Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 262–266, January 1994
Immunology

Isolation and characterization of major histocompatibility complex class IIB genes from the nurse shark

(polymerase chain reaction/polymorphism/molecular evolution/cartilaginous fish)

SIMONA BARTL* AND IRVING L. WEISSMAN

Hopkins Marine Station, Stanford University, Pacific Grove, CA 93950; and Department of Pathology, Stanford University School of Medicine, Stanford, CA 94305

Contributed by Irving L. Weissman, September 15, 1993

ABSTRACT **The major histocompatibility complex (MHC) contains a set of linked genes which encode cell surface proteins involved in the binding of small peptide antigens for their subsequent recognition by T lymphocytes. MHC proteins share structural features and the presence and location of polymorphic residues which play a role in the binding of antigens. In order to compare the structure of these molecules and gain insights into their evolution, we have isolated two MHC class IIB genes from the nurse shark, *Ginglymostoma cirratum*. Two clones, most probably alleles, encode proteins which differ by 13 amino acids located in the putative antigen-binding cleft. The protein structure and the location of polymorphic residues are similar to their mammalian counterparts. Although these genes appear to encode a typical MHC protein, no T-cell-mediated responses have been demonstrated in cartilaginous fish. The nurse shark represents the most phylogenetically primitive organism in which both class IIA [Kasahara, M., Vazquez, M., Sato, K., McKinney, E. C. & Flajnik, M. F. (1992) *Proc. Natl. Acad. Sci USA* 89, 6688–6692] and class IIB genes, presumably encoding the α/β heterodimer, have been isolated.**

Mammalian immune systems have the ability to recognize soluble antigens as well as antigens on cell surfaces. B cells bind soluble molecules which are perceived as foreign. T cells respond to small antigenic peptides as part of a complex on cells. This complex consists of an antigenic peptide bound in a groove of a membrane protein encoded in the major histocompatibility complex (MHC). The requirement for T cells but not B cells to bind antigen as a cell surface complex allows for differential targeting of these two components of the immune system.

The MHC-encoded proteins which are bound by T cells can be placed into two classes, I and II (1). Both classes of proteins are involved in the recognition of antigenic peptides by T cells. Class I proteins generally bind endogenously derived peptides from intracellular proteins encoded in the cellular genome or by infecting intracellular agents—e.g., viral genomes. Class II proteins, in general, bind exogenous peptides derived from soluble proteins and extracellular pathogens (2).

Both classes of MHC proteins have a similar tertiary structure but differ at the primary and secondary structure level (3, 4). Class I molecules have a three-domain structure with a fourth domain contributed by a noncovalently bound molecule called $\beta 2$ microglobulin. The class II molecule is a heterodimer with each protein, α and β, contributing two extracellular domains. For both classes of MHC proteins, the two membrane-proximal domains form a characteristic fold which identifies these molecules as immunoglobulin superfamily members. The two membrane-distal domains form a cleft which binds peptides. The sides of the cleft are composed of two α-helices which lie on a β-pleated sheet. The uppermost surface of the cleft with its associated peptide is recognized by the T-cell receptor (5).

Within mammalian populations, any single species may contain many loci encoding each type of MHC protein (class I, class IIα, or class IIβ) and up to 100 alleles for each locus (6). MHC alleles exhibit a high rate of nonsynonymous substitution within the domain encoding the peptide-binding cleft (7, 8). The resultant amino acid differences usually point into the cleft or along the upper surface of the α-helices, potentially affecting both peptide binding and T-cell recognition (3). It is the recognition of allelic MHC products with bound peptides that causes vigorous rejection of intraspecific tissue grafts between MHC-disparate individuals.

To understand the evolution of the MHC protein-based immune recognition system, it is important to isolate and characterize contemporary representatives of genes or subgenic elements encoding MHC molecules in more primitive species. Through an analysis of their structure, we may gain insights into the nature of ancestral precursors of both mammalian and primitive chordate MHC genes. Did such precursors bind peptides, perhaps as some type of transporter molecules? Were they highly polymorphic? Did interacting cells express receptors that recognized these precursors? When did the two MHC classes diverge, and which class is ancestral?

Within the last three years MHC gene homologues have been found in several species of bony fish (9–14) and cartilaginous fish (15, 16). Consistent with the presence of MHC homologues, bony fish display many hallmarks of T-cell-dependent immune responses, such as acute allograft rejection (17), mixed leukocyte reactions (18), and antibody production resulting from T/B-cell collaboration (19). In contrast, none of these T-cell-dependent responses have been documented in cartilaginous fish (20–22), yet genes which encode what appear to be typical MHC proteins are present in these organisms.

In this paper, we describe the isolation of two cDNA clones encoding MHC class IIB alleles from the nurse shark.† The sequences of these genes and the polypeptides they might encode are similar to mammalian class IIB genes in deduced amino acid sequence, proposed tertiary structure, and the extent and distribution of polymorphism. These data, along with previously published data on the class IIA molecules (16), indicate that both class IIA and class IIB genes, and

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

Abbreviations: MHC, major histocompatibility complex; ORF, open reading frame.

