Immunology 2003 **108** 288–295

Cloning of the complete rat immunoglobulin δ gene: evolutionary implications

YAOFENG ZHAO & LENNART HAMMARSTRÖM *Center for Biotechnology and Center for Oral Biology, Department of Bioscience at Novum, Karolinska Institutet, Huddinge, Sweden*

SUMMARY

The recent discovery of a Cδ encoding gene in artiodactyls has raised questions regarding the evolution of the gene. In the present study, we have analysed the complete rat Cδ gene both at the cDNA and genomic levels, showing that the rat Cδ gene is structurally similar to the corresponding mouse gene. Analysis of the rat immunoglobulin D heavy chain cDNA tail sequences, revealed two transcripts for the secreted form with varying sizes of their 3′ untranslated region (UTR), resulting from usage of two different poly(A) addition signals. Furthermore, a membrane-bound form encoding transcript, possessing a long 3′ UTR, was also observed. Phylogenetic analysis supports that the Cδ gene appeared early in the evolution of vertebrates, and it was probably duplicated from the Cμ gene more than 400 million years ago.

INTRODUCTION

Mammals express five classes of immunoglobulin heavy chain constant region genes, including μ, δ, γ, ϵ, and α. The δ gene, encoding immunoglobulin D (IgD), has been thought to be a relatively novel immunoglobulin class, being present only in rodents and primates. IgD was first identified in a myeloma patient and later, it was also found in mice and rats.[1] The molecule is mainly expressed on the surface of B cells and only low concentrations of the secreted form can be found in serum.[2,3] For the past three decades, there has been no convincing evidence to support the presence of the protein or the corresponding gene in additional species of mammals. In a recent study,[4] however, we have identified a δ gene in artiodactyl species such as cow, sheep and pig, showing that the gene in these species is structurally similar to the human δ gene, consisting of three constant region domains (CH1, CH2 and CH3) and one to two hinge encoding exons. Studies in these species have offered significant insight into IgD phylogeny, as it could be shown that the present δCH1 exon in ruminants was rather recently (approximately 20 million years ago) duplicated from the μCH1 together with its upstream sequence, leading to an approximately 96% sequence similarity between the μCH1 and δCH1 domains.[4] The duplication of the μCH1 upstream sequence, encompassing the Sμ region, has supplied the bovine δ gene with a short functional Sδ region. Thus, in contrast to primates and rodents, the bovine δ gene can be expressed through class switch recombination, involving the bovine Sμ and Sδ. The short Sδ has so far only been demonstrated in the cow.

Based on our findings, one may speculate that the Cδ gene may be present in most, if not all mammalian species. However, in other lower vertebrates, such as birds, reptiles and amphibians, the existence of the gene remains unknown.[5,6] Surprisingly, a δ like gene with multiple constant region domains was found several years ago downstream of the μ gene in teleost fish.[7–9] The gene thus appears to have evolved differently in vertebrates.

In order to study the evolution of the vertebrate Cδ gene, it is necessary to have access to additional sequences from different species to allow a comprehensive phylogenetic analysis. As only a partial sequence of the rat δ gene cDNA was previously described[10] we therefore analysed the complete rat δ gene at both the cDNA and genomic levels.

Received 6 September 2002; revised 4 November 2002; accepted 20 December 2002.

Abbreviations: CSR, class switch recombination; EST, expressed sequence tag; RACE, rapid amplification of cDNA ends; UTR, untranslated region.

The sequences reported in this paper have been deposited into NCBI GenBank with accession numbers AY148494-148497.

Correspondence: Dr L. Hammarström, Center for Biotechnology, Novum, Karolinska Institutet, SE-14157, Huddinge, Sweden. E-mail: lennart.hammarstrom@biosci.ki.se

MATERIALS AND METHODS

RNA isolations and reverse transcriptase–polymerase chain reaction (RT–PCR) amplification of the rat IgD constant region encoding cDNA

Total cellular RNA was extracted using the RNeasy® minikit (QIAGEN, Valencia, CA) from rat (Sprague–Dawley) spleen,

and subjected to cDNA synthesis with a First-Strand cDNA Synthesis Kit (Amersham Pharmacia Biotech, Uppsala, Sweden). RT–PCR was carried out under the following conditions: 94° 3 min, then 40 cycles of 94° 30 s, 55° 30 s and 72° 1 min. The primers used for amplification were: rat IgDs (5′-TAA GCT AAC CAG ACA CAG AAA G-3′) and rat IgDas (5′-CT CCT GAA AAG GAA CCT GAC T-3′).

Cloning of both the secreted and transmembrane tail of rat IgD heavy chain cDNA

3′ rapid amplification of cDNA ends (RACE) was carried out to clone both the secreted and transmembrane tail encoding sequences. About 5 μg rat spleen total RNA was reverse-transcribed into cDNA using a First-Strand cDNA Synthesis Kit (Amersham Pharmacia Biotech). The first amplification was performed using the primers rat IgD Hs (hinge exon derived, 5′-CCA ATC CAA CTC CCC AAG CAA A-3′) and cDN (designed based on the sequence of the *Not*I-d$(T)_{18}$ primer in the kit, 5′-AAC TGG AAG AAT TCG CGG CC-3′) under conditions of 94° 3 min for denaturing, 30 cycles of 94° 30 s, 58° 30 s, 72° 30 s. One μl PCR mixture from the first amplification was employed as template for the second amplification using the primers rat IgD CH3s (CH3 domain derived, 5′-ATC TGG AGT GTC CTG AGA CTG C-3′), and Fs-common (designed based on the sequence of the *Not*I-d$(T)_{18}$ primer in the kit, 5′-TGG AAG AAT TCG CGG CCG CAG GAA-3′) under conditions of 94° 3 min for denaturing, 30 cycles of 94° 30 s, 60° 30 seconds, 72° 30 s.

