The α Chain Gene of H-2O Has an Unexpected Location in the Major Histocompatibility Complex

By Lars Karlsson and Per A. Peterson

From the Department of Immunology, The Scripps Research Institute, La Jolla, California 92037

Summary

A previously unknown major histocompatibility complex class II molecule consisting of the β chain encoded by the *H-2Ob* gene and an unknown α chain was recently described. We now report that the α chain occurs in two allelic forms distinguished by charge difference. Using inbred recombinant mouse strains we were able to map the *H-2Oa* gene to a location between the A.TL and B10.MBR recombination points. Cosmids covering this region were used to isolate the gene. Sequence analysis revealed that the *H-2Oa* gene is the murine equivalent of the human *HLA-DNA* gene. These results indicate that the human *HLA-DNA* gene, the existence of which has long been known, is indeed coding for DOα, the α chain pairing with DOβ.

H-2A and H-2E are highly polymorphic cell surface glycoproteins that have a critical role in antigen presentation to T lymphocytes. The genes encoding these heterodimeric molecules are arranged pairwise and are located in the class II region of the MHC (1). This genetic region also contains several other genes, some of which have homology with the H-2A and E genes (2–6). A novel class II molecule, H-2O, whose β chain is encoded in the MHC region, was recently described (7). In contrast to H-2A and H-2E, which are expressed by B cells, macrophages, dendritic cells, and epithelial cells of the thymic cortex and medulla, H-2O is expressed only on B cells and on thymic medullary epithelium. We now report the structure and location in the MHC region of the *H-2Oa* gene, which like the *H-2Ob* gene, is largely nonpolymorphic. The sequence of the novel α chain gene revealed that it is the homologue of the human *HLA-DNA* gene, whose role has remained enigmatic for some time.

Materials and Methods

Animals. C57BL/6 (B6), B10.M, CBA/J, BALB/c, DBA/1, SJL, (B6 × CBA/J)F_1, and B10.A(5R) mice were obtained from The Scripps Research Institute breeding colony, as were New Zealand White rabbits. A.TL mice were obtained from The Jackson Laboratory (Bar Harbor, ME), and B10.AQR and B10.MBR mice were obtained from the Department of Genetics, Washington University, St. Louis, MI.

Antibodies. BSA (Sigma Chemical Co., St. Louis, MO) was conjugated with synthetic peptides, using m-Maleimidobenzoyl-*N*-hydroxysuccinimide ester (MBS) (Pierce Chemical Co., Rockford, IL) as described (8). Rabbits were bled and immunized with 500 μg BSA conjugate in Freunds adjuvant (Gibco Laboratories, Grand Island, NY) every 2–3 wk. The K545 rabbit antiserum, reacting with H-2Oα, was raised against the BSA-conjugated peptide TGTRRPSIRR, (single-letter amino acid code). The resulting serum did not crossreact with cytoplasmic tail peptides from H-2Aα, H-2Aβ, H-2Eα, H-2Eβ, H-2Mα, H-2Mβ, or H-2Oβ in ELISA assays. In immunoprecipitation analyses of molecules from transfected HeLa cells expressing either H-2Oα or H-2Oβ, the K545 antiserum only reacted with H-2Oα. The K507 antiserum has previously been shown to react only with the H-2Oβ chain (7).

Cell Labeling and Immunoprecipitation. Cell labeling, lysis, immunoprecipitation, and two-dimensional IEF gels were done essentially as described by P. Jones (9), except that both [^{35}S]cysteine and [^{35}S]methionine were used. In summary, splenocytes were cleared of red blood cells by osmotic lysis and labeled in methionine- and cysteine-free RPMI 1640 (Biofluids Inc., Rockville, MD) with 0.5 mCi [^{35}S]cysteine and 0.5 mCi [^{35}S]methionine for the time indicated in the figure legends. Cells were washed twice with PBS and lysed in 1% Triton X-100 with 2 mM PMSF in PBS. After preclearing with protein A-Sepharose (Pharmacia LKB Biotechnology Inc., Piscataway, NJ), lysates were incubated with the specific antisera. Immunoprecipitates were harvested with protein A-Sepharose, washed, and resuspended in IEF sample buffer before being subjected to two-dimensional gel electrophoresis. The first dimension IEF was done using ampholines pH 5–7 (Pharmacia LKB Biotechnology Inc.). Second-dimension slab gels were 7.5–12.5% polyacrylamide. After electrophoresis, gels were fixed and treated with Amplify (Amersham Corp., Arlington Heights, IL) before drying and autoradiography.

DNA Techniques. Two degenerate oligonucleotides (5′-CCC-AACACCCTTATCTG-3′ and 5′-TCCACTTTGCAGTCATA-3′)

GT CG T CCGA

C

were used to amplify DNA from cosmids under the following conditions: 30 cycles of 92° 45 s, 49° 1 min, and 72° 45 s. The PCR

 J. Exp. Med. © The Rockefeller University Press • 0022-1007/92/08/0477/07 $2.00
Volume 176 August 1992 477–483

product from cosmid II 2.27A (2.27α) was cloned into Bluescript KS+ (Stratagene Inc., La Jolla, CA) and sequenced (10). The sequence was used to construct two pairs of oligonucleotides. One pair (5'-TCCCGCCTGTGATCAAT-3' and 5'-GAGTGGCCCA-GACCAGC-3') was derived from the coding strand, and the other pair (5'-GAAGCTGGTCTGGGCCA-3' and 5'-TGGGCTGGC-TGTTGCGC-3') was complementary to the coding strand. These two primer sets, together with vector-derived primers (5'-CAG-ACCAACTGGTAATGGTA-3' and 5'-CCTGGAGCCCGTCAG-TAT-3', respectively) were used in subsequent PCR reactions to amplify H-2Oa DNA from a B10.M splenocyte cDNA library cloned into λgt11 (11). Amplified material was cloned into Bluescript KS+ and sequenced using internal primers. The sequences from different mouse strains (including, as a control, B10.M) were determined as follows: total RNA was prepared and converted to first strand cDNA according to a standard protocol (12). The first strands were used as templates in PCR reactions with primers derived from the 5' and 3' untranslated sequences. In a second round of amplification, a single nested biotinylated primer was used, either in the 5' or in the 3' end. Noncloned sequencing templates representing the entire open reading frame were prepared as described (13) and sequenced according to a standard protocol (10), using H-2Oa-specific sequencing primers.

To determine the gene structure cosmid II 2.27A was digested with restriction enzymes, blotted onto nylon filters, and probed with the 2.27α fragment. The hybridizing 3.5-kb HindIII and 1.1-kb KpnI fragments were cloned. PCR with primers derived from the untranslated parts of the cDNA sequence showed that the entire coding sequence was contained in the HindIII clone. PCR with cDNA-derived primers (see Fig. 2 *B*) were used to determine the lengths of the introns. Exon-intron boundaries were sequenced using cDNA-derived primers.

Results and Discussion

The coding sequence of the *H-2Ob* gene displays very limited polymorphism, and most substitutions are silent. Analysis of immunoprecipitated H-2O molecules by two-dimensional gel electrophoresis indicated that the degree of polymorphism is also low in the H-2O α chain, which is coprecipitated with the H-2O β chain (7). However, since we wished to molecularly clone the H-2O α chain gene, we tried to identify a charge polymorphism in this chain to trace its location in the genome. Further analyses by two-dimensional gel electrophoresis of H-2O molecules from several mouse strains confirmed the low degree of α chain polymorphism, but one case of allelic polymorphism was identified by two-dimensional gel electrophoresis. Thus, H-2O molecules from the CBA/J strain expressed a H-2O α chain that was slightly more basic than the α chains of other mouse strains (Fig. 1; compare α chain spots from B10.M and CBA/J mice). Since the H-2A and E loci contain genes encoding both the α and the β chains, and since the H-2O β chain is encoded in the MHC region, it seemed reasonable to assume that the H-2O α chain would also be located in this region. This assumption was supported by the observation that the CBA/J and B10.BR strains shared allelic forms of the H-2O α chain (not shown). Since these mouse strains have identical MHC regions (the $H\text{-}2^k$ haplotype) but differ in a number of other loci, we immunoprecipitated H-2O molecules from several inbred, recombinant mouse strains, where intra-MHC crossovers between the $H\text{-}2^k$ haplotype and other H-2 haplotypes have occurred, and subjected the molecules to two-dimensional gel electrophoresis. The results of these analyses are summarized in Fig. 1. It can be seen that the H-2O α chain spots, which typify the $H\text{-}2^k$ haplotype (CBA/J), were present only when the H-2O molecules were derived from the A.TL recombinant strain. Thus, the other recombinant strains, B10.A(5R), B10.AQR, and B10.MBR displayed H-2O α chains of the non-$H\text{-}2^k$ allelic form (compare to H-2O α chain spots of the B10.M strain ($H\text{-}2^f$) in Fig. 1).

