Biochemical comparison of major histocompatibility complex molecules from different subspecies of *Mus musculus*: Evidence for *trans*-specific evolution of alleles

(polymorphism/H-2 complex/peptide mapping)

BERNHARD ARDEN AND JAN KLEIN*

Max-Planck-Institut für Biologie, Abteilung Immungenetik, Corrensstr. 42, 7400 Tübingen, Federal Republic of Germany

Communicated by Jack L. Strominger, December 18, 1981

ABSTRACT H-2 haplotypes were extracted from wild mice of three subspecies, Mus musculus domesticus, M. m. molossinus, and M. m. castaneus, that are known to have been separated from one another for some 1 to 2 million years. Serologically indistinguishable molecules controlled by some of the polymorphic H-2 loci were compared by tryptic peptide mapping, and the maps were found to be identical. In addition, a number of instances of biochemically indistinguishable H-2 molecules were found among wild mice and inbred strains of the M. m. domesticus subspecies. These findings suggest that some of the H-2 alleles have not altered for >1 million years. To reconcile this apparent stability of H-2 genes with their extraordinary polymorphism (some 100 alleles at each of the polymorphic H-2 loci), it is proposed that the H-2 alleles evolve as if they were separate loci.

To a geneticist, the H-2 complex, the major histocompatibility complex (Mhc) of the mouse, presents two puzzling features. First, some of the H-2 loci display a degree of polymorphism not found at any other locus (1). Some 50 H-2K and 30 H-2D alleles have already been identified among wild mice and the actual number may be double this figure. Furthermore, in most populations, there does not appear to be any high-frequency common allele, as in most other genetic systems; instead, the populations contain a large number of alleles occurring at relatively low frequencies (1). Second, when one compares two Kor two D alleles biochemically, one can estimate that they probably differ at 40 or more of their amino acids (2-5). This large number of differences has again not been found at any other locus, with the possible exception of certain immunoglobulin loci (6); the usual situation is that polypeptide chains produced by allelic genes differ in one or two amino acids.

These findings raise several questions, among which not the least important is, how has the polymorphism arisen? Three principal answers to this question are possible. First, the H-2 alleles diversify much more rapidly than alleles at most other loci and the \approx 40 amino acid substitutions differentiating two H-2 alleles have accumulated since the time the mouse became established as a species (7). Second, what we regard as alleles are in fact pseudoalleles at distinct but closely linked and closely homologous loci that, however, evolve at the same pace as many other loci (8). Third, the Mhc alleles evolve more or less independently of speciation and are therefore much older than the species (9). In this communication, we provide evidence that argues against the first possibility and can be interpreted as favoring the third possibility.

MATERIALS AND METHODS

Mice. The critical strains were B10.MOL1, B10.CAS2, and STU. The first was derived from a cross between B10.BR (H-

 2^{k}) laboratory mice and wild mice belonging to the subspecies M. musculus molossinus. The wild mice were captured in Japan and reached us via M. Potter, National Institutes of Health, Bethesda, MD. From the original cross, a congeneic line was produced by repeated backcrossing to mice carrying the genetic background of strain C57BL/10Sn (B10) and selection, in each generation, for the $H-2^{wl2}$ haplotype of the wild mouse (10). The B10.CAS2 line was derived in a similar manner from a wild mouse representative of the subspecies M.m. castaneus; it now carries the $H-2^{w17}$ haplotype of this mouse on the C57BL/10Sn background (10). The STU strain was produced by W. Schäfer, Max-Planck-Institut für Virusforschung, Tübingen, from noninbred European albino mice of uncertain ancestry (11). Lines B10.GAA37, B10.KPB128, B10.CHA2, B10.STC77, and B10. SAA48 carry H-2 haplotypes extracted from Michigan wild mice; the W12A strain carries the H-2 haplotype of a wild mouse captured in Texas (12). Other mice used in this study belonged to common inbred and H-2 recombinant strains.

Antisera. A list of antisera used is given in Table 1. The antisera were produced by immunization of mice with lymphoid cells of appropriate donors, as described elsewhere (10).