Long PCR amplification of genomic DNA fragments of rat immunoglubulin Cδ gene

To analyse exon–intron boundaries of the rat Cδ gene, genomic fragments spanning exon–intron–exon were amplified using a long PCR Kit (Expand™ Long Template PCR System Kit, Roche Diagnostics Scandinavia AB, Bromma, Sweden). The primer pairs used were, Rat IgD Hs (sequence as indicated above) and Rat IgD CH3as (5′-GTG AAT GTG GTC AGG ATG TT-3′) yielding an approximately 1·4 kb fragment, Rat IgD CH3s (sequence as indicated above) and Rat IgD Secas (5′-CTC AGG GAA AGG CAG GAC CAT C-3′) yielding a 4·8-kb band, Rat IgD Secs (5′-GAT GGT CCT GCC TTT CCC TGA G-3′) and Rat IgD TMas (5′-GAC ATG CCT CCT AAT CCT ATC A-3′) yielding a 2·5-kb band. The extending temperature for all amplifications was 68°, as annealing temperature for each amplification was adjusted according to the primers used.

Cloning and sequencing of PCR products

All PCR products were run on agarose gels and purified using a QIAquick Gel Extraction Kit (QIAGEN), and subsequently cloned into the pGEM-FT vector.[11] Sequencing of the recombinant plasmid was performed using an ABI PRISM® BigDye™ Terminator Ready Reaction Kit (Perkin-Elmer, Foster, CA). The exon–intron boundaries were deduced by comparing the genomic DNA with the cDNA sequences.

Computational analysis of DNA sequences and construction of phylogenetic tree

All DNA sequence editing, comparisons and alignments were performed using the MegAlign program (DNASTAR, Inc, Madison, WI). Some rat Cδ genomic DNA sequences were obtained by search of the rat whole genome shotgun database (http://www.ncbi.nlm.nih.gov/blast/mmtrace.html). Construction of phylogenetic tree was made using the *protpars* programs from the PHYLIP package.[12] The consensus tree were taken from 1000 bootstrapped phylogenetic trees (using IgM and IgD heavy chain constant region membrane bound tail peptide sequences). Except for the rat IgD sequence obtained in this study, all other sequences were extracted from the NCBI GenBank with the following accession numbers: cow IgM (U63637), human IgM (X14940), mouse IgM (J00443), pig IgM (U50149), sheep IgM (L04260), rabbit IgM (J00666), catfish IgM (M27230), salmon IgM (s48658), clawed frog IgM (M20484), duck IgM (U27213, AJ314754), human IgD (X57331), mouse IgD (J00447), cow IgD (AF411240), sheep IgD (AF411238), pig IgD (AF411239), catfish IgD (U67437), cod IgD (AF155203-AF155205), salmon IgD (AF141605). The membrane-bound form encoding cDNA tails of IgM heavy chain in species such as rat (BF551033), chicken (AJ395400, AJ444200, BG625537, AJ395139, AJ395400, AJ395578) and frog (BI314212), were identifieded from EST clones deposited in the NCBI GenBank.

RESULTS

Cloning of the rat full length IgD heavy chain constant region encoding cDNA

Part of the rat IgD constant region – encoding cDNA, lacking part of the CH1 sequence, has been published previously.[10] To clone the full rat IgD heavy chain constant region gene, a BLAST search against the rat EST database in the NCBI GenBank using the published δ sequence, was first carried out to look for EST clones, which may span the entire CH1 domain. One clone (AI555192) was found, containing a rat VH sequence and a putative δCH1 sequence. Two primers, rat IgDs and rat IgDas, were designed based on the sequence of the EST clone and 3′ UTR (untranslated region) of the published δ chain encoding sequence, and employed in RT–PCR amplification using total RNA isolated from rat spleen as an initial template. Sequencing of the cloned, approximately 800 bp PCR product (accession number: AY148494), revealed that the EST clone identified above indeed encoded the δ chain.

Alignment of the deduced rat IgD heavy chain constant region amino acid sequence with that of the mouse shows an overall 62·2% similarity. However, a comparison of each individual domain indicates that the CH3 domain is more conserved (76·6%) than the CH1 (44%) and hinge region (55·9%) domains. The potential disulphide bond donor, cysteine, appears at invariable positions throughout the entire constant region, if potential amino acid deletions or insertions are taken into account. In contrast, the tripeptide acceptor sequence for GlcN oligosaccharides attachment (N-X-T/S) is not positionally conserved (data not shown).