The H-2 haplotypes of the mouse strains used in the genetic analyses are depicted in Fig. 2 *A* (1, 14). The intra-H-2 recombination points are indicated and they show that the H-2O α chain must be encoded in the region between the recombination points in B10.MBR and A.TL. This is the only interpretation consistent with the observation that the A.TL strain, but no other recombinant strain expressed the $H\text{-}2^k$ form of the H-2O α chain. This information is surprising inasmuch as all expressed class II loci contain α and β chain genes in juxtaposition. The genetic analyses revealed that the *H-2Oa* gene must reside centromeric of the *H-2Ob* gene, and that it is separated from this gene by some 150 kb of DNA. In fact, two clusters of recently described genes (4–6), four of which are involved in antigen processing and transport of peptides for class I molecules (15, 16), are interspersed between the two H-2O genes.

The DNA segment separating the recombination points in the A.TL and B10.MBR strains was available in four overlapping cosmid clones (14) (Fig. 2 *B*). To molecularly clone the H-2O α chain we used these cosmids and adopted a PCR strategy. Conserved regions in the third exon of previously identified class II α chain genes were identified, and a set of degenerate primers were synthesized. In a PCR reaction these primers should be able to amplify a 193-bp fragment from all known α chain genes. Cosmid II 2.27A (Fig. 2 *B*) was the only one that gave rise to a fragment of the expected size. The amplified fragment, 2.27α, was cloned and sequenced and revealed striking homology to other α chain genes. To obtain the full-length sequence of the putative H-2Oα chain, we used nested sets of oligonucleotides derived from the sequence of the 2.27α fragment together with vector sequence-derived oligonucleotides to amplify material from a B10.M spleen cell cDNA library. This strategy allowed us to obtain the complete cDNA sequence of the putative H-2Oα chain. The exact location of the novel α chain gene was determined by Southern blot analysis of restriction enzyme-digested II 2.27A DNA, using the 2.27α fragment as a probe. This strategy allowed us to identify and clone a 3.5-kb HindIII fragment and an overlapping 1.1-kb KpnI fragment. PCR analyses of the clones showed that the HindIII clone contained the entire translated gene sequence. The KpnI clone contained most of the 3' untranslated sequence. (Fig. 2 *B*). By comparing the restriction maps of II 2.27A with those of the two subclones, the gene orientation could be determined. PCR with primers corresponding to the cDNA se-