Biochemical Methods. These are described in detail elsewhere (13). Briefly, splenocytes were labeled by incubating them with ³H- or ¹⁴C-labeled arginine, lysine, leucine, and tyrosine. The radiolabeled cells were solubilized with Nonidet P-40, the extracts were passed through an affinity column containing Lens culinaris hemagglutinin conjugated to Sepharose 4B, and the glycoprotein-containing fraction was eluted with α methyl-D-mannoside. H-2 molecules were precipitated from the glycoprotein pool by specific alloantisera (Table 1), and the samples were boiled with 5% 2-mercaptoethanol and isolated by discontinuous NaDodSO4/polyacrylamide gel electrophoresis. The samples were dissolved in NH4HCO3 and digested with N-tosylphenylalanine chloromethyl ketone-treated trypsin, and the peptides were lyophilized. Before use, the lyophilized samples were dissolved in pyridine acetate and the insoluble material was removed by centrifugation. Soluble peptides were loaded onto a cation-exchange column containing Technicon Chromobeads type P equilibrated in pyridine acetate. Peptides were eluted by using a linear pH-ionic strength gradient and the fractions were collected into scintillation vials. After evaporation to dryness, the samples were dissolved in NaDodSO₄ solution and assayed in Beckman HP cocktail in a Beckman LS7500 scintillation counter. The ³H and ¹⁴C radioactivities of each fraction were recorded in MACRO-11 on a file and from there were corrected for channel spillover and constant background and plotted by a FORTRAN program on a PDP11.

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.

Abbreviation: Mhc, major histocompatibility complex.

^{*} To whom reprint requests should be addressed.

Code	Recipient	Donor	<i>H-2</i> haplotype	Class I product detected
K-323	$(DBA/2 \times B10.RIII) F_1$	B10.A	(d/r)a	K ^k
K-26	$(C57BL/10 \times A) F_1$	AQR	(b/a)y1	Kq
K-548	$(B10.A \times A) F_1$	B10.D2	(a/a)d	K ^d
T-35	$(B10.D2 \times CBA) F_1$	B10.PL	(d/k)u	Ku
T-21	$(\mathbf{A} \times \mathbf{B10}) \mathbf{F}_1$	B10.NZW	(a/b)z	K ^u , D ^{w14}
G-25	$(C3H \times B10.D2) F_1$	STU	(k/d)w34	D ^{w3}
T-3	$(R107 \times C3H) F_1$	B10.A(4R)	(i7/k)h4	\mathbf{D}^{b}

RESULTS

Rationale of the Experiment. We have demonstrated previously that H-2 alleles found in inbred strains also occur in wild mice and that the molecules encoded in these alleles are indistinguishable serologically and biochemically. In those studies, however, we could not estimate how long the alleles had been separated and hence we could only very tentatively argue that H-2 alleles were not diversifying at a rapid rate. In this study, we compared H-2 alleles for which the minimal time of separation is known. The alleles were extracted from three subspecies of the house mouse Mus musculus: M. m. domesticus, represented by inbred strains and by wild mice from the northern part of the United States, and M. m. molossinus and M. m. castaneus, represented by the congeneic lines B10.MOL1 and B10.CAS2, respectively. Restriction enzyme mapping of mtDNA isolated from these mice did not show any differences among mice within the same subspecies (for example, between inbred and wild mice of M. m. domesticus) but did show substantial differences between subspecies (14). Yonekawa et al. (14) have found the following numbers of common/different cleavage sites between M.m. castaneus and M.m. domesticus for various restriction enzymes: BamHI, 2/4; EcoRI, 2/2; HindII, 3/3; HindIII, 2/1; Hpa I, 2/2; Pst I, 1/1. For M. m. molossinus and M. m. domesticus, the values are BamHI, 2/2; EcoRI, 2/2; HindII, 5/0; HindIII, 2/1; Hpa I, 3/0; Pst I, 1/1. From these data, Yonekawa et al. estimated that the mtDNA of M. m. castaneus and M.m. domesticus differ in 7.2% of their nucleotides; the corresponding value for M. m. molossinus and M. m. domesticus is 3.1%. They calculated that M.m. domesticus has been separated from M.m. molossinus and M.m. castaneus for



FIG. 1. Paired-label comparative peptide maps of H-2K polypeptides. Tryptic peptides from ¹⁴C-labeled H-2K^k antigens (---) of B10.A vs. ³H-labeled H-2K antigens (---) of CBA(M523) (A) or ³H-labeled H-2K antigens (---) of B10.MOL1 (B) were analyzed by ion-exchange chromatography. \blacklozenge , Distