

BRIEF COMMUNICATION

Michael F. Criscitiello · Ralph Benedetto
 Aurita Antao · Melanie R. Wilson
 V. Gregory Chinchar · Norman W. Miller
 L. William Clem · Thomas J. McConnell

β_2 -microglobulin of Ictalurid catfishes

Received: 15 December 1997 / Revised: 13 March 1998

Key words Beta-2 microglobulin · Catfish · MHC · Phylogeny

Major histocompatibility complex (MHC) class I molecules are expressed on the surface of virtually every vertebrate nucleated cell. The MHC-encoded class I heavy chain is non-covalently bound to the β_2 -microglobulin (β_2m), an association necessary for proper folding, peptide binding, and surface display of class I antigens (Hansen et al. 1989; Vitiello et al. 1990). Unlike the classical class I heavy chain genes, the β_2m chain gene is not linked to the MHC loci in human and mouse and is not polymorphic (Klein 1986).

Although β_2m sequences have been determined in only a few species of fish and birds, a high degree of sequence conservation has been noted in β_2m among all vertebrate species in which it has been identified. β_2m closely resembles the structure of an immunoglobulin (Ig) constant region domain (Peterson et al. 1972; Smithes et al. 1972) and also exhibits similarity to the MHC class I $\alpha 3$ domain (Williams et al. 1988). Both of the genes encoding these (Ig) superfamily proteins are thought to have arisen by duplication of a common ancestral gene (Burnet 1970; Gally et al. 1972). Hence, the study of β_2m in lower vertebrates may shed light on the origin of the MHC.

In addition, the fact that the channel catfish is presently the best characterized fish in vitro system with regard to cellular aspects of adaptive immunity makes it a logical choice for detailed studies of MHC and related molecules (Miller et al. 1985, 1986, 1994). For example, previous in vitro studies have shown that catfish have

the functional equivalents of T, B, NK, and accessory cells which interact in an alloantigen (presumably MHC)-restricted fashion (Clem et al. 1991; Vallejo et al. 1992). Consequently the work reported here was undertaken to identify and sequence the β_2m chain gene from six North American catfishes: channel catfish, *Ictalurus punctatus*; Headwater catfish, *I. pricei*; blue catfish, *I. furcatus*; *I. n.sp.* (*Ictalurus* new species, referred to here as the chihuahua catfish, Humphries and W.W. Miller, personal communication); white catfish, *Ameiurus catus*; and yellow bullhead *Ameiurus natalis*.

A partial cDNA sequence for β_2m was amplified from the cloned B-cell line, 1B10 (Miller et al. 1994), and used to probe a channel catfish 42TA macrophage cDNA library (Luft et al. 1996). Five phage clones were plaque purified and two were sequenced. Both these clones exhibited significant similarity to known β_2m sequences in BLAST (Altschul et al. 1990) searches of the GenBank database. The longest cDNA clone [1170 base pairs (bp)] contained a 5' untranslated region, the entire β_2m coding sequence, and a 3' untranslated region containing four possible polyadenylation signal sites. Figure 1A shows the nucleotide and inferred amino acid sequences of this cDNA. The channel catfish β_2m mature protein is predicted to be 97 amino acids in length with a 19 aa leader. The characteristic cysteines forming the intradomain disulfide bridge (Williams and Barclay 1988) are found at residues 25 and 80. Based on sequences obtained from polymerase chain reaction (PCR) amplification of genomic DNA from different catfish species (see below) the exon boundaries of channel catfish β_2m are readily identified. At the 3' end of exon 3, i.e., amino acids 90–93, is a potential carbohydrate acceptor site (Asn-Ile-Ser). To date no β_2m chain gene containing an N-linked glycosylation signal sequence(s) has been reported in any species.

