residues, consistent with the notion that most if not all the cysteine residues in the extracellular domain will form disulfide bonds. The different location of cysteine residues between species suggests that the placement of disulfide bonds will vary between species. Disulfide bonds are known to vary in position between highly related domain structures (48).

It is interesting that the fibronectin type III domains of CD45 are enriched in cysteine residues. This is an unusual feature not commonly found in other fibronectin type III domains. To our knowledge, there has been only one documented example of a disulfide bond within a fibronectin domain (49). This feature may be important in maintaining the positioning of carbohydrate structures. Interestingly, the amino-terminal domains of some cytokine receptors are similar in structure to fibronectin type III domains and contain two to three disulfide bridges that vary in position between different cytokine receptors (48).

The low conservation of protein sequence of vertebrate extracellular domains is consistent with the idea that interactions with potential CD45 ligands are mediated through the carbohydrate groups. In contrast, the transmembrane and cytoplasmic domains are more highly conserved between vertebrate species, suggesting that protein-protein interactions for these domains may be more important. In support of this idea, we have recently demonstrated

that the interaction between mouse CD45 and CD45-associated protein is through the transmembrane region (50).

Transmembrane PTPases can be grouped into at least six subfamilies. All transmembrane PTPases with a substantial extracellular region contain fibronectin type III domains. The identification of additional fibronectin type III domains in the CD45 extracellular regions suggests that CD45 is more similar to other transmembrane PTPases than previously appreciated.

In mammals, CD45 is expressed by all nucleated cells of hematopoietic origin. CD45 expression has not been detected on any other cell type, suggesting that CD45 function is unique to the immune system. Indeed, the development of CD45-deficient lines and mice supports this conjecture (8–13). Sharks diverged from mammals 450 million years ago. The identification of an apparent CD45 homologue in elasmobranchs suggests that CD45 function is required in all vertebrate immune systems and supports a model in which all vertebrate immune systems have common regulatory features. Interestingly, shark lymphocytes, similar to mammal lymphocytes, contain Ag receptors (51, 52). In mammals, CD45 has been demonstrated to be essential for lymphocyte Ag receptor signal transduction (8, 9, 11–13). Thus, it is possible that the requirement for CD45 in Ag receptor signaling is a feature conserved throughout vertebrate species.

The comparison of the avian and elasmobranch sequences with mammalian CD45 supports the idea that they represent CD45 orthologues. However, it is possible that the chicken and *Heterodontus* sequences represent subfamily members rather than orthologues. We view this to be unlikely for the following reasons. First, the PTPase domain sequences are most closely related to each other and contain unique features not found in other transmembrane PTPase sequences. Second, the alignment of the extracellular domains demonstrates similar structural features despite poor sequence conservation. Third, the isolation of multiple cDNAs encoding the amino-terminal region indicates alternative splicing of a region enriched in potential *O*-linked carbohydrate sites. This feature has not been demonstrated for other transmembrane PTPases. Fourth, Northern blot analysis suggests hematopoietic-specific expression of the *Heterodontus* mRNA. The ChPTPλ mRNA also appears to be primarily expressed by hematopoietic cells (44). Finally, repeated screening of the *Heterodontus* cDNA library did not reveal other related cDNAs. Indeed, no other cDNAs within a single species have been identified that could be classified as a CD45 subfamily member. Thus, it appears that the elasmobranch and avian sequences identified represent CD45 orthologues and that this subfamily of transmembrane PTPases may contain only one member, CD45.