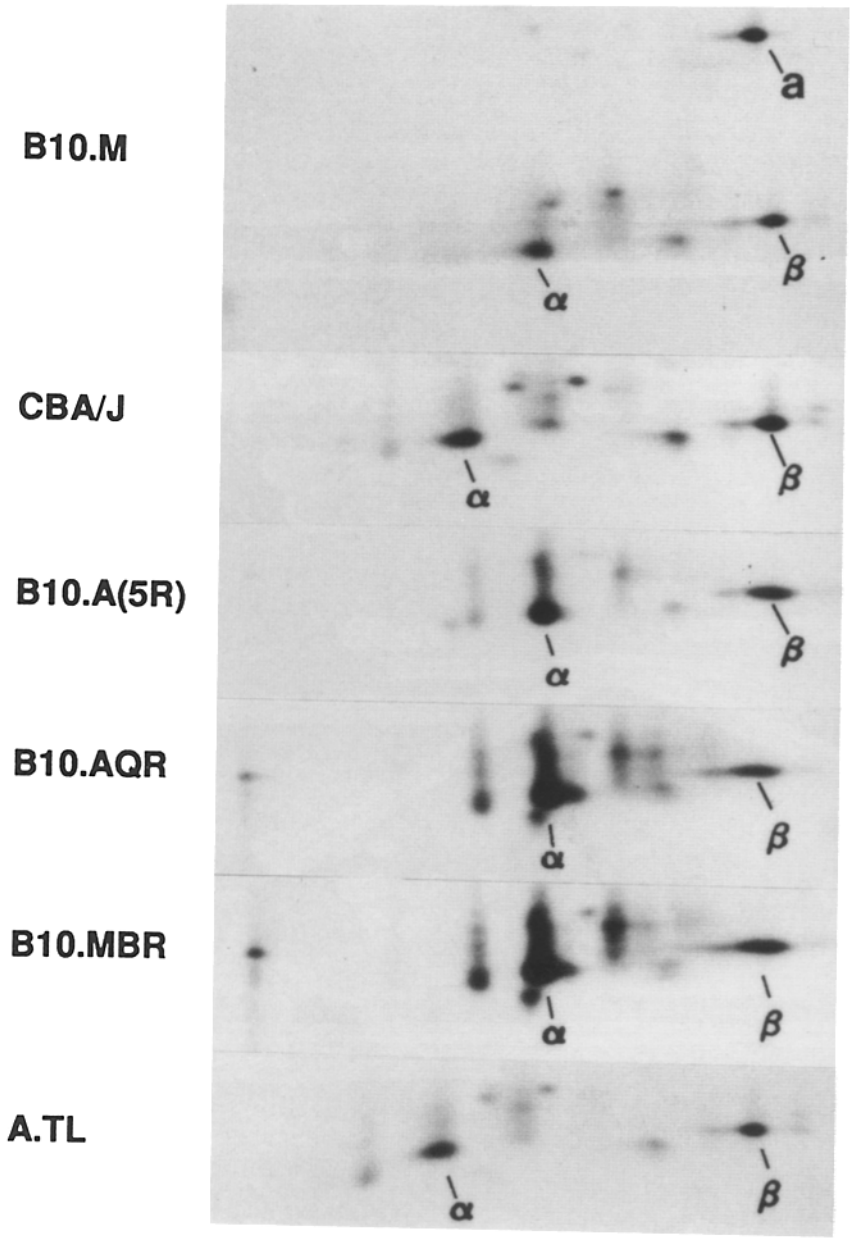
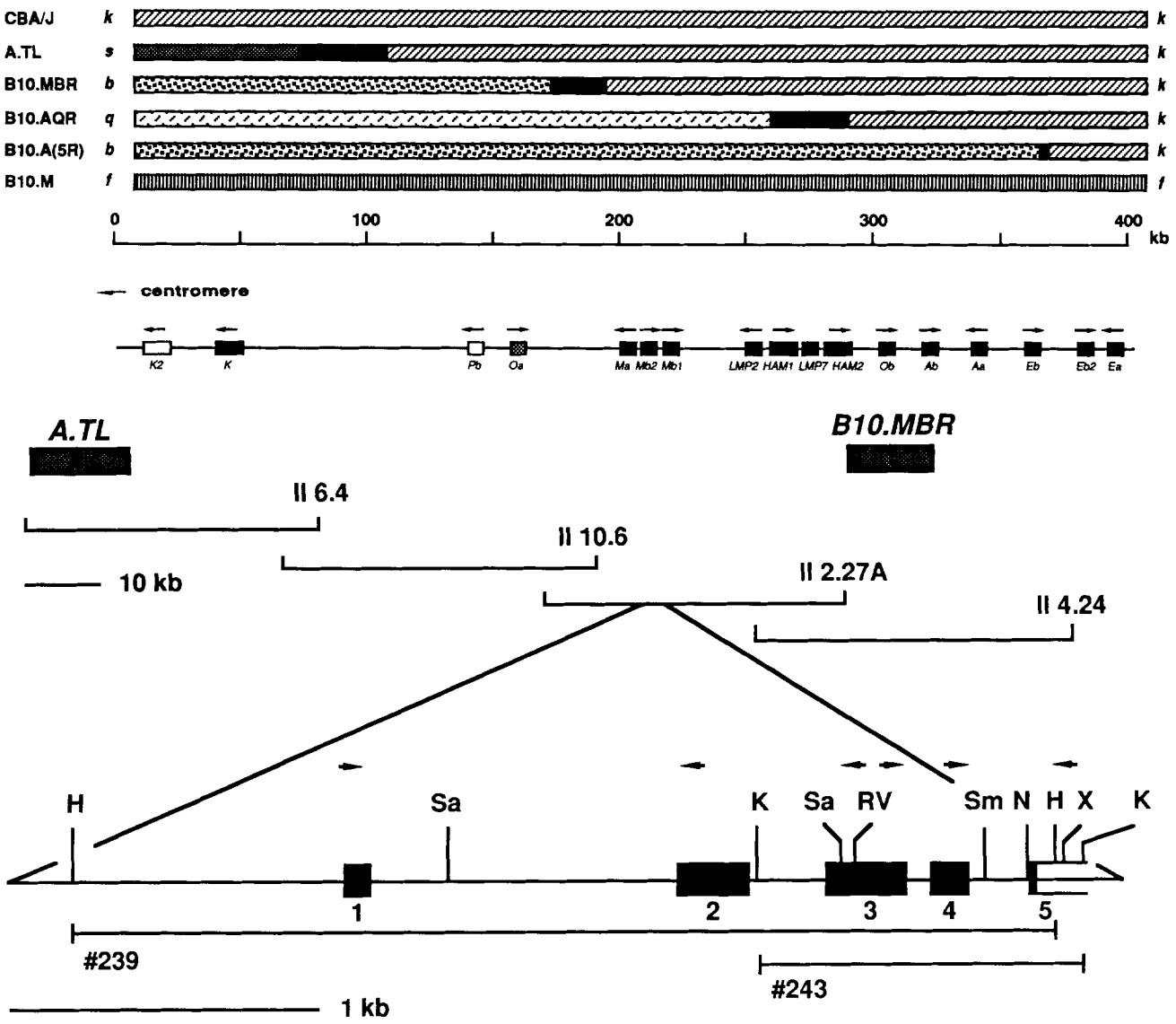