*Present address: Department of Biological Sciences, University of North Carolina, 601 South College Road, Wilmington, NC 28403-3297.

†The sequences reported in this paper have been deposited in the GenBank data base [accession nos. L20274 (clone 11) and L20275 (clone 8)].

presumably the α/β heterodimer they encode, exist in cartilaginous fish.

MATERIALS AND METHODS

Isolation of MHC-Related Gene Segments by PCR Amplification. We decided to probe the nurse shark genome for MHC-related genes by amplifying segments between two conserved regions within known MHC class I and II genes (see Fig. 1). Primer 1 is 5′-CCAAGCTTTG(TC)IIIGTI(AT)-(CAG)IGGITT(TC)TA(TC)CC-3′ and primer 2 is 5′-CCGG-ATCCA(GA)I(GC)(AT)I(GC)(AT)(AG)TGI(TA)(CT)I-ACI(TC)(TG)(AG)CA-3′. Nucleotides in parentheses are degenerate and I is deoxyinosine. Restriction sites were added at the 5′ ends of each primer.

Peripheral blood lymphocyte RNA from the nurse shark, *Ginglymostoma cirratum* (subclass Elasmobranchii), was kindly provided by Martin F. Flajnik (University of Miami School of Medicine). Total RNA (3.5 μg) was converted to cDNA and 1/10th of the resultant product was used as PCR template in a 50-μl reaction mixture containing 100 pmol of each primer, 0.2 mM each dNTP, 2.5 units of AmpliTaq DNA polymerase (Perkin–Elmer/Cetus), 10 mM Tris·HCl (pH 8.3), 50 mM KCl, and 1.5 mM $MgCl_2$ overlaid with mineral oil. The first five PCR cycles were 94°C for 1 min, 40°C for 2 min, a slow rise over 2 min to 72°C for 3 min. The next 35 cycles were 94°C for 1 min, 45°C for 2 min, and 72°C for 3 min, with a final extension step of 72°C for 15 min.

PCR products were cloned into pBluescript II SK(+) (Stratagene) and sequenced as single-stranded DNA with the Sequenase kit (United States Biochemical). Four of eight clones sequenced exhibited sequence similarity to known MHC genes. One of these clones (named 60C) was used to probe a library.

Library Screening and Southern Blots. A nurse shark spleen cDNA library was generously provided by Martin F. Flajnik. The insert from clone 60C was labeled with ^{32}P by T4 polynucleotide kinase (United States Biochemical). The library was screened by using Quikhyb (Stratagene) with a final wash at 60°C with 0.2× standard saline citrate (SSC)/0.1% SDS.

Whole blood from two captive adult nurse sharks was kindly provided by Sylvia Smith (Florida International University). Blood was shipped at room temperature in lysis buffer and genomic DNA was extracted according to Seutin *et al.* (23). Southern hybridization was carried out on supported nitrocellulose membranes (Schleicher & Schuell) with randomly primed probe (Prime-it, Stratagene). Filters were hybridized with Quikhyb (Stratagene) at 68°C, with the highest stringency wash being 60°C with 2× SSC/0.1% SDS.

Alignments and Phylogenetic Trees. The inferred protein sequences of the *Ginglymostoma* cDNA clones were aligned with other MHC sequences retrieved from the Swiss-Prot 23 data base by the CLUSTAL V computer program (24). Minor adjustments to the alignment were done manually. The aligned amino acid sequences were used for distance and parsimony analyses. Two data sets were used for distance analyses. First, the membrane-proximal domains of class I, class IIα, and class IIβ were analyzed with β_2-microglobulin as the outgroup. Second, 15 complete class IIβ sequences were analyzed with 4 class I sequences (without the α1 domain) as the outgroup. Pairwise distances were calculated according to the formula of Kimura (25). Evolutionary relationships were determined by using neighbor joining (26), and reliability of the branching order was determined by 2000 bootstrap replications. For maximum parsimony analyses, 8 complete class IIβ sequences were analyzed with 4 class I sequences as the outgroup by using the PAUP 3.0s program to perform heuristic (with bootstrapping) and branch and bound searches (27).

RESULTS

Isolation of Nurse Shark MHC Genes. We designed primers to two regions each encoding a conserved cysteine found within the membrane-proximal domains of all MHC class I and II proteins. However, there is some variation in the amino acids encoded within these regions and we accounted for this by using highly degenerate, inosine-containing PCR primers. At the position corresponding to the codon downstream of the first cysteine codon, we choose to use three inosines to allow for complete variation in the template sequence. Primers 1 and 2 are 31-mers with six and five inosines, 48- and 2048-fold degeneracy, and 196,608 and 2,097,152 potential target sequences, respectively (see *Materials and Methods*).

PCR amplification of cDNA from nurse shark peripheral blood lymphocytes yielded a band at ≈200 bp. Eight PCR products were cloned and sequenced, four of which exhibited sequence similarity to MHC genes. Clones 60A and 60C differed by only 1 nucleotide and clones 61B and 61C were identical. Clones 60A and 60C encoded the amino acid stretch NGDWT which is found in many class IIB genes. A search of the Swiss-Prot 20 data base using the deduced amino acid sequence found that the 16 most similar sequences were MHC class IIβ proteins. A similar analysis of sub-clone 61B found that the two most similar sequences were β_2-microglobulin while the rest of the sequences were not MHC proteins. We chose to screen a nurse shark spleen cDNA library with clone 60C.

Seventeen clones isolated from the library could be divided into two groups based on sequence similarity. Two of the longest clones of each group, called 8 and 11, were sequenced in their entirety in both directions and the DNA and deduced amino acid sequences are shown in Fig. 1. The two clones differ by 21 nucleotides in the coding region which result in 13 amino acid changes. This degree of similarity is on the order of allelic differences between mammalian MHC genes; therefore we propose that these clones will turn out to be allelic (6). Like allelic forms of MHC genes, the majority of amino acid differences between these clones are encoded to the membrane-distal domain, which contains the putative antigen-binding cleft. In particular, 11 of the 12 polymorphic residues in the membrane-distal domain lie within the proposed α-helical sides (H1–3) or the β-stranded floor (S1–4) of the antigen-binding cleft. In the second β-strand (S2), the polymorphic residues alternate such that all three variable amino acids could be pointing into the cleft.

Surprisingly, the 3′ untranslated regions of both clones are identical except for an additional AC before the poly(A) tail for clone 11. To rule out the possibility of an artifact particular to these two clones, we sequenced the 5′ and 3′ ends of the remaining 15 clones. From 5′ sequences, eight clones were identical to clone 8 and nine to clone 11. At the 3′ end, all clones were identical for up to 250 nucleotides except for variation in the number of AC repeats (0–2) before the poly(A) tail. Therefore, the conserved 3′ untranslated region is not a peculiarity of clones 11 and 8.

Southern Blots. A convenient *Hinc*II site exists at the junction between the coding sequence and the 3′ untranslated region of the shark class IIB genes. We used the two fragments generated in this digest as separate probes in Southern blots of genomic DNA from two different nurse sharks. The sharks used in the Southern blot analysis were independently captured off the coast of Florida and are presumed to be unrelated to one another and to the shark from which the cDNA library was made.

With the coding-region probe, up to five bands were detected by hybridization (Fig. 2). Banding patterns differed for the two individuals. This polymorphism may indicate heterozygosity at the class IIB locus. Common bands may represent a shared allele and differing bands may signify unique alleles. Alternatively, the nurse shark may have more

than one class IIB locus. The possibility of more than one MHC-like locus is consistent with data from PCR amplifications using genomic DNA as a template. Eight unique PCR products were isolated from the DNA of one shark by using primers 1 and 2 (see *Materials and Methods*) and a third primer. Sequence similarity indicated that all of these products were immunoglobulin gene superfamily members and some shared similarity to MHC genes. One of these products was identical to clone 60C. Some of the other PCR products may represent different MHC loci, including pseudogenes (42). However, the sequences of these DNA fragments are sufficiently divergent that cross-hybridization with the clone 8 ORF probe would be unlikely.

The 3′ untranslated-region probe hybridizes to the same single band within the two DNAs (Fig. 2). There is no detectable polymorphism between these two individuals. The presence of this band indicates that this untranslated region is conserved within at least three presumably unrelated individuals. If the ORF-detected restriction length polymorphisms are representative of allelic differences within a single gene, then the 3′ regions of these alleles are remarkably conserved. If they represent more than one class IIB gene, then the 3′ regions of the other loci must have diverged beyond our level of detection. Fully understanding the complexity of the shark class IIB locus will require analysis of genomic clones.

Amino Acid Comparisons of MHC Proteins. The deduced amino acid sequence of clone 11 was used to search the Swiss-Prot 23 data base. The top 130 sequences found were all MHC proteins, with the top 43 being class IIβ proteins (29–34% similarity).