Channel catfish primers TM301 and TM307, based on the cDNA sequence, were used to amplify the β_2m chain gene from genomic DNA of six species of catfish: *I. punctatus*, *I. n.sp.*, *I. pricei*, *I. furcatus*, *A. catus*, and

M.F. Criscitiello · R. Benedetto · T.J. Mc Connell (✉)
 Department of Biology, 108 Howell Science Complex,
 East Carolina University, Greenville, NC 27858, USA

A. Antao · M.R. Wilson · V.G. Chinchar · N.W. Miller ·
 L.W. Clem
 Department of Microbiology,
 University of Mississippi Medical Center,
 Jackson, MS 39216, USA

	1	11	21	31	41	51	61	71	81	91
CONSENSUS	KEAPPKIQVY	SRNPGFEGKE	NTLICHVSDF	HPPDI-IDLL	KNGEVIPNAE	QTDLAFEKGW	KFHLTKSV-F	TPTS-DEY-C	RVRH..LKET	K--SWEPDM
<i>I. punctatus</i>	-----	-----	-----	-----T-----	-----	-----	-----S-----	-----N-KFT-----	-----	-----NI-----
<i>I. n.sp.</i>	-----	-----	-----	-----T-----	-----	-----	-----S-----	-----N-----	-----	-----NI-----
<i>I. pricei</i>	-----	-----	-----	-----T-----	-----A-----	-----	-----S-----	-----N-----	-----	-----NI-----
<i>I. furcatus</i>	-----	-----A-----	-----	-----T-E-----	-----	-----	-----S-----	-----S-KFT-----	-----D-----	-----NI-----
<i>A. natalis</i>	-----T-----	-----	-----R-----	-----E-----	-----G-K-----	-----	-----T-----	-----S-R-V-----	-----TE-----	-----IFT-----
<i>A. catus</i>	-----T-----	-----	-----	-----E-----	-----A-G-K-----	-----	-----Q-----	-----G-SFA-----	-----NE-----	-----TYT-----
<i>O. mykiss</i>	-----S-V-----	-----N-DK-----	-----G-----	-----S-Q-I-----	-----VE-D-K-----	-----Q-----	-----Q-----	-----A-GE-T-----	-----H-----	-----NL-TYT--A-----
<i>C. carpio</i>	-TSS--V--	-HF--Y--	-----G-----	-----T-E-----	-D--IL--TQ-----	-----	-----Q-----	-K-ERGQN-A-S-----	-----MNNK NIY-----	-----N-----
<i>D. rerio</i>	-VST--VH--	-HF--Y--P-----	-----Y--S-----	-----S-E-----	-----Q-MSDTK-----	-----	-----Q-----	-----A-----	-----ERK--T-S-----	-----M-----
<i>G. gallus</i>	ADLT--V--	--F-ASA-TK	-V-N-FAAG--	---K-S-T-M-----	-D-VPMEG-Q	YS-MS-NDD-	T-QRLVHAD-	--S-GST-A-	K-E-ET--P	QVYK-D-EF
<i>M. musculus</i>	IQKT-Q----	--H-P-N--P-----	-I-N-Y-TQ-----	---H-E-QM-----	---KK--KV-----	MS-MS-S-D-	S-YILAHTE-	---ET-T-A-	--K-DSMA-P	-TVY-DR--
<i>H. sapiens</i>	IQRT-----	--H-A-N--S-----	-F-N-Y--GH-----	-QS--EV-----	-----R-EKV-----	HS--S-S-D-	S-Y-LYTE-	---EK--A-----	--N-VT-SQR	-IVK-DR--
	i+ +ii	ib i + i+	i i i	i +i+o+o	i+i o +	+io i++ i +i		i +i + ++o	i+ii o +	+ o ia +o

Fig. 2 Amino acid sequence alignment of the six catfish β_2m mature protein sequences with those of other vertebrate species. The majority consensus sequence is shown at the top; identity with consensus is shown by dashes, and spaces for amino acids absent in fish, but present in mammals and birds, are indicated by dots. The bottom row indicates the nature of the substitutions. Amino acid groupings follow the scheme of Smith and Smith (1990): acidic (Asp, Glu), basic (Lys, Arg, His), amino (Asn, Gln), small polar (Ser, Thr), aliphatic (Ile, Leu, Val), aromatic uncharged (Phe, Trp, Tyr), and small hydrophobic (Ala, Gly). Changes are indicated by the letters *a*, *b*, *n*, *p*, *c*, *r*, and *h*. Invariant, polar, and nonpolar residues are indicated by *i*, *+*, and *o*, respectively. The N-linked glycosylation sequence of the *Ictalurus* species is in boldface. References for other included sequences: *Oncorhynchus mykiss* Jb-10 (rainbow trout) (Shum et al. 1996), *Cyprinus carpio* (carp) (Dixon et al. 1993), *Brachydanio rerio* (zebrafish) (Ono et al. 1993), *Gallus gallus* (chicken) (Kaufman et al. 1992), *Mus musculus* (mouse) (Daniel et al. 1983), *Homo sapiens* (human) (Gussow et al. 1987).