Figure 1. The isoelectric point of H-2Oα is haplotype dependent. Splenocytes of the indicated mouse strains were metabolically labeled for 4 h. H-2O molecules were precipitated with K507 (anti–H-2Oβ) and separated by two-dimensional gel electrophoresis. (α and β) Mature forms of the H-2O chains. Immature forms are not indicated. (*a*) Actin.

quence was used to determine the exon–intron organization of the gene, and the exact boundaries were determined by sequencing. The genomic organization resembles that of other α chain genes in general, and that of the human *HLA-DNA* in particular (17). The location in the class II region is similar to that of *HLA-DNA*, and the existence of a murine homologue of *HLA-DNA* has recently been implicated by cross-hybridization (18).

The open reading frame encoded a typical class II α chain, with a predicted disulfide bridge in the second domain and two glycosylation sites in positions identical to those of other α chains (except H-2Mα [6]). Sequence comparisons between the novel chain and previously identified α chains, revealed that the putative H-2O α chain did indeed most closely resemble the human HLA-DNα chain (Table 1). This fact, in conjunction with the observation that the *HLA-DNA* gene occupies the same relative position in the human MHC region as does the *H-2Oa* gene in the murine MHC, suggests that the DNα chain may be pairing with the human DOβ chain.

The genetic analyses described above predicted that if the gene encoded in cosmid II 2.27A was indeed the *H-2Oa* gene, its sequence from the H-2^k haplotype should reveal a more basic net charge than the sequences of the other haplotypes. To examine this, we sequenced the coding region of the novel gene from five additional haplotypes. Table 2 summarizes the



few differences observed between the listed haplotypes and the H-2^f sequence. Apart from differences in the signal sequences, only two replacement substitutions were observed in the coding regions and only one involved a charged residue. Thus, the novel gene encoded an alanine rather than a glutamic acid in the H-2^k allele, which gave a calculated pI of 6.93 for the H-2^k chain rather than 6.71 for the chains of the other haplotypes. Thus, the more basic character of the H-2^k allelic form is fully consistent with the two-dimensional gel electrophoretic analyses shown in Fig. 1.

Further evidence for the novel gene encoding the H-2O α chain was obtained by immunoprecipitation analyses. A rabbit antiserum, K545, raised against a peptide corresponding to the cytoplasmic tail of the novel chain, was used in immunoprecipitation analyses of metabolically labeled proteins of C57BL/6 (B6), CBA/J and (B6 × CBA/J)F$_1$ splenocytes. The two-dimensional gel patterns of polypeptides immunoprecipitated with the antibodies against the α chain was compared with the patterns obtained with an antiserum against the H-2O β chain (Fig. 3). Although the intensities of the spots varied, depending upon which antiserum was used, their locations were identical. Accordingly, it can be concluded that the antibodies against the novel α chain coprecipitates the H-2O β chain.