Alignment of the deduced amino acid sequence of clone 11 and representative protein sequences for all three MHC loci is shown in Fig. 3. This alignment provides convincing evidence that clone 11 does indeed encode a class IIβ protein. The overall length of the deduced shark protein indicates the two extracellular domain structure of class II molecules. Class I molecules have three extracellular domains (the first domain was removed for ease of alignment). Like most class IIB genes, clone 11 encodes four conserved cysteine residues, two within each extracellular domain. This is in contrast to class IIα chains, which do not contain cysteine residues in the membrane-distal domain. Since the membrane-distal domain contains most of the allelic variation, a clearer picture of the locus level differences is seen in the membrane-proximal domain. In this domain, the shark class IIβ chain exhibits several stretches of amino acid identity with other class IIβ proteins. In particular, almost all class IIB genes encode a conserved NGDWT(Y/F)Q sequence. This sequence is shared between shark and the endothermic vertebrates. Surprisingly, the fish sequences contain only five of these seven conserved residues, and the recently published frog class IIβ chains contain six of the seven residues (37). When overall sequence similarity within the membrane-proximal domain is compared, the deduced amino acid sequence for the shark class IIβ chain is more similar to mammalian (40–44%) and chicken (44%) class IIβ proteins than to the fish proteins (38%). The membrane-proximal domain of shark class IIβ chain also has more similarity to other shark MHC proteins, class IIα (40%) and class I (37%), than to mammalian class IIα (31%) and class I (31%) proteins.

Phylogenetic Analysis. In the initial analysis, the phylogenetic relationships of all three MHC loci, class I and class IIA

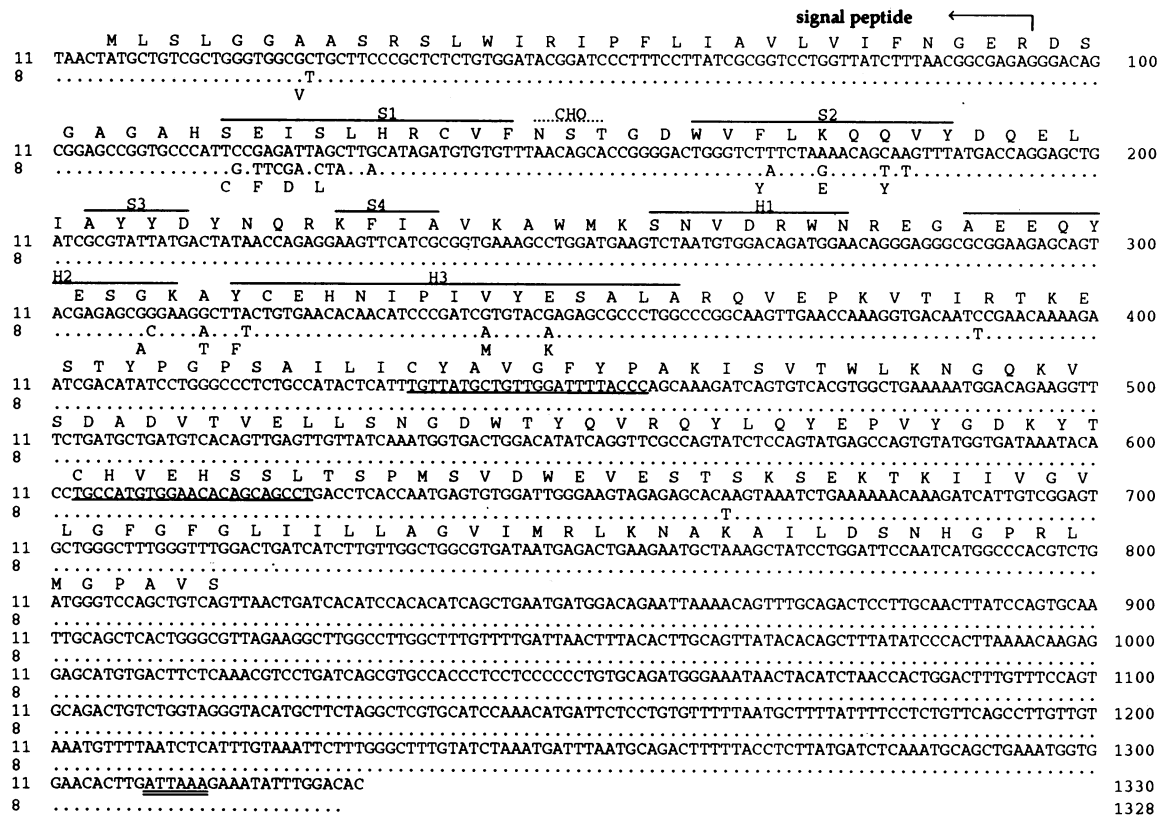

FIG. 1. The nucleotide and deduced amino acid sequences of nurse shark cDNA clones 11 and 8. Clone 11 contained two cloning adapters but was missing 12 nucleotides of coding sequence at the 5′ end. Thus, the first 17 nucleotides in the sequence for clone 11 were inferred from two additional clones of the same gene. By using the weight matrix of Von Heijne (28) and alignments with other MHC proteins, the signal peptide was assigned to the first 29 amino acids. CHO indicates a potential N-linked glycosylation site. The annealing sites of primers 1 and 2 are underlined. The putative polyadenylylation signal is doubly underlined. The predicted locations of the β-strands (S1–4) and the α-helices (H1–3) which compose the putative antigen-binding cleft were taken from alignments made with the predicted structure of Brown *et al.* (4).