References

1. Mourey, R. J., and J. E. Dixon. 1994. Protein tyrosine phosphatases: characterization of extracellular and intracellular domains. *Curr. Opin. Genet. Dev. 4:31.*
2. Stone, R., and J. Dixon. 1994. Protein-tyrosine phosphatases. *J. Biol. Chem. 269:31323.*
3. Gebbink, M. F. B. G., G. C. M. Zondag, R. W. Wubbolts, R. L. Beijersbergen, I. van Etten, and W. H. Moolenaar. 1993. Cell-cell adhesion mediated by a receptor-like protein tyrosine phosphatase. *J. Biol. Chem. 268:16101.*
4. Brady-Kalnay, S. M., A. J. Flint, N. K. Tonks, A. Singer, and H. Westphal. 1993. Homophilic binding of PTPμ, a receptor-type protein tyrosine phosphatase, can mediate cell-cell aggregation. *J. Cell Biol. 122:961.*
5. Sap, J., Y. Jiang, D. Friedlander, M. Grumet, and J. Schlessinger. 1994. Receptor tyrosine phosphatase R-PTP-κ mediates homophilic binding. *Mol. Cell. Biol. 14:1.*
6. Milev, P., D. Friedlander, T. Sakurai, L. Karthikeyan, M. Flad, R. Margolis, and M. Grumet. 1994. Interactions of the chondroitin sulfate proteoglycan, the extracellular domain of a receptor-type protein tyrosine phosphatase, with neurons, glia, and neural cell adhesion molecules. *J. Cell Biol. 127:1703.*
7. Peles, E., M. Nativ, P. L. Campbell, T. Sakurai, R. Martinez, S. Lev, D. O. Clary, J. Schilling, G. Barnea, G. D. Plowman, M. Grumet, and J. Schlessinger. 1995. The carbonic anhydrase domain of receptor tyrosine phosphatase β is a functional ligand for the axonal cell recognition molecule contactin. *Cell 82:251.*
8. Pingel, J. T., and M. L. Thomas. 1989. Evidence that the leukocyte-common antigen is required for antigen-induced T lymphocyte proliferation. *Cell 58:1055.*
9. Justement, L. B., K. S. Campbell, N. C. Chien, and J. C. Cambier. 1991. Regulation of B cell antigen receptor signal transduction and phosphorylation by CD45. *Science 252:1839.*
10. Peyron, J. F., S. Verma, R. de Waal Malefyt, J. Sancho, C. Terhorst, and H. Spits. 1991. The CD45 protein tyrosine phosphatase is required for the completion of the activation program leading to lymphokine production in the Jurkat human T cell line. *Int. Immunol. 3:1357.*
11. Weaver, C. T., J. T. Pingel, J. O. Nelson, and M. L. Thomas. 1991. $CD8^+$ T-cell clones deficient in the expression of the CD45 protein tyrosine phosphatase have impaired responses to T-cell receptor stimuli. *Mol. Cell. Biol. 11:4415.*
12. Koretzky, G. A., J. Picus, M. L. Thomas, and A. Weiss. 1990. Tyrosine phosphatase CD45 is essential for coupling T-cell antigen receptor to the phosphatidyl inositol pathway. *Nature 346:66.*
13. Kishihara, K., J. Penninger, V. A. Wallace, T. M. Kundig, K. Kawai, A. Wakeham, E. Timms, K. Pfeffer, P. S. Ohashi, M. L. Thomas, C. Furlonger, C. J. Paige, and T. W. Mak. 1993. Normal B lymphocyte development but impaired T cell maturation in CD45-exon6 protein tyrosine phosphatase-deficient mice. *Cell 74:143.*
14. Cahir McFarland, E. D., T. R. Hurley, J. T. Pingel, B. M. Sefton, A. Shaw, and M. L. Thomas. 1993. Correlation between Src-family member regulation by the protein tyrosine phosphatase, CD45, and transmembrane signaling through the T-cell receptor. *Proc. Natl. Acad. Sci. USA 90:1402.*
15. Sieh, M., J. B. Bolen, and A. Weiss. 1993. CD45 specifically modulates binding of Lck to a phosphopeptide encompassing the negative regulatory tyrosine of Lck. *EMBO J. 12:315.*
16. Shiroo, M., L. Goff, M. Biffen, E. Shivnan, and D. Alexander. 1992. CD45 tyrosine phosphatase-activated $p59^{fyn}$ couples the T cell antigen receptor to pathways of diacylglycerol production, protein kinase C activation and calcium influx. *EMBO J. 11:4887.*