A. natalis. PCR yielded products of 1279, 1547, 1548, 1562, 1438, and 1474 bp sequences, respectively. Figure 1B is a schematic of channel catfish β_2m showing the exon and intron structure. As commonly found for teleost MHC genes, each of the catfish β_2m chain introns are phase one, with the intron interrupting the codon between the first and second nucleotides. The exon/intron splice sites for the six catfish genes correlate precisely with the only other published teleost genomic β_2m sequence, that of the zebrafish *Brachydanio rerio* (Ono et al. 1993). In catfish β_2m , the leader sequence (19 codons), the first two codons of the mature peptide and the first base of the 3rd codon (serine) are encoded by exon 1. Exon 2 encodes the two remaining bases of the serine codon through the first base of codon 94 (glutamate). The remainder of the glutamate 94, the last three amino acids, and the stop codon are encoded by exon 3. Each of the six catfish β_2m chain genes show high sequence similarity to one another, particularly in the coding regions. The most similar (*I. punctatus* and *I. n.sp.*) differ only in their intron sequences; *I. pricei* differs from *I. punctatus* by only two nucleotides (one amino acid) in the coding region; and *I. pricei*, *I. furcatus*, *A. catus*, and *A. natalis* are 99%, 96%, 87%, and 83% similar to *I. punctatus*, respectively, in amino acid sequence.

An alignment of the six catfish β_2m amino acid sequences with those of other vertebrate species is shown in Fig. 2. Twenty-two amino acids that are invariant in all other published vertebrate β_2m sequences are also invariant in the catfish species (Dixon et al. 1993; Ellis

and Martin 1993; Ellis et al. 1993, 1995; Ono et al. 1993; Milland et al. 1993; Riegert et al. 1996; Ruiz et al. 1994; Shum et al. 1996). Each of the four *Ictalurus* species contain the exon 3 glycosylation signal (marked in bold) encoded by the identical nine nucleotide sequence, whereas the *Ameirus* catfish species, rainbow trout (*Oncorhynchus mykiss*), and carp (*Cyprinus carpio*) miss having an N-linked glycosylation signal sequence by one nucleotide (data not shown). When compared with mammalian and bird β_2m , each of the known teleost β_2m (Dixon et al. 1993; Ono et al. 1993; Shum et al. 1996) are two amino acids shorter, i.e., they lack amino acids 85 and 86 in exon 3. This shortening of exon 3 could affect, at least for the *Ictalurus* species, the location of the glycosylation signal sequence in the folded protein. Using diagrammatic representations of human β_2m as a guide (Bjorkman et al. 1987), one can postulate the structure of β_2m lacking these two amino acids. Since these amino acids occur in the turn between β strands 6 and 7, their absence could shift the location of the potential glycosylation site (amino acids 92–94 by human numbering) to the more exposed turn region.

The appearance of a single band in Southern blot analyses of genomic DNA from five individual fish digested separately with three different restriction enzymes (data not shown) suggest the presence of only one β_2m chain locus in the channel catfish. Although β_2m is usually encoded by a single-copy gene (Klein 1986), recently two examples of multiple loci have been described in teleosts. Dixon and co-workers (1993) identified two loci in gynogenetic carp, probably the result of tetraploidy in this species. Ten different β_2m sequences have been cloned and sequenced from an individual rainbow trout, a number that the tetraploidy of this species cannot account for (Shum et al. 1996). The authors of that study speculate that the β_2m chain gene in this salmonid species has remained in the MHC proper and has consequently been subjected to duplication events that are thought to give rise to the large diversity of MHC alleles. Their Southern blot data also indicate the presence of multiple β_2m chain loci, but the manner in which their sequences fall into homologous groups suggest polymorphism. Perhaps both polymorphism and multiple alleles contribute to the high number of rainbow trout β_2m sequences, as opposed to the simple non-polymorphic locus probably present in channel catfish.