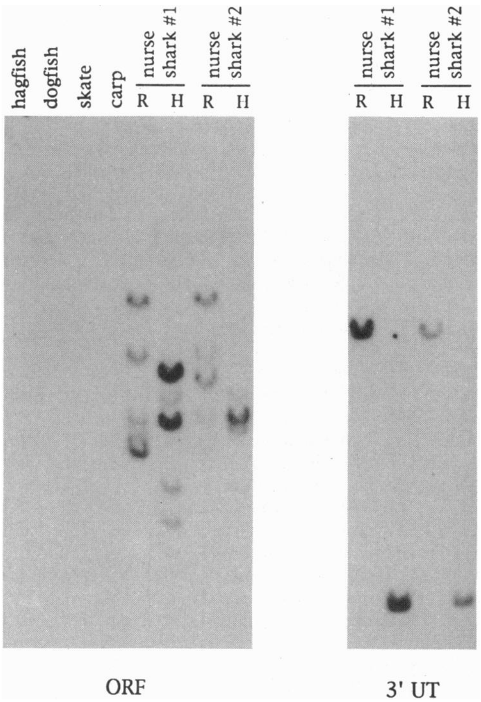

FIG. 2. Southern blot analyses of genomic DNA from cartilaginous and bony fish and two different nurse sharks (nos. 1 and 2). DNAs were digested with restriction enzymes *Eco*RI (unmarked and marked R) or *Hin*dIII (H). The blots were probed with either the open reading frame (ORF) or the 3′ untranslated region (3′ UT) of clone 8.

and IIB, were compared. We used only the deduced amino acids in the membrane-proximal domain and found that representatives of the three loci formed separate clusters in a neighbor-joining tree but that the bootstrap values for each node were low (data not shown). However, the shark β-chain sequence did cluster with the other β chains in this tree. Since only 93 amino acids from each protein were used in this first analysis and resulted in poor support, the complete sequences of only the class II β chains were used for the next analysis.

Fourteen class II β-chain protein sequences from nine species were used in both distance and parsimony analyses. Multiple sequences from the same species represent different class IIB loci within that species. All of these analyses yielded trees with the same topology with regard to the branching order of the shark and fish β chains relative to the other vertebrate β chains. Fig. 4 shows the resultant neighbor-joining tree with bootstrapping. The taxa branch in the generally accepted order of relatedness to mammals. Thus, the proteins from cartilaginous fish branch first, followed by bony fish, amphibian, and avian sequences. The simplest interpretation of the tree presented here is a single ancestral gene which gave rise to all the known β chains. In that view, subsequent gene duplications along the mammalian and bony fish lineage then gave rise to the different class IIB loci.