The H-2O molecule is different from conventional class II molecules because of its nonpolymorphic nature and its restricted tissue distribution. The present observation that the two H-2O genes, like their human counterparts *HLA-DNA* and *HLA-DOB*, are separated by a cluster of genes that are involved in antigen processing and peptide transport for class I molecules, raises the possibility that the H-2O genes are more ancient than the class I-related cluster of genes, which must have been introduced into the MHC region before speciation. Nonetheless, the occurrence of H-2O genes in all species examined (17, 19–22) strongly argues for these genes having remained functional during evolution.

Figure 2. Schematic picture of the murine MHC class II region together with the location, genomic structure, and cDNA sequence of H-2Oa. (*A*) The genes of the murine class II region. Recombination regions of mouse strains used in Fig. 1 are indicated above (14). Letters flanking the "recombination bars" indicate the MHC haplotypes. (*Arrows*) Direction of gene transcription. (*Solid boxes*) Functional genes; (*open boxes*) pseudogenes (1, 4–6), (K. Früh, unpublished data). (*Dotted*) Oa, the gene described. (*B*) Four cosmids overlap the region between the A.TL and B10.MBR recombination regions (14). The *H-2Oa* gene is located in cosmid II 2.27A. #239 is the 3.5-kb HindIII clone, #243 is the 1.1-kb KpnI clone. (*Arrows*) Locations of the PCR primers used for the gene structure determination. (*Boxes*) Exons numbered 1–5. Restriction enzymes are abbreviated as follows: H, HindIII; Sa, SacI; K, KpnI; RV, EcoRV; Sm, SmaI; N, NruI; and X, XhoI. (*C*) Nucleotide sequence of H-2Oa^f (from B10.M) with the translation product indicated below. +1 indicates the first amino acid of the mature protein. (*Boxed*) Sites for Asn-linked carbohydrate addition. (*Underlined*) The predicted transmembrane region. (*Arrows*) Exon–intron boundaries. (*Double underlined*) The potential poly(A) addition signal. These sequence data are available from EMBL/GenBank/DDBJ under accession number M95514.

Table 1. *Homology between H-2Oα and Other Class II α Chains*

	Mα	Aα	Eα	DPα	DNα	DMα	DQα	DRα
Oα	27.4	51.6	47.8	52.7	**76.4**	28.2	52.7	52.2
Mα	–	30.7	25.7	26.6	30.5	**74.5**	26.2	27.8
Aα		–	53.3	61.1	57.8	33.2	**73.2**	57.2
Eα				51.8	51.8	27.0	51.8	**77.3**
DPα				–	58.0	29.7	61.4	62.0
DNα					–	28.7	59.4	55.4
DMα						–	30.8	28.9
DQα							–	56.6

Figures indicate percent identity of the mature protein sequences (i.e., not including the signal sequences), calculated using the program GAP (23). The highest homologies are indicated with bold text. The sequences used for comparison are for Oα, Aα (24), and Eα (25) of the k-haplotype. For DPα the sequence used is DPα1*0103 (26), and for DQα the sequence used is DQα1*0501 (27).

Table 2. *The Allelic Polymorphism in the H-2Oa Coding Sequence is Limited*

Strain	Haplotype	Base changes		Amino-acid changes	
C57BL/6	*b*	41	T→C	–12	Val→Ala
		647	C→T	191	Thr→Ile
BALB/c	*d*	324	T→C		– –
		351	T→G		– –
		647	C→T	191	Thr→Ile
DBA/1	*q*	135	T→C		– –
SJL	*s*	28	G→A	–16	Val→Ile
		129	C→T		– –
CBA/J	*k*	26–34	Deleted	–17 – –15	Deleted
		197	A→C	41	**Glu→Ala**
		201	C→T		– –
		647	C→T	191	Thr→Ile

The sequences were compared to the H-2Oa^f sequence (See Fig. 2 *c*). Base 1 is the A in the translation initiation codon (ATG).