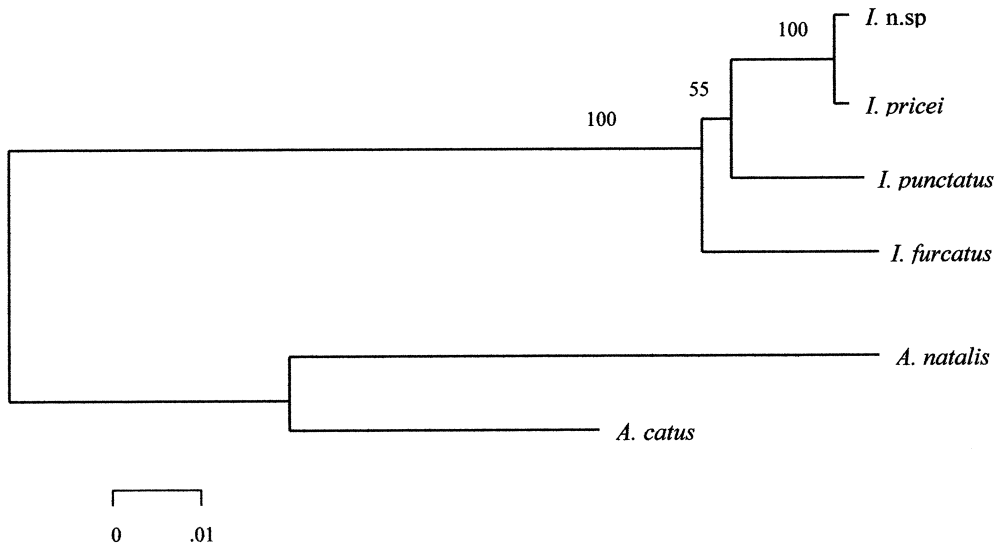


Fig. 3 Phylogenetic tree of β_2m mature peptide-coding nucleotide sequences from different vertebrates. Searches for sequences similar to catfish β_2m were performed and preliminarily aligned using the NCBI BLAST E-mail server (Altschul et al. 1990). The final nucleotide sequence alignments were performed using the PILEUP and PRETTY programs of the Genetics Computer Group (GCG) (Devereux et al. 1984). Pairwise distances were calculated for intron 1 using the P -distance algorithm β_2m and the neighbor-joining method of Saitou and Nei (1987). Calculations and dendrogram construction were completed using the Molecular Evolutionary Genetic Analysis (MEGA) (Penn State University, University Park, Penn.) programs. Numbers on nodes indicate the frequency with which this node was recovered per 100 bootstrap replications in a total of 500 replications. The MEGA program treated both *Amerius* species as a single outgroup, therefore no bootstrap value was generated for this node. References for individual sequences are as in Fig. 2

A dendrogram was constructed employing the neighbor-joining method on distances calculated using the p -distance algorithm (MEGA, Penn State University, University Park, Penn.) for the intron 1 sequence of the six catfish species (Fig. 3). Intron 1 is about 900 bp long in zebrafish (incomplete sequence) compared with 810, 1078, 1079, 1091, 974, and 1010 bp in *I. punctatus*, *I. n.sp.*, *I. pricei*, *I. furcatus*, *A. catus*, and *A. natalis*, respectively. Insertions/deletions in the sequence comparisons were not weighted in the calculations used for generation of the dendrogram. The topology of this dendrogram, unlike a dendrogram based on nucleotide coding regions, is identical to one generated using 384 nucleotides of the 5' end of the mitochondrial cytochrome *B* gene (data not shown). This dendrogram (Fig. 3) is consistent with phylogenetic relationships predicted using morphological characteristics (Lundberg 1992). This analysis places *I. n.sp.* with *I. pricei*, and places *I. punctatus* outside of and ancestral to these two. The phylogeny of these closely related *Ictalurus* species has never been resolved. The phylogeny offered in Fig. 3 is consistent with the fishes' present geographic range: *I. n.sp.* and *I. pricei* are restricted in geographic range to Mexican and Southwestern arid re-

gions, while the native range of *I. punctatus* includes much of North America (Peterson 1991). *I. furcatus* places as ancestral to the other *Ictalurus* species as expected (Lundberg 1992). The white catfish (*A. catus*) is placed with *A. natalis*, supporting the movement of the white catfish from the *Ictalurus* to the *Amerius* genus (Hubbs and Lagler 1958; Lundberg 1982). Also, *A. natalis* and *A. catus* β_2m do not contain the glycosylation signal found in the four *Ictalurus* catfish sequences.