RACE amplifications of rat IgD heavy chain tail encoding sequences

Since the transmembrane tail of the rat IgD heavy chain was not known, the 3′ RACE technique was employed to clone both the

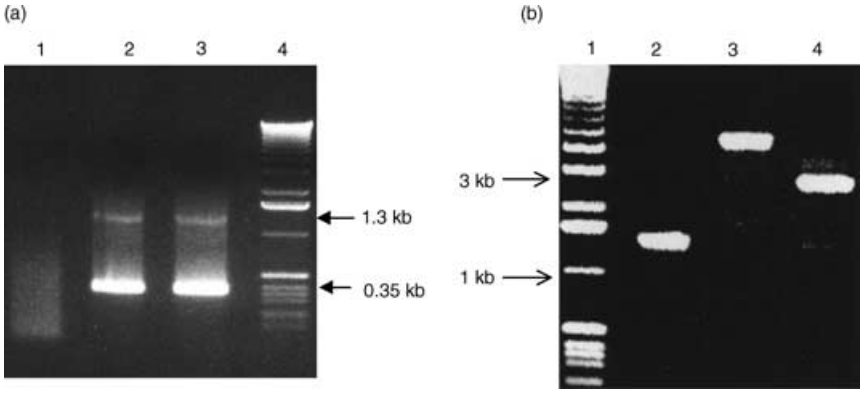

Figure 1. PCR amplifications of the rat Cδ gene fragments. (a) 3′ RACE PCR amplification of rat IgD heavy chain cDNA tails. 1. First PCR; 2, 3. Second PCR; 4. 1 kb DNA ladder. (b) Long PCR amplification of the rat genomic fragments. 1, 1 kb DNA ladder; 2, 3, 4, amplified genomic fragments b, c, d as indicated in Fig. 4.

transmembrane and secreted tails. The first RT–PCR amplification was performed using a hinge region derived primer, rat IgD Hs, and an anchor primer, cDN, using rat spleen total RNA as a template. The second amplification was subsequently carried out using 1 μl of the PCR mixture from the first amplification, employing the primers rat IgD CH3s (CH3 domain derived) and Fs-common. While the first amplification did not generate any visible bands, the second PCR produced two distinct bands, approximately 1·3 kb and 0·35 kb in size (Fig. 1a). Cloning and sequencing of the two bands showed that the long band represented the transcript for the membrane-anchored form, and the short band corresponded to the secreted form-encoding sequence. The transmembrane transcript has an approximately 1 kb long-3′ UTR sequence, containing $(GTTT)_8$ and $(AC)_{22}$ or $(CA)_{22}$ repeats representing two micro-(mini)-satellite sites (Fig. 2a), that are frequently involved in the regulation of gene expression. Furthermore, an examination of the 3′ UTR region revealed two potential poly(A) addition signal sites, indicated in Fig. 2(a).

Two different forms of mRNA transcripts for the secreted form were cloned, showing different lengths of the 3′ UTR and utilizations of two different poly(A) addition signal sites, located approximately 80 bp apart (Fig. 2b, c). However, the two transcripts encode the same peptide sequence of the secreted tail, and most likely, the shorter one is only rarely transcribed in rat spleen, as it is hardly visible on a gel separating the 3′ RACE products (Fig. 1a).

BLAST searches using both the rat secreted and membrane tails against the rat EST database revealed the presence of the same transcripts as we cloned (AW534436, BI291798). One EST (accession number: BI291798) clone contains an $(AC)_{20}$ repeat, indicating that the mini-satellite site within the 3′UTR is polymorphic in terms of the number of repeats in different animals.

Genomic organization of rat Cδ gene

Because a database for the rat whole genome shotgun sequence is available in the NCBI GenBank, it allows possibilities to map some, if not all, rat genes, by comparing the known cDNA sequence with genomic fragments deposited in the database. Employing this strategy, we found more than 10 genomic fragments containing the rat IgD CH1, Hinge and CH3 encoding sequences. One of these spans the CH1–intron–hinge region (gnl|ti|39491863 rta1lv58.y), allowing us to deduce the exon–intron boundaries and intron size. The latter was also confirmed by sequencing of cloned PCR products, whereas the hinge–CH3, CH3–Sec and Sec–TM genomic fragments were all generated by long PCR amplification using rat spleen DNA as a template and partially sequenced (Fig. 1b). We were thus able to construct a physical map of the gene (Fig. 3), where the intron sizes were determined based on the length of the PCR products. All the exon–intron boundaries were deduced based on comparisons of cDNA and genomic sequences (Table 1).

As this paper was being revised, a long genomic rough sequence, containing the rat δ gene (accession number: AC111740), was deposited in the NCBI GenBank by the Rat Genome Sequencing Consortium. However, the sequence has not yet been annotated. Analysis of this sequence revealed that the rat Cδ gene is located 2 kb downstream of the μTM2 exon, and the exon–intron boundaries are identical to the data deduced by sequencing of the PCR products. The size of introns (as indicated in parentheses in Table 1) are also consistent with the PCR results. A restriction map of the Cδ gene was constructed based on the sequence and is consistent with the pattern of digestion of the amplified PCR products (Fig. 3).

Phylogenetic aspects of Cδ genes in vertebrates

It is known that the transmembrane tails of IgM in vertebrates are evolutionarily conserved and typically encoded by two exons. As more δ sequences are becoming available, it appears that the δTM show a high degree of sequence similarity to the μTM, consistent with the hypothesis that the Cδ gene was evolutionarily duplicated from the Cμ gene. A thorough comparison of the μTM and δTM sequences may therefore be informative for understanding the evolution of the two genes. As seen in Fig. 4(a) (including the chicken and frog μTM sequences that have not previously published, and were identified from the NCBI GenBank EST database), the known IgM and IgD membrane-bound form in vertebrates possess a nearly invariant intracellular cytoplasmic tail (KVK), with

(a)