17. Mustelin, T., and A. Altman. 1990. Dephosphorylation and activation of the T cell tyrosine kinase $p56^{lck}$ by the leukocyte common antigen (CD45). *Oncogene 5:809.*
18. Mustelin, T., T. Pessa-Morikawa, M. Autero, M. Gassmann, L. Andersson, C. G. Gahmberg, and P. Burn. 1992. Regulation of the $p59^{fyn}$ protein tyrosine kinase by the CD45 phosphotyrosine phosphatase. *Eur. J. Immunol. 22:1173.*
19. Pingel, J. T., E. D. Cahir McFarland, and M. L. Thomas. 1994. Activation of CD45-deficient T cell clones by lectin mitogens but not anti-Thy-1. *Int. Immunol. 6:169.*
20. Brown, W. R. A., A. N. Barclay, C. A. Sunderland, and A. F. Williams. 1981. Identification of a glycophorin-like molecules at the cell surface of rat thymocytes. *Nature 289:456.*
21. Hall, L. R., M. Streuli, S. F. Schlossman, and H. Saito. 1988. Complete exon-intron organization of the human leukocyte common antigen (CD45) gene. *J. Immunol. 141:2781.*
22. Saga, Y., J.-S. Tung, F.-W. Shen, T. C. Pancoast, and E. A. Boyse. 1988. Organization of the Ly-5 gene. *Mol. Cell. Biol. 8:4889.*
23. Johnson, N. A., C. M. Meyer, J. T. Pingel, and M. L. Thomas. 1989. Sequence conservation in potential regulatory regions of the mouse and human leukocyte common antigen gene. *J. Biol. Chem. 264:6220.*
24. Barclay, A. N., D. I. Jackson, A. C. Willis, and A. F. Williams. 1987. Lymphocyte specific heterogeneity in the rat leucocyte common antigen (T200) is due to differences in polypeptide sequences near the NH_2-terminus. *EMBO J. 6:1259.*
25. Thomas, M. L., P. J. Reynolds, A. Chain, Y. Ben-Neriah, and I. S. Trowbridge. 1987. B-cell variant of mouse T200 (Ly-5): evidence for alternative mRNA splicing. *Proc. Natl. Acad. Sci. USA 84:5360.*
26. Saga, Y., J.-S. Tung, F.-W. Shen, and E. A. Boyse. 1987. Alternative use of 5′ exons in the specification of Ly-5 isoforms distinguishing hematopoietic cell lineages. *Proc. Natl. Acad. Sci. USA 84:5364.*
27. Streuli, M., L. R. Hall, Y. Saga, S. F. Schlossman, and H. Saito. 1987. Differential usage of three exons generates at least five different mRNAs encoding human leukocyte common antigens. *J. Exp. Med. 166:1548.*
28. Thomas, M. L. 1989. The leukocyte common antigen family. *Annu. Rev. Immunol. 7:339.*
29. Powell, L. D., D. Sgroi, E. R. Sjoberg, I. Stamenkovic, and A. Varki. 1993. Natural ligands of the B cell adhesion molecule CD22β carry *N*-linked oligosaccharides with α-2,6-sialic acids that are required for recognition. *J. Biol. Chem. 268:7019.*
30. Munro, S., B. J. E. G. Bast, K. J. Colley, and T. F. Tedder. 1992. The B lymphocyte surface antigen CD75 is not an α-2,6-sialytransferase but is a carbohydrate antigen, the production of which requires the enzyme. *Cell 68:1003.*
31. Stamenkovic, I., D. Sgroi, A. Aruffo, M. S. Sy, and T. Anderson. 1991. The B lymphocyte adhesion molecule CD22 interacts with leukocyte common antigen CD45RO on T cells and α-2,6-sialytransferase, CD75, on B cells. *Cell 66:1133.*
32. Lefrancois, L. 1987. Expression of carbohydrate differentiation antigens during ontogeny of the murine thymus. *J. Immunol. 139:2220.*
33. Chui, D., C. J. Ong, P. Johnson, H-S. Teh, and J. D. Marth. 1994. Specific CD45 isoforms differentially regulate T cell receptor signaling. *EMBO J. 13:798.*
34. Novak, T. J., D. Farber, D. Leitenberg, S-C. Hong, P. Johnson, and K. Bottomly. 1994. Isoforms of the transmembrane tyrosine phosphatase CD45 differentially affect T cell recognition. *Immunity 1:109.*
35. Ong, C. J., D. Chui, H-S. Teh, and J. D. Marth. 1994. Thymic CD45 tyrosine phosphatase regulates apoptosis and MHC-restricted negative selection. *J. Immunol. 152:3793.*