In conclusion, we identified a unique glycosylation signal sequence of β_2m in *Ictalurus punctatus* and found the site in the only three other members of the genus that were sampled. Two species of closely related *Amerius* catfishes lack the encoded consensus glycosylation site. No previously reported β_2m chain genes contain sequences signaling for glycosylation. Experiments are underway to determine whether or not the catfish β_2m protein product is glycosylated. A glycosylated β_2m could possibly associate with calnexin and/or calreticulin in the endoplasmic reticulum, as proposed for class I molecules during folding and assembly (Parham 1996), in *Ictalurus* catfishes. Also, a phylogenetic tree generated from the catfish β_2m sequences corroborates and extends recent morphological analyses of the phylogeny of these fishes, subsequent work should further elucidate the phylogeny of the family Ictaluridae.

Acknowledgments We thank B. Jenson (Dexter National Fish Hatchery, Dexter, New Mexico), J. Layton (Layton and Smith Fishery, Windsor, North Carolina), and A.M. Snyder (Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico) for generously providing fish samples. We thank R. Miller (Museum of Zoology, University of Michigan) for discussions about the chihuahua catfish. This work was supported by grants from ECU (95-RS-37), NIH (R37-AI-19530), and USDA (95-37204-2225).

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. Basic local alignment search tool. *J Mol Biol* 215: 403–410, 1990
- Bjorkman, P.J., Saper, M.A., Samraoui, B., Bennet, W.S., Strominger, J.L., and Wiley, D.C. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329: 512–515, 1987
- Burnet, F.M. A certain symmetry: histocompatibility antigens compared with immunocyte receptors. *Nature* 226: 123–126, 1970
- Clem, L.W., Miller, N.W., and Bly, J.E. Evolution of lymphocyte subpopulations, their interactions and temperature sensitivities. In SONDZEICHEN symbol “Z” (eds.): *The phylogenesis of immune functions*, pp. 191–212, CRC Press, Boca Raton 1991
- Daniel, F., Morello, D., Le Bail, O., Chambon, P., Cayre, Y., and Kourilsky, P. Structure and expression of the mouse β_2 -microglobulin gene isolated from somatic and non-expressing teratocarcinoma cells. *EMBO J* 2: 1061–1065, 1983
- Devereux, J., Haeblerli, P., and Smithies, O. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12: 387–395, 1984
- Dixon, B., Stet, R.J.M., van Erp, S.H.M., and Pohajdak, B. Characterization of β_2 -microglobulin transcripts from two teleost species. *Immunogenetics* 38: 27–34, 1993
- Ellis, S.A. and Martin, A.J. Nucleotide sequence of horse beta 2-microglobulin cDNA. *Immunogenetics* 38: 383, 1993
- Ellis, S.A., Braem, K.A., and Payne, L.K. Nucleotide sequence of cattle beta 2-microglobulin cDNA. *Immunogenetics* 38: 310, 1993
- Ellis, S.A., Pichowski, J.S., and Staines, K.A. Sequence divergence of *B2m* in different cattle populations. *Immunogenetics* 42: 229–230, 1995
- Gally, J.A. and Edelman, G.M. The genetic control of immunoglobulin synthesis. *Annu Rev Genet* 6: 1–46, 1972
- Gussow, D., Rein, R., Ginjaar, I., Hochstenbach, F., Seemann, G., Kottman, A., and Ploegh, H.L. The human β_2 -microglobulin gene: primary structure and definition of the transcriptional unit. *J Immunol* 139: 3132–3138, 1987
- Hansen, T.H., Myers, N.B., and Lee, D.R. Studies of two antigenic forms of L^d with disparate β_2 -microglobulin (β_2m) associations suggest that β_2m facilitates the folding of the $\alpha 1$ and $\alpha 2$ domains during de novo synthesis. *J Immunol* 140: 3522–3527, 1989
- Hubbs, C.L. and Lagler, K.F. *Fishes of the Great Lakes region* (revised edition), Cranbrook Institute Scientific, Bloomfield Hills, 1958