GGC ATG GTG GAC ACC ATC CCA AAC TCA TGC ATC AGG GAT GAG CAA ACC GAC AGC
G M V D T I P N S C I R D E Q T D S
TAC GTG GAC CTA GAG GAG GAG AAC GGC CTG TGG CCC ACA CTG TGC ACC TTC GTG
Y V D L E E E N G L W P T L C T F V
TM1 | TM2
GCC CTC TTT CTA CTC ACA CTG CTC TAC AGT GGC TTT GTC ACC TTC ATC AAG | GTG
A L F L L T L L Y S G F V T F I K V
AAG TAG ATC AGG ACA ACA GAA TCC TGC AAC TAC AAA GAA AAG TGC CTT CCC TCA
K *
ACATGAAGCCCACAGAGAGAATGCCTTCTATGCAGAGAGAAGCTCCAAGCCCTCCTCTCAGTCCTCGCTCT
CCACCCAGCCTCCCAGACTCAAAATGCCTGTCATCCTCGGCTATAGGCAAAGAAGGTGTCCTGCTGACCCA
GTCCTACTCATCCCTGTTCCCTGTCTCCCTCCAAGTCCCCATCATTCTATAAGCTTGCTGTGGTTCACCCC
ACCTGCCTCCT**GTTTGTTTGTTTGTTTGTTTGTTTGTTTGTTT**CTGTGGTTTA**AATAAA**CTCAAAGTGCAT
GCAATCTGAGGCAAACAGTGGAGAGGCCACTAGGGAGGAGTATCAGGGTCACACTGAAGAAAACTGTTCAG
TGAACTAGGACACTCTCCTCCCACCATAAACTCACAGTTGACTTCTCACATACACATAAGAATA**CACACAC**
ACACACACACACACACACACACACACACACACACACACATGCATGTGTTCACATCACAGTCCCTCAAACAAAA
TGAACATACATGCATGATTACAGAAGCAGATGTACATGAGTGCATATGCATACCTACTCAGATGAACACAA
ATTTACACCAGGGCCTTAGGTGACAGCAATGGGGACTGTGGTCATTTCAATAAGAATAGACTCCATGGGTT
CATATAGTGGAATGCTTAGTCACCAGGGAGTGGAACTCTTTGATAGGATTAGGAGGCATGTCCTTCTAAGA
AATGTGTGTCACTGGAGCACAGGAAGGCTTTCAGGTTTCAAAAGCTCATGTCAGGACCAATGTCTGTTTGT
CTCTCTGTGTCTCTGTGTTATCTCTTTAGAACTCTAGGTTATTTCTCCAGCACCATGTCCACTGGCATACC
ACCATGCTCCCCAACATGATAATGAACTAACCTCTGAAATTGTTAGCAAGCCCC**AATTAAA**TGCTTTCTTT
TATAAGAAAAAAAAAAAAAAAAAAAAAAA

(b)

GGA TGC TAC CAC CTC CTG CCT GAG TCC GAC GGC CCT CCC AGG AGA CCT GAT GGT
G C Y H L L P E S D G P P R R P D G
CCT GCC TTT CCC TGA GAC CTC TCT GTG TCT GCT TAG CTT ATT TGC CTC TGT GTT
P A F P *
AAGTTGCTTC**ATAAAA**GCAAAAACCAAAAAAAAAAAAAAAAAAA

(c)

GGA TGC TAC CAC CTC CTG CCT GAG TCC GAC GGC CCT CCC AGG AGA CCT GAT GGT
G C Y H L L P E S D G P P R R P D G
CCT GCC TTT CCC TGA GAC CTT TCT GTG TCT GCT TAG CTT ATT TGC CTC TGT GTT
P A F P *
AAGTTGCTTC**ATAAAA**GCAAAAACCATTGCCTTTCCACTTCTCATTATGTTGTAGAAAAGCACTTCTGTAA
ATTACAGTAAGAAGTGAACAAT**AATAAA**TGAGACAAACAAAAAAAAAAAAAAAA

Figure 2. Nucleotide sequences of both rat IgD secreted and transmembrane encoding cDNA tails. The coding sequences, repeats and poly(A) addition signals are in bold. The repeats and poly(A) addition signal are underlined. The stop codons are indicated as stars. The sequences have been deposited into NCBI GenBank with accession numbers AY148495–148497.

the exception of catfish δTM (KVKIA), salmon δTM (KTK) and cod δTM (KTK). In salmon and cod, the neutral amino acid, valine, is substituted by threonine[8,9] and the catfish IgD molecule contains two extra amino acids in the cytoplasmic tail.[7]

A distinct usage of stop codons in δ and μ heavy chain transcripts was also revealed when the known mammalian IgM and IgD heavy chain transcripts were compared. As shown in Table 2, the opal codon UGA is utilized in all the known secreted and membrane-bound form encoding transcripts of IgM, while the amber codon UAG appears to be used in the IgD membrane-bound encoding transcripts. None of these transcripts in mammals adopts the ochre codon UAA as a termination codon, although it is preferentially used in bird IgM transcripts. In contrast to mammals, all three teleost fish species investigated to date frequently utilize the opal codon UGA as a stop codon in IgD membrane-bound form encoding transcripts instead of the amber codon that is utilized by secreted forms of μ

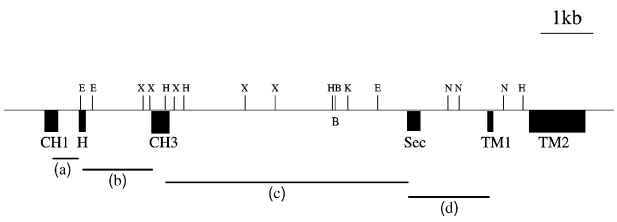

Figure 3. Genomic organization of the rat Cδ gene. (a) A genomic fragment obtained from the rat whole genome shotgun sequence (gnl|til39491863 rta11v58.y), and confirmed by sequencing of a cloned PCR product; (b, c, d) Cloned PCR products. B *Bam*HI; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; N, *Nco*I; X, *Xba*I.