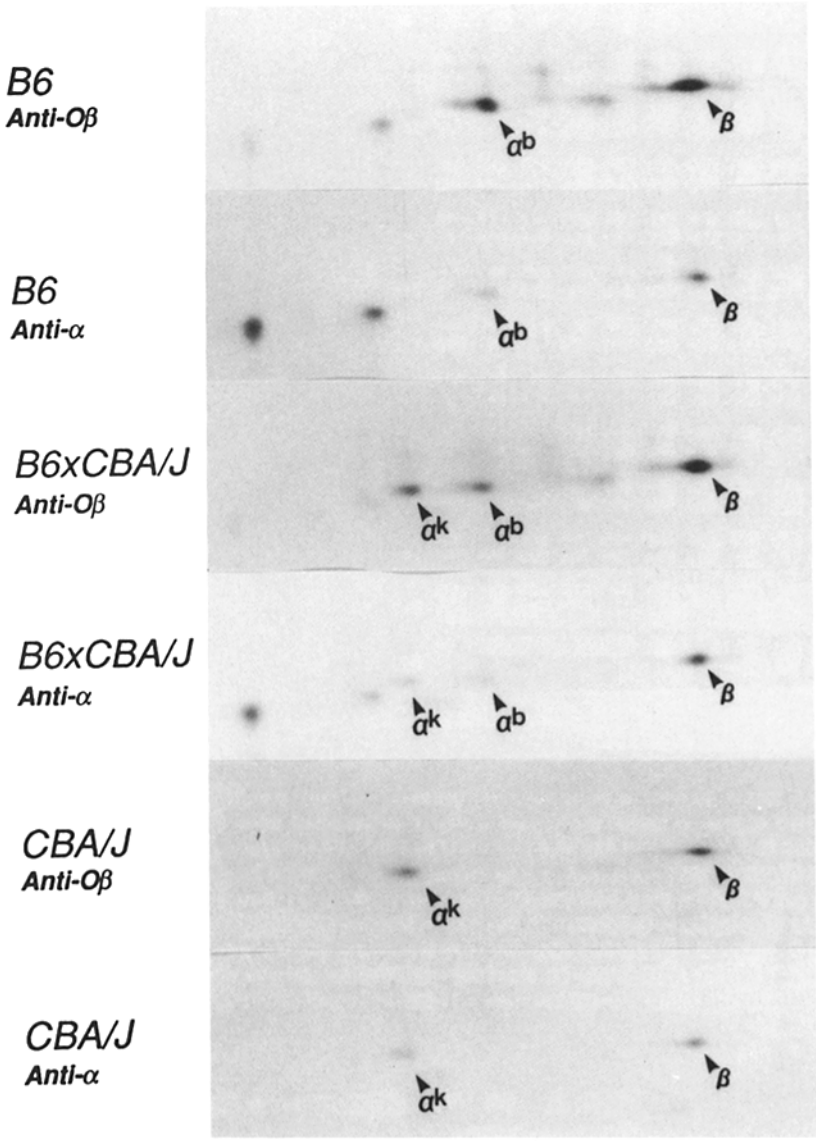


Figure 3. Coprecipitation of H-2Oα with H-2Oβ. C57BL/6 (B6), (B6 × CBA/J)F_1 and CBA/J splenocytes were labeled for 3 h. Lysates were split and immunoprecipitated with either K507 (anti-H-2Oβ) or K545, directed against the cytoplasmic tail of the new α chain. α^b and α^k indicate the mature forms of H-2Oα^b and H-2Oα^k, respectively. β indicates the mature form of H-2Oβ. Immature forms are not indicated.

We thank C. Surh and J. Sprent for discussions; Y. Uematsu and M. Steinmetz for MHC cosmids; J. Chambers for aid in DNA sequencing; and D. Uranowski for secretarial assistance.

This work was supported by National Institutes of Health grant AI-26610 (P. A. Peterson), and a Cancer Research Institute/F.M. Kirby Foundation Fellowship (L. Karlsson).

Address correspondence to Dr. Lars Karlsson, Department of Immunology, IMM-8, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037.

Received for publication 6 April 1992.

References

1. Klein, J. 1986. Natural History of the Major Histocompatibility Complex. John Wiley & Sons, Inc., New York. 775 pp.
2. Braunstein, N.S., and R.N. Germain. 1986. The mouse Eβ2 gene: a class II MHC β gene with limited intraspecies polymorphism and an unusual pattern of transcription. *EMBO (Eur. Mol. Biol. Organ.) J.* 5:2469.