- Kaufman, J., Anderson, R., Avila, D., Engberg, J., Lambris, J., Salomonsen, J., Welinder, K., and Skjodt, K. Different features of the MHC class I heterodimer have evolved at different rates. *J Immunol* 148: 1532–1546, 1992
- Klein, J. *The Natural History of the Major Histocompatibility Complex*, John Wiley, New York, 1986
- Luft, J.C., Wilson, M.R., Bly, J.E., Miller, N.W., and Clem, L.W. Identification and characterization of a heat shock protein 70 family member in channel catfish (*Ictalurus punctatus*). *Comp Biochem Physiol B* 113: 169–174, 1996
- Lundberg, J.G. The comparative anatomy of the toothless blindcat, *Trogloglanis pattersoni* Eigenmann, with a phylogenetic analysis of the ictalurid catfishes. *Misc Publ Mus Zool Univ Mich* 163: 1–85, 1982
- Lundberg, J.G. The phylogeny of Ictalurid catfishes: a synthesis of recent work. In SONDZEICHEN symbol “Z” (eds.): *Systematics, historical ecology, and North American freshwater fishes*, pp. SONDZEICHEN symbol “Z”, Stanford: Stanford University Press, 392–420, 1992
- Milland, J., Loveland, B.E., and McKenzie, I.F. Isolation of a clone for pig beta 2-microglobulin cDNA. *Immunogenetics* 38: 464, 1993
- Miller, N.W., Sizemore, R.C., and Clem, L.W. Phylogeny of lymphocyte heterogeneity: the cellular requirements for in vitro antibody responses of channel catfish leukocytes. *J Immunol* 134: 2884–2888, 1985
- Miller, N.W., Deuter, A., and Clem, L.W. Phylogeny of lymphocyte heterogeneity: the cellular requirements for the mixed leukocyte reaction with channel catfish. *Immunol* 59: 123–128, 1986
- Miller, N.W., Ryczyn, M.A., Wilson, M.R., Warr, G.W., Naftel, J.P., and Clem L.W. Development and characterization of channel catfish long term B cell lines. *J Immunol* 152: 2180–2189, 1994
- Ono, H., Figueroa, F., O’huigin, C., and Klein, J. Cloning of the β_2 -microglobulin gene in the zebrafish. *Immunogenetics* 38: 1–10, 1993
- Parham, P. Functions for MHC class I carbohydrates inside and outside the cell. *Trends Biochem Sci* 21: 427–433, 1996
- Peterson, P.A., Cunningham, B.A., Berggard, I., and Edelman, G.M. β_2 -microglobulin – a free immunoglobulin domain. *Proc Natl Acad Sci USA* 69: 1697–1701, 1972
- Peterson, R.T. (ed.): *A field guide to freshwater fishes of North America North of Mexico*, Houghton, Boston, 1991
- Riegert, P., Anderson, R., Bumstead N., Dohring, C., Dominguez-Steglich, M., Engberg, J., Salomonsen, J., Schmid, M., Schwager, J., Skjodt, K., and Kaufman, J. The chicken beta 2-microglobulin gene is located on a non-major histocompatibility complex microchromosome: a small, G+C rich gene with X and Y boxes in the promoter. *Proc Natl Acad Sci USA* 93: 1243–1248, 1996
- Ruiz, R.E., Hall, B.L., Doyle, C., and Ward, F.E. Baboon and cotton-top tamarin B2 m cDNA sequences and the evolution of primate beta 2-microglobulin. *Hum Immunol* 39: 188–194, 1994
- Saitou, N. and Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406–425, 1987
- Shum, B.P., Azumi, K., Zhang, S., Kehrer, S.R., Raison, R.L., Detrich, H.W., and Parham, P. Unexpected β_2 -microglobulin sequence diversity in individual rainbow trout. *Proc Natl Acad Sci USA* 93: 2779–2784, 1996
- Smith, R.F. and Smith, T.F. Automatic generation of primary sequence patterns from sets of related protein sequences. *Proc Natl Acad Sci USA* 87: 118–122, 1990
- Smithies, O. and Poulik, M.D. Dog homologue of human β_2 -microglobulin. *Proc Natl Acad Sci USA* 69: 2914–2917, 1972
- Vallejo, A., Miller, N., and Clem, W. Antigen processing and presentation in teleost immune responses. *Annu Rev Fish Dis* 2: 73–89, 1992
- Vitiello, A., Potter, T.A., and Sherman, L.A. The role of β_2 -microglobulin in peptide binding by class I molecules. *Science* 250: 1423–1426, 1990
- Williams, A.F. and Barclay, A.N. The immunoglobulin superfamily domains for cell surface recognition. *Annu Rev Immunol* 6: 381–405, 1988