36. McKenney, D. W., H. Onodera, L. Gorman, T. Mimura, and D. M. Rothstein. 1995. Distinct isoforms of the CD45 protein-tyrosine phosphatase differentially regulate interleukin 2 secretion and activation signal pathways involving Vav in T cells. *J. Biol. Chem. 270:24949.*
37. Bork, P., and R. F. Doolittle. 1993. Fibronectin type III modules in the receptor phosphatase CD45 and tapeworm antigens. *Protein Sci. 2:1185.*
38. Kokubu, F., K. Hinds, R. Litman, M. J. Shamblott, and G. W. Litman. 1988. Complete structure and organization of immunoglobulin heavy chain constant region genes in a phylogenetically primitive vertebrate. *EMBO J. 7:1979.*
39. Matthews, R. J., E. D. Cahir, and M. L. Thomas. 1990. Identification of an additional member of the protein-tyrosine-phosphatase family: evidence for alternative slicing in the tyrosine phosphatase domain. *Proc. Natl. Acad. Sci. USA 87:4444.*
40. Matthews, R. J., D. B. Bowne, E. Flores, and M. L. Thomas. 1992. Characterization of hematopoietic intracellular protein tyrosine phosphatases: description of a phosphatase containing an SH2 domain and another enriched in proline-, glutamic acid-, serine-, and threonine-rich sequences. *Mol. Cell. Biol. 12:2396.*
41. Koonin, E. V., P. Bork, and C. Sander. 1994. Yeast chromosome III: new gene functions. *EMBO J. 13:493.*
42. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 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.*
43. Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol. 115:887.*
44. Fang, K. S., K. Barker, M. Sudol, and H. Hanafusa. 1994. A transmembrane protein-tyrosine phosphatase contains spectrin-like repeats in its extracellular domain. *J. Biol. Chem. 269:14056.*
45. Jackson, D. I., and A. N. Barclay. 1989. The extra segments of sequence in rat leucocyte common antigen (L-CA) are derived by alternative splicing of only three exons and show extensive *O*-linked glycosylation. *Immunogenetics 29:281.*
46. Leahy, D. J., W. Hendrickson, I. Aukhil, and H. Erickson. 1992. Structure of a fibronectin type III domain from tenascin phased by MAD analysis of the selenomethionyl protein. *Science 258:987.*
47. Main, A., T. Harvey, M. Baron, J. Boyd, and I. Campbell. 1992. The three-dimensional structure of the tenth type III module of fibronectin: an insight into RGD-mediated interactions. *Cell 71:671.*
48. Bork, P., L. Holm, and C. Sander. 1994. The immunoglobulin fold: structural classification, sequence patterns and common core. *J. Mol. Biol. 242:309.*
49. Huber, A., Y. Wang, A. J. Bieber, and P. J. Bjorkman. 1994. The crystal structure of tandem type III fibronectin domains from *Drosophila* neuroglian at 2.0 A. *Neuron 12:717.*
50. Cahir McFarland, E. D., and M. L. Thomas. 1995. CD45 protein tyrosine phosphatase associates with the WW domain-containing protein, CD45AP, through the transmembrane region. *J. Biol. Chem. 270:28103.*
51. Shamblott, M. J., and G. W. Litman. 1989. Genomic organization and sequences of immunoglobulin light chain genes in a primitive vertebrate suggest coevolution of immunoglobulin gene organization. *EMBO J. 8:3733.*
52. Rast, J., and G. Litman. 1995. T-cell receptor gene homologs are present in the most primitive jawed vertebrates. *Proc. Natl. Acad. Sci. USA 91:9248.*