3. Widera, G., and R.A. Flavell. 1985. The I-region of the C57BL/10 mouse: characterization and physical linkage to H-2K of a novel SBβ-like pseudogene *Aβ3. *Proc. Natl. Acad. Sci. USA.* 82:5500.
4. Monaco, J.J., S. Cho, and M. Attaya. 1990. Transport protein genes in the murine MHC: possible implications for antigen processing. *Science (Wash. DC).* 250:1723.
5. Martinez, C.K., and J.J. Monaco. 1991. Homology of proteasome subunits to major histocompatibility-linked LMP gene. *Nature (Lond.).* 353:664.
6. Cho, S., M. Attaya, and J.J. Monaco. 1991. New class II-like genes in the murine MHC. *Nature (Lond.).* 353:573.
7. Karlsson, L., C.D. Suhr, J. Sprent, and P.A. Peterson. 1991. A novel class II MHC molecule with unusual tissue distribution. *Nature (Lond.).* 351:485.
8. Harlow, E., and D. Lane. 1988. Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 726 pp.
9. Jones, P.P. 1980. Analysis of radiolabeled lymphocyte proteins by one- and two-dimensional gel electrophoresis. In Selected Methods in Cellular Immunology. B.P. Mishell, and S.P. Shigii, editors. Freeman Publications, San Francisco. 398–440.
10. Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463.
11. Young, R.A., and R.W. Davis. 1983. Efficient isolation of genes by using antibody probes. *Proc. Natl. Acad. Sci. USA.* 80:1194.
12. Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1988. Current Protocols in Molecular Biology. John Wiley & Sons, Inc., New York.
13. Hultman, T., S. Bergh, T. Moks, and M. Uhlén. 1991. Bidirectional solid-phase sequencing of in vitro-amplified plasmid DNA. *Biotechniques.* 10:84.
14. Steinmetz, M., D. Stephan, and K. Fisher-Lindahl. 1986. Gene organization and recombinational hotspots in the murine major histocompatibility complex. *Cell.* 44:895.
15. Powis, S.J., A.R.M. Townsend, E.V. Deverson, J. Bastin, G.W. Butcher, and J.C. Howard. 1991. Restoration of antigen presentation to the mutant cell line RMA-S by an MHC-linked transporter. *Nature (Lond.).* 354:528.
16. Attaya, M., S. Jameson, C.K. Martinez, E. Hermel, C. Aldrich, J. Forman, K. Fisher Lindahl, M.J. Bevan, and J.J. Monaco. 1992. Ham-2 corrects the class I antigen-processing defect in RMA-S cells. *Nature (Lond.).* 355:647.
17. Trowsdale, J., and A. Kelly. 1985. The human HLA class II α chain gene DZα is distinct from genes in the DP, DQ, and DR subregions. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2231.
18. Hanson, I.M., and J. Trowsdale. 1991. Colinearity of novel genes in the class II regions of the MHC in mouse and human. *Immunogenetics.* 34:5.
19. Andersson, L., and L. Rask. 1988. Characterization of the MHC class II region in cattle. The number of DQ genes varies between haplotypes. *Immunogenetics.* 27:110.
20. Sachs, D.H., S. Germana, M. El-Gamil, K. Gustafsson, F. Hirsch, and K. Pratt. 1988. Class II genes of miniature swine. *Immunogenetics.* 28:22.
21. Schøller, J., and Å Lernmark. 1985. Isolation of a Rat Immune Response gene Identical to an Alleged Mouse a Class II β-Chain Pseudogene. *Immunogenetics.* 22:601.
22. Tonelle, C., R. DeMars, and E.O. Long. 1985. DOβ: A new β chain gene in HLA-D with a distinct regulation of expression. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:2839.
23. Devereux, J.H., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids. Res.* 12:387.
24. Benoist, C.O., D.J. Mathis, M.R. Katner, V.E. Williams II, and H.O. McDevitt. 1983. Regions of allelic hypervariability in the murine Aα immune response gene. *Cell.* 34:169.
25. Mathis, D.J., C.O. Benoist, V.E. Williams II, M.R. Kanter, and H.O. McDevitt. 1983. The murine Eα immune response gene. *Cell.* 32:745.
26. Gustafsson, K., E. Widmark, A.-K. Jonson, B. Servenius, D.H. Sachs, D. Larhammar, L. Rask, and P.A. Peterson. 1987. Class II genes of the human major histocompatibility complex. Evolution of the DP region as deduced from nucleotide sequences of the four genes. *J. Biol. Chem.* 262:8778.
27. Schenning, L., D. Larhammar, P. Bill, K. Wiman, A.-K. Jonsson, L. Rask, and P.A. Peterson. 1984. Both α and β chains of HLA-DC class II histocompatibility antigens display extensive polymorphism in their amino-terminal domains. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:447.