DISCUSSION

The data on the shark class IIB genes presented here, along with data on other shark MHC genes, indicate that class I and class IIA and IIB genes diverged before the split of the shark lineage and the other vertebrates. The class IIB genes of shark have structural similarity to other class IIB genes: they encode the same basic two-domain structure, each domain being disulfide-linked; the membrane-proximal domain contains most of the allelic polymorphisms. The polymorphic residues are mainly due to nonsynonymous nucleotide substitutions and are clustered along the α-helices and β-pleated sheets which compose the putative antigen-binding pocket. All of these structural attributes are shared by the well-characterized class IIβ molecules of mammals. On the basis of structure, it is reasonable to predict that the shark homologues could carry out the functions documented for the mammalian system, such as peptide binding and T-cell-receptor interaction.

```
              Membrane-distal domain
                             ▽                                                    ▽
shark β        DSGAGAHSEISLHRCVFNSTG-DWV-FLKQQVYDQELIAYYDYNQRKFIAVKAWMKSNVDRWN--REGAEEQYESGKAYCEHNIPIVYE--S-ALARQV   92
mouse Aβ       GNSERHFVHQFQPF.Y.TNGT.QRIRLVIRYI.NR.EYVRF.SDVGEYR..TELGRPDAEY....KQYL.RTRAELDTV.R..YEKTET..PTS.R.LE   93
mouse Eβ      VRDSRPWFL.YCKSE.H.YNGT.QR.RL.ERYF.NL.ENLRF.SDVGE.H..TELGRPDAEN..SQP.FL.QKRAEVDTV.R..YE-ISD..KFLVR.R.   91
human DPβ      RATPENYLFQGRQE.YAFNGT.Q--R..ERYI.NR.EF.RF.SDVGE.R..TELGRPAAEY..SQKDIL..KRAVPDRM.R..YE-LGG..PMT.Q.R.   93
human DQβ      RDSPEDFVYQFKGM.Y.TNGT.ER.RLVSRSI.NR.E.VRF.SDVGE.R..TLLGLPAAEY..SQKDIL.RKRAAVDRV.R..YQ-LEL..RTT.Q.R.   95
human DRβ      GDTRPRFL.Y.TSE.H.FNGT.ER.R..DRYF.N..EYVRF.SDVGE.R..TELGRPDEEY..SQKDFL.DRRAAVDT..R..YG-.G....FTVQ.R.   95
chicken β          FFQWSATVE.H.LNGT.ER.R..VRH..NRQQYVHF.SDVGL.V.DTVLGEPSAKLF.SQPDVL.KNRAAVEML.NY.YE-IVA..PLT.Q.R-   89
zebrafish β   FTGT.DGYYQYTMLE.IYSTSDYSDMVL.ESGSFNKVVDVQ.NSTVG.YVGYTEQGVIFARNF.KNQAYLQQRKAEVESF.R..AQ.SDS..AVR-DKA.   97
carp β             HG.YGYVQML.HASISQNVEFT.--SVN.NKFEYLR.NSTEK.VVGYTELGEKWAEDY...MILLAKGGLPVQQ-.RQ-LRMLTD..PF.-.LT.   86
mouse Iα2           .T.TLQWMYG.DVG.D..RLLRGYE.FA..GCDYIALNEDLKTWTFADMSSQITRRK.E.QAGA..YYRAYLEGE.VEWLHRYLKNGNAT.L.TD  180

              Membrane-proximal domain
                                  ▽                                                    ▽
shark β       EPKVTIRTKESTYPGPSAILICYAVGFYPAKISVTWLKNGQKVSDADVTVELLSNGDWTYQVRQYLQYEPVYGDKYTCHVEHSSLTSPMSVDW   185
mouse Aβ      Q.S.V.SLSRTEALNHHNT.V.SVTD......K.R.FR...EETVGVSSTQ.IR.....F..LVM.EMT.RR.EV.......P..K..IT.E.   187
mouse Eβ      ..T..VYPTKTQPLEHHNL.V.SVSD...GN.E.R.FR..KEEKTGI.STG.VR.....F.TLVM.ETV.QS.EV...Q...P...D.VT.E.   184
human DPβ     Q.R.NVSPSKKGPLQHHNL.V.HVTD...GS.Q.R.FL...EETAGV.STN.IR.....F.ILVM.EMT.QQ..V...Q...T..D..VT.E.   186
human DQβ     ..T...SPSRTEALNHHNL.V.SVTD....Q.K.R.FR.D.EETAGV.STP.IR.....F.ILVM.EMT.QR..V.......P..Q..IT.E.   188
human DRβ     H....VYPSKTQPLQHHNL.V.SVS....GS.E.R.FR...EEKTGV.STG.IH.....F.TLVM.ETV.RS.EV...Q...P.V...LT.E.   188
chicken β     ....R.FALQ.GSLPQTDR.A..VT....PE.E.K.FQ...EETERV.STDVIQ........LVV.EIS.RH..S.V.Q...T..QQ.ITQR.   182
zebrafish β   K.....QSVMQAEGKHP.M.L.D.YE...K..KMS..RDDKV.TSDVTSTIEMA..N.Y..IHSH.E.T.KS.E.IQ.V...A.S.Q.ITKE.   190
carp β        K.E.I..SDREAKGNEK.V.V.S.YD...KP.KL..MRDDKE.TTDVTST.E.AD...Y..IHSH.E.F.KP.E.IS.V...A.SHK..IYH.   179
shark Iα3     K.S.N.T.HRLHRGKDPLL...HVN....SG.NA...H..GTIQQEVLSSRI.P.T.G.F.TTLQISVT.QSR.T...Q.....S.DKLTAT.
mouse Iα3     S..AHV--THHPRSKGEVT.R.W.L.....D.TL..QL..EELTQDMEL..TRPA..G.F.KWASVVVPLGKEQN...R.Y.EG.PE.LTLR.   271
shark α2      P.E.SVYSEDLVEW.QLNT...F.D....PH.TMK.RR.NEPMT.G.NIT.FYIKD.F..RRFS..SIV.SP..M.S........QD.VT.F.   179
mouse α2      A.QA.VFP.SPVLL.QPNT...FVDNIF.PV.NI...R.SKS.T.GVYETSFFV.R.YSFHKLS..TFI.SDD.I.D.K...WG.EE.VLKH.   182

              Connecting peptide    Transmembrane region        Cytoplasmic region

shark β       EVESTSKSEKTK          IIVGVLGFGFGLIILLAGVIM       RLKNAKAILDSNHGPRLMGPAVS                        241
mouse Aβ      RAQ--.E.ARS.          MLS.IG.CVL.V.F.GL.LFI       .HRSQ.GPRGPPPAGL.Q                             236
mouse Eβ      KAQ--.T.AQN.          MLS..G..VL..LF.G..LFIYF     .NQKGQSG.QPTGLLS                               233
human DPβ     KAQ--.D.ARS.          TLT.AG..VL....CGV.IFM       HRRSK.VQRG.A                                   229
human DQβ     RAQ--.E.AQS.          MLS.IG..VL...F.GL.L.I       HHRSQ.GL.H                                     229
human DRβ     RAR--.E.AQS.          MLS..G..VL..LF.G..LFIYF     .NQKGHSG.QPRGFLS                               237
chicken β     .PP--GDVSRS.          LLM..G..VL..VY.AL.IFFFL     CS.KGQPDPT.PGILN                               231
zebrafish β   NPH-I.E.DRN.          FAI.AS.LVL.I..AI..L.YY      KK.STGR..VP.                                   235
```