Table 1. Exon–intron boundaries of the rat immunoglobulin δ gene

	Ex/In	Intron size (kb)	In/Ex
CH1-In-H	AGTTGCCTG/**gt**gagtatc	0·373 (0·371)	tgactcc**ag**/AGACAAGGA
H-In-CH3	ACATAGTAG/**gt**aagaagt	1·4 (1·353)	tccctgc**ag**/GGGCTATGG
CH3-In-Sec	AAATCAGTG/**gt**aagtcac	4·6 (4·681)	tgattct**ag**/GATGCTACC
CH3-In-TM1	AAATCAGTG/**gt**aagtcac	5·9 (6·190)	cttcccc**ag**/GCATGGTGG
TM1-In-TM2	TTCATCAAG/**gt**gggcctc	0·301 (0·301)	ctctggc**ag**/GTGAAGTAG

The intron size indicated in parentheses is based on the sequence derived from the NCBI GenBank (accession number: AC111740).

encoding transcripts. The differential usage of these termination codons in IgM and IgD transcripts may reflect an evolutionary difference and that the expression of IgM and IgD may be differentially regulated in vertebrates.

As shown in Fig. 4(a), both the δTM and μTM sequence in vertebrates are well conserved from the lower to the higher vertebrates. We have constructed a phylogenetic tree using the peptide sequences presented in Fig. 4(b). The unrooted phylo-

(a)

```
MouseMTM     --------EGEVNAEE--------EGFENLWTTASTFIVLFLLSLFYSTTVTLFKVK--
RatMTM       --------EGEVNAEE--------EGFENLWTTASTFIVLFLLSLFYSTTVTLFKVK--
RabbitMTM    --------EGEVGAEE--------EGFENLWTTASTFIVLFLLSLFYSTTVTLFKVK--
CowMTM       --------EGEVSAEE--------EGFENLNTMASTFIVLFLLSLFYSTTVTLFKVK--
PigMTM       --------EGEVSAEE--------EGFENLNTMASTFIVLFLLSLFYSTTVTLFKVK--
SheepMTM     --------EGEVSAEE--------EGFENLNTMASTFIVLFLLSLFYSTTVTLFKVK--
HumanMTM     --------EGEVSADE--------EGFENLWATASTFIVLFLLSLFYSTTVTLFKVK--
FrogMTM      --------AKVLFVDEG-------EEMASLWTTASTFIVLFLLSLFYSATVTLFKVK--
ChickenMTM   ------LLHWPLEAEED-------DDIQRLWATTSTFIVLFILSLFYSATVTLIKVK--
DuckMTM      --------DAYTEVEED-------DDLDNLWTTASTFIVLFILSLFYSATVTLIKVK--
CatfishMTM   ----VCIWSTEI-FHYE-----MEMDDDNMANTALTFVFLFLITLFYSIGVTVFKVK--
SalmonMTM    -----CLVLTDCPCSNT-----IETDRDSMGKTAFTFIILFLITLLYGVGATAIKVK--
CowDTM       --GLATWPPWS--QDESSDDGTDVEDASPLWLT---FLALFLVTVVYGGFVTFIKVK--
SheepDTM     --GLATWPPWS--QDESSDDSADAEDASPLWLT---FLALFLATVVYSGFVTFIKVK--
PigDTM       --GLATLTPGS--QDEGSDDYVDLEDAGRLWLT---FTVLFLVTLLYSGFVTFLKVK--
HumanDTM     YLAMTPLIPQS--KDENSDDYTTFDDVGSLWTTLSTFVALFILTLLYSGIVTFIKVK--
RatDTM       --GMVDTIPNSCIRDEQTDSYVDLEEENGLWPTLCTFVALFLLTLLYSGFVTFIKVK--
MouseDTM     ----VNTIQHSCIMDEQSDSYMDLEEENGLWPTMCTFVALFLLTLLYSGFVTFIKVK—
CatfishDTM   -----EFTEETIYFDENTSETSTLDQAEETWNMACAFIILFVISLLYGCSVTLVKVKIA
SalmonDTM    ----VLLGVEGVFFDENRSDPILADTAEENWNMACIFLVLFLISLLYSITVTLVKTK--
CodDTM       ----PEILEESVLFDENK-QVFVQDMGEDNWMMTITFLLLFFVSLFYGILVTVIKTK--
```

(b)