FIG. 3. Alignment of MHC protein sequences. Assignment of domains is based on the alignment and previously published information. Triangles point to the four conserved cysteines. Dashes indicate gaps and dots signify identity with the top sequence. The number of the last residue is shown on the right. The sources of class IIβ sequences (marked β) are as follows: shark (clone 11, this paper), mouse A (29) and E (30), human DP (31), DQ (32) and DR (33), chicken (34), zebrafish (10), and carp (9). Other sequences aligned include shark class I (15), mouse class I (35) (α2 and α3 domains shown), shark class IIα (16), and mouse class IIα (36) (only α2 domains are aligned).

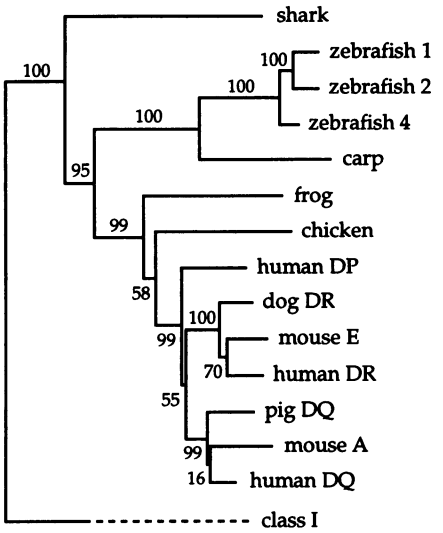

FIG. 4. Phylogenetic tree showing the relationship of representative MHC class IIβ chains. Class I molecules, without the α1 domain, from human (38), mouse (35) and chicken (39) were used as the outgroup and are subsumed under the label class I. Both nurse shark sequences from this paper were used and are combined under the label shark. The three zebrafish sequences represent separate loci (10). Numbers along branches are percentages of 2000 bootstraps supporting each partitioning. All sources of sequences are listed in the legend to Fig. 3 except dog DR (40), pig DQ (EMBL no. M31497), and frog (37).

It has been difficult to isolate MHC genes, especially both the class IIA and IIB genes, from nonmammalian vertebrates. The isolation from shark of class IIB genes (this study) and class IIA genes (16) represents one of the first documentations of both loci in a nonmammalian organism. Since both genes were isolated from the same cDNA library, both genes are expressed in the same individual. The next stage shall be to test for linkage of these two loci, as well as for the degree of polymorphism contained in the presumed heterodimers they encode.

Bony fish display many hallmarks of T-cell-dependent immune responses (17–19). In contrast, no T-cell-mediated immune responses such as allograft rejection have been documented in cartilaginous fish (20–22), despite the presence of MHC class IIA and IIB gene products. This paradox could be due to technical difficulties at several levels in revealing T-cell immunity, especially if the only test is allograft rejection. It is also possible that shark MHC molecules may be less polymorphic and thus unable to stimulate acute graft rejections. However, the degree of polymorphism detected thus far would argue against this possibility. Third, MHC molecules may be expressed on fewer cells in shark than in other vertebrates, limiting shark T-cell responses to fewer tissues. Lastly, MHC restriction of T cells may be different in shark than in other vertebrates. With the isolation of MHC genes from shark, one can now characterize the role of polymorphism, tissue distribution, and MHC restriction in the shark and start to address some of these issues.

The isolation of MHC genes in shark provides encouragement for the search for similar molecules in organisms which share even more distant ancestors with mammals, such as jawless fish or protochordates. Like shark, these chordates do not have documented T-cell responses, yet they do exhibit blood cell-mediated allorecognition. For example, a colonial tunicate, *Botryllus schlosseri*, possesses a highly polymorphic gene locus which dictates historecognition (41).

We thank Drs. Joseph M. Quattro and David W. Stock for their patient assistance with the phylogenetic analyses, and we are grateful to Dr. Jonathan B. Geller for critical reading of the manuscript. This work is supported by National Cancer Institute Grant CA42551.