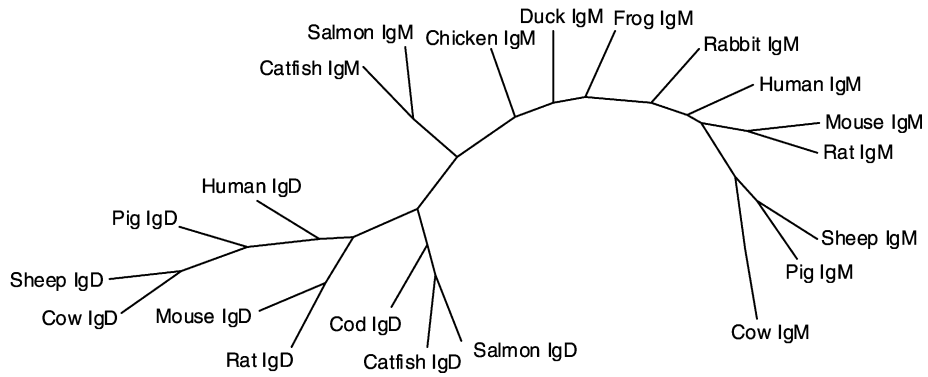

Figure 4. Phylogeny of the δTM and μTM in vertebrates, (a) Sequence comparison of δTM and μTM from vertebrates. Dashes indicate potential deletions. The rat δTM sequence that was obtained from this study, and the rat μTM (BF551033), frog μTM (BI314212) and chicken μTM (AJ444200, BG625537, AJ395139, AJ395400 and AJ395578) were identified by us from the NCBI GenBank EST database. All other sequences were based on the following references: mouse μTM,[24] rabbit μTM,[25] cow μTM,[26] pig μTM,[27] sheep μTM,[28] human μTM, δTM,[29] duck μTM,[6] catfish μTM,[30] salmon μTM,[31] cow δTM, sheep δTM, pig δTM,[4] mouse δTM,[32] catfish δTM,[7] salmon δTM[8] and cod δTM.[9] (b) An unrooted phylogenetic tree constructed using both the IgD and IgM transmembrane amino acid sequences. The tree was constructed using *propars* program, and taken from 1000 trees by running *consenus* from the PHYLIP package.[12]

Table 2. Stop codon usage of the known IgD and IgM heavy chain transcripts in vertebrates

Species	δTM	δSec	μTM	μSec
Cow	TAG		TGA	TGA
Sheep	TAG		TGA	TGA
Pig	TAG	TAG	TGA	TGA
Human	TAG	TGA	TGA	TGA
Mouse	TAG	TGA	TGA	TGA
Rat	TAG	TGA	TGA	TGA
Rabbit			TGA	TGA
Hamster				TGA
Horse				TGA
Cat				TGA
Chicken			TGA	TAA
Duck			TAA	TAA
frog			TAA	TAG
Cod	TGA		TGA	TAG
Salmon	TGA		TGA	TAG
Catfish	TGA	TAG	TGA	TAG

The rat μTM (BF551033), chicken μTM (AJ444200, BG625537, AJ395139, AJ395400 and AJ395578) and frog μTM (BI314212) were identified by us from the NCBI EST database, and the rat δTM, δSec was derived from this study. All other sequence data in this table are based on the following references: cow δTM,[4] cow μTM, μSec,[26] sheep δTM,[4] sheep μTM,[28] sheep μSec,[33] pig δTM, δSec,[4] pig μTM, μSec,[27] human δTM, δSec, μTM, μSec,[29] mouse δTM, δSec,[32] mouse μTM, μSec,[24] rat μSec,[34] rabbit μTM, μSec,[25] hamster μSec,[35] horse μSec,[36] cat μSec,[37] chicken μSec,[38] duck μTM, μSec,[6] frog μSec,[39] cod δTM,[9] cod μTM, μSec,[40] salmon δTM,[8] salmon μTM, μSec,[31] catfish δTM, δSec,[7] catfish μTM, μSec.[30]

genetic tree displayed a distinct IgD clade from the IgM, indicating that the vertebrate Cδ gene was duplicated from the Cμ gene early in evolution, probably before divergence of the teleost fish from the lineage leading to mammals, approximately 400 million years ago.[13,14]

DISCUSSION

Rodent IgD is distinguished by a short hinge region and a lack of the δCH2 domain[3,10,15,16] but the biological consequence of the short IgD heavy chain remains unclear. The absence of δCH2 domain and in rats and mice is probably caused by a genetic event that deleted a pre-existing genomic sequence containing δCH2 and one of the hinge encoding exons during evolution. However, a sequence comparison of human and mouse Cδ loci did not generate evidence to support this view[17] because of the low sequence similarity between them. An examination of the Cδ gene from species such as hamster and guinea-pig, which are phylogenetically closer to rats and mice, may provide useful information for understanding the Cδ gene evolution in rodents.

As more and more Cδ gene sequences are being identified in a variety of vertebrate species, several interesting points regarding their phylogeny, expression patterns and structural diversities have emerged. In contrast to the previously held view, IgD has a wide distribution in vertebrate species, and it is probably present in most, if not all mammals, implicating that IgD is a functionally important antibody class. Our previous phylogenetic analysis indicated that teleost fish δ genes were true homologues to the mammalian genes[4] suggesting that IgD may also be present in birds, amphibians and reptiles, which are evolutionarily located between teleost fish and mammals. However, a δ gene does not seem to be present in birds according to investigations on the bird IGHC gene loci.[5,6] In addition, a survey of an EST database representing the African clawed frog expressed genes (containing more than 220 000 clones) showed that there were no putative antibody classes other than the three known IgM, IgY, IgX (unpublished data). Furthermore, no translated sequence displayed a higher similarity to the conserved IgD transmembrane domain sequence, than the defined frog IgM transmembrane sequence. However, before a conclusion that the IgD is absent in these species is drawn, there are still some points that need to be taken into account. One major key factor is that the topology of the bird IGHC loci are distinct from the mammalian loci, as the Cα gene is situated in an inverted transcriptional orientation between the Cμ and the Cυ genes.[5,6] The unusual structure of the bird IGHC locus indicates that it may have undergone gene rearrangements during evolution, which may have deleted the Cδ gene or translocated it to downstream of the Cυ gene.