1. Hood, L., Steinmetz, M. & Malissen, B. (1983) *Annu. Rev. Immunol.* **1**, 529–568.
2. Germain, R. (1988) *Cell* **54**, 441–444.
3. Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. (1987) *Nature (London)* **329**, 506–512.
4. Brown, J. H., Jardetzky, T., Saper, M. A., Samraoui, B., Bjorkman, P. J. & Wiley, D. C. (1988) *Nature (London)* **332**, 845–850.
5. Bjorkman, P. J., Saper, M. A., Samraoui, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. (1987) *Nature (London)* **329**, 512–518.
6. Klein, J. (1987) *Natural History of the Major Histocompatibility Complex* (Wiley, New York).
7. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335**, 167–170.
8. Hughes, A. L. & Nei, M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 958–962.
9. Hashimoto, K., Nakanishi, T. & Kurosawa, Y. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6863–6867.
10. Ono, H., Klein, D., Vincek, V., Figueroa, F., O'hUigin, C., Tichy, H. & Klein, J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 11886–11890.
11. Hordvik, I., Grimholt, U., Fosse, V. M., Lie, Ø. & Endresen, C. (1993) *Immunogenetics* **37**, 437–441.
12. Grimholt, U., Hordvik, I., Fosse, V. M., Olsaker, I., Endresen, C. & Lie, Ø. (1993) *Immunogenetics* **37**, 469–473.
13. Ono, H., Figueroa, F., O'hUigin, C. & Klein, J. (1993) *Immunogenetics* **38**, 1–10.
14. Dixon, B., Stet, R. J. M., van Erp, S. H. M. & Pohajdak, B. (1993) *Immunogenetics* **38**, 27–34.
15. Hashimoto, K., Nakanishi, T. & Kurosawa, Y. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 2209–2212.
16. Kasahara, M., Vazquez, M., Sato, K., McKinney, E. C. & Flajnik, M. F. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 6688–6692.
17. Hildemann, W. H. (1958) *Immunology* **1**, 46–53.
18. Miller, N. W., Deuter, A. & Clem, L. W. (1986) *Immunology* **59**, 123–128.
19. Miller, N. W., Sizemore, R. C. & Clem, L. W. (1985) *J. Immunol.* **134**, 2884–2888.
20. Borysenko, M. & Hildemann, W. H. (1970) *Transplantation* **10**, 545–551.
21. Tam, M. R., Reddy, A. L., Karp, R. D. & Hildemann, W. H. (1976) in *Comparative Immunology*, ed. Marchalonis, J. J. (Wiley, New York), pp. 98–119.
22. McCumber, L. J., Sigel, M. M., Trauger, R. J. & Cuchens, M. A. (1982) in *The Reticuloendothelial System*, eds. Cohen, N. & Sigel, M. M. (Plenum, New York), Vol. 3, pp. 393–422.
23. Seutin, G., White, B. N. & Boag, P. T. (1990) *Can. J. Zool.* **69**, 82–90.
24. Higgins, D. G. & Sharp, P. M. (1988) *Gene* **73**, 237–244.
25. Kimura, M. (1983) *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press, Cambridge, U.K.).
26. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
27. Swofford, D. L. (1991) *PAUP: Phylogenetic Analysis Using Parsimony, Version 3.0s* (Illinois Nat. Hist. Survey, Champaign).
28. Von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683–4690.
29. Estess, P., Begovich, A. B., Koo, M., Jones, P. P. & McDevitt, H. O. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 3594–3598.
30. Widera, G. & Flavell, R. A. (1984) *EMBO J.* **3**, 1221–1225.
31. Kelly, A. & Trowsdale, J. (1985) *Nucleic Acids Res.* **13**, 1607–1621.
32. Boss, J. M. & Strominger, J. L. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 5199–5203.
33. Tieber, V. L., Abruzzini, L. F., Didier, D. K., Schwartz, B. D. & Rotwein, P. (1986) *J. Biol. Chem.* **261**, 2738–2742.
34. Bourlet, Y., Behar, G., Guillemot, F., Frechin, N., Billault, A., Chausse, A. M., Zoorob, R. & Auffray, C. (1988) *EMBO J.* **7**, 1031–1039.
35. Moore, K. W., Sher, B. T., Sun, Y. H., Eakle, K. A. & Hood, L. E. (1982) *Science* **215**, 679–682.
36. Benoist, C. O., Mathis, D. J., Kanter, M. R., Williams, V. E. & McDevitt, H. O. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 534–538.
37. Sato, K., Flajnik, M. F., DuPasquier, L., Katagiri, M. & Kasahara, M. (1993) *J. Immunol.* **150**, 2831–2843.
38. Holmes, N. & Parham, P. (1985) *EMBO J.* **4**, 2849–2854.
39. Guillemot, F., Billault, A., Pourquie, O., Behar, G., Chausse, A.M., Zoorob, R., Kreibich, G. & Auffray, C. (1988) *EMBO J.* **7**, 2775–2785.
40. Sarmiento, U. M. & Storb, R. (1990) *Immunogenetics* **31**, 396–399.
41. Weissman, I. L., Saito, Y. & Rinkevich, B. (1990) *Immunol. Rev.* **113**, 227–241.
42. Bartl, S. & Weissman, I. L. (1994) *Ann. N.Y. Acad. Sci.*, in press.