The second important point concerning IgD in vertebrates is their structural diversities. In teleost fish, up to 10 domain-encoding Cδ exons, resulting from the multiple duplications, have been observed.[7,8] Exon deletions have also played a role in the formation of the Cδ gene in Atlantic cod.[9] Furthermore, in teleost fish, the hinge region has not yet been developed. In mammals, the lack of the δCH2 domain makes rodent IgD shorter than other known mammlian IgD molecules. Rodents also display a short, one-exon encoded, hinge region, while the hinge regions are encoded by two exons in most other mammalian species. In the pig, however, we recently found a one-exon encoded IgD hinge region, although it only shows a low degree of similarity to the rodent IgD hinge,[4] suggesting that they arose by independent genetic events.

In all known cases, the Cδ genes are located immediately downstream of the Cμ gene. However, consistent with their structural diversities, the expression of the Cδ genes display several distinct pathways, which can basically be classified into two major categories; cotranscription with the Cμ gene and class switch recombination (CSR). The former mechanism, which is performed through alternative splicing of a long Cμ and Cδ containing transcript, is more frequently employed in the expression of mammalian Cδ genes.[18] In teleost fish, however, the μCH1 is spliced onto the δ sequence to form a chimeric antibody heavy chain,[7–9] a mechanism that has never been observed in primates or rodents. CSR is a mechanism involving in expression of other antibody classes than IgM and IgD in mammals[19] and only in very rare cases, direct switching to IgD has been observed in humans and mice.[20–22] The reason is the lack of an authentic switch delta region in these species.[17] However, we have recently demonstrated in the cow, that there is a short Sδ region located in the Cμ and Cδ intronic region, which can direct switching.[4]

Our recent findings support the notion that IgD is an ancient antibody class, present in most mammals, indicating that there are probably other unidentified biologically important functions for IgD, in addition to its ability to replace the function of IgM in driving development of B cells as previously demonstrated in μ gene knockout mice.[23] More work is however, required to generate a complete blueprint of the role of IgD during B-cell

development in order to understand the functional importance of this antibody class.

ACKNOWLEDGMENTS

This work was supported by the Swedish Research Council. The authors want to thank Janos Geli for technical assistance.

REFERENCES

1 Abney ER, Parkhouse RM. Candidate for immunoglobulin D present on murine B lymphocytes. Nature 1974; **252:**600–2.
2 Pernis B, Chiappino G, Rowe DS. Cells producing IgD immunoglobulins in human spleen. Nature 1966; **211:**424–5.
3 Takahashi N, Tetaert D, Debuire B, Lin LC, Putnam FW. Complete amino acid sequence of the delta heavy chain of human immunoglobulin D. Proc Natl Acad Sci U S A 1982; **79:**2850–4.
4 Zhao Y, Kacskovics I, Pan Q, Liberles DA, Geli J, Davis SK, Rabbani H, Hammarstrom L. Artiodactyl IgD: the missing link. J Immunol 2002; **169:**4408–16.
5 Zhao Y, Rabbani H, Shimizu A, Hammarstrom L. Mapping of the chicken immunoglobulin heavy-chain constant region gene locus reveals an inverted alpha gene upstream of a condensed upsilon gene. Immunology 2000; **101:**348–53.
6 Lundqvist ML, Middleton DL, Hazard S, Warr GW. The immunoglobulin heavy chain locus of the duck. Genomic organization and expression of D, J, and C region genes. J Biol Chem 2001; **276:**46729–36.
7 Wilson M, Bengten E, Miller NW, Clem LW, Du Pasquier L, Warr GW. A novel chimeric Ig heavy chain from a teleost fish shares similarities to IgD. Proc Natl Acad Sci U S A 1997; **94:**4593–7.
8 Hordvik I, Thevarajan J, Samdal I, Bastani N, Krossoy B. Molecular cloning and phylogenetic analysis of the Atlantic salmon immunoglobulin D gene. Scand J Immunol 1999; **50:**202–10.
9 Stenvik J, Jorgensen TO. Immunoglobulin D (IgD) of Atlantic cod has a unique structure. Immunogenetics 2000; **51:**452–61.
10 Sire J, Auffray C, Jordan BR. Rat immunoglobulin delta heavy chain gene: nucleotide sequence derived from cloned cDNA. Gene 1982; **20:**377–86.
11 Zhao Y, Liu C, Wen S, Dai S, Rabbani H, Li N, Hammarstrom L. Construction of a high efficiency PCR products cloning T vector using pGEM-5zf (+). Prog Natural Sci (Chinese); in press.
12 Felsenstein J 1989 Phylip–Phylogeny Interface Package Version 3.2. Cladistics 2002; **5:**164–6.
13 Kumar S, Hedges SB. A molecular timescale for vertebrate evolution. Nature 1998; **392:**917–20.
14 Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. Annu Rev Immunol 1999; **17:**109–47.
15 Liu CP, Tucker PW, Mushinski JF, Blattner FR. Mapping of heavy chain genes for mouse immunoglobulins M and D. Science 1980; **209:**1348–53.
16 Tucker PW, Liu CP, Mushinski JF, Blattner FR. Mouse immunoglobulin D. messenger RNA and genomic DNA sequences. Science 1980; **209:**1353–60.
17 Koop BF, Richards JE, Durfee TD *et al.* Analysis and comparison of the mouse and human immunoglobulin heavy chain JH-Cmu-Cdelta locus. Mol Phylogenet Evol 1996; **5:**33–49.
18 Maki R, Roeder W, Traunecker A, Sidman C, Wabl M, Raschke W, Tonegawa S. The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta genes. Cell 1981; **24:**353–65.
19 Harriman W, Volk H, Defranoux N, Wabl M. Immunoglobulin class switch recombination. Annu Rev Immunol 1993; **11:**361–84.
20 Owens JD Jr, Finkelman FD, Mountz JD, Mushinski JF. Nonhomologous recombination at sites within the mouse JH-C delta locus accompanies C mu deletion and switch to immunoglobulin D secretion. Mol Cell Biol 1991; **11:**5660–70.
21 White MB, Word CJ, Humphries CG, Blattner FR, Tucker PW. Immunoglobulin D switching can occur through homologous recombination in human B cells. Mol Cell Biol 1990; **10:**3690–9.
22 Kluin PM, Kayano H, Zani VJ, Kluin-Nelemans HC, Tucker PW, Satterwhite E, Dyer MJ. IgD class switching: identification of a novel recombination site in neoplastic and normal B cells. Eur J Immunol 1995; **25:**3504–8.
23 Lutz C, Ledermann B, Kosco-Vilbois MH, Ochsenbein AF, Zinkernagel RM, Kohler G, Brombacher F. IgD can largely substitute for loss of IgM function in B cells. Nature 1998; **393:**797–801.
24 Richards JE, Gilliam AC, Shen A, Tucker PW, Blattner FR. Unusual sequences in the murine immunoglobulin mu-delta heavy-chain region. Nature 1983; **306:**483–7.
25 Bernstein KE, Alexander CB, Reddy EP, Mage RG. Complete sequence of a cloned cDNA encoding rabbit secreted mu-chain of VHa2 allotype: comparisons with VHa1 and membrane mu sequences. J Immunol 1984; **132:**490–5.
26 Mousavi M, Rabbani H, Pilstrom L, Hammarstrom L. Characterization of the gene for the membrane and secretory form of the IgM heavy-chain constant region gene (C mu) of the cow (*Bos taurus*). Immunology 1998; **93:**581–8.
27 Sun J, Butler JE. Sequence analysis of pig switch mu, C mu, and C mu m. Immunogenetics 1997; **46:**452–60.
28 Hein WR, Dudler L. Nucleotide sequence of the membrane form of sheep IgM and identification of two C mu allotypes. Mol Immunol 1993; **30:**783–4.
29 Word CJ, White MB, Kuziel WA, Shen AL, Blattner FR, Tucker PW. The human immunoglobulin C mu-C delta locus: complete nucleotide sequence and structural analysis. Int Immunol 1989; **1:**296–309.
30 Wilson MR, Marcuz A, van Ginkel F, Miller NW, Clem LW, Middleton D, Warr GW. The immunoglobulin M heavy chain constant region gene of the channel catfish, *Ictalurus punctatus*: an unusual mRNA splice pattern produces the membrane form of the molecule. Nucl Acids Res 1990; **18:**5227–33.
31 Hordvik I, De Vries Lindstrom C, Voie AM, Lilybert A, Jacob J, Endresen C. Structure and organization of the immunoglobulin M heavy chain genes in Atlantic salmon, *Salmo salar*. Mol Immunol 1997; **34:**631–9.
32 Cheng HL, Blattner FR, Fitzmaurice L, Mushinski JF, Tucker PW. Structure of genes for membrane and secreted murine IgD heavy chains. Nature 1982; **296:**410–5.
33 Patri S, Nau F. Isolation and sequence of a cDNA coding for the immunoglobulin mu chain of the sheep. Mol Immunol 1992; **29:**829–36.
34 Parker KE, Bugeon L, Cuturi MC, Soulillou JP. Cloning of cDNA coding for the rat mu heavy chain constant region: differences between rat allotypes. Immunogenetics 1994; **39:**159.
35 McGuire KL, Duncan WR, Tucker PW. Phylogenetic conservation of immunoglobulin heavy chains: direct comparison of hamster and mouse Cmu genes. Nucl Acids Res 1985; **13:**5611–28.
36 Schrenzel MD, King DP, McKnight ML, Ferrick DA. Characterization of horse (*Equus caballus*) immunoglobulin mu chain- encoding genes. Immunogenetics 1997; **45:**386–93.
37 Kanai TH, Ueda S, Nakamura T. Identification of two allelic IgG1 C (H) coding regions (Cgamma1) of cat. Vet Immunol Immunopathol 2000; **73:**53–62.
38 Dahan A, Reynaud CA, Weill JC. Nucleotide sequence of the constant region of a chicken mu heavy chain immunoglobulin mRNA. Nucl Acids Res 1983; **11:**5381–9.

39 Schwager J, Mikoryak CA, Steiner LA. Amino acid sequence of heavy chain from *Xenopus laevis* IgM deduced from cDNA sequence: implications for evolution of immunoglobulin domains. Proc Natl Acad Sci USA 1988; **85:**2245–9.

40 Bengten E, Leanderson T, Pilstrom L. Immunoglobulin heavy chain cDNA from the teleost Atlantic cod (*Gadus morhua* L.): nucleotide sequences of secretory and membrane form show an unusual splicing pattern. Eur J Immunol 1991; **21:**3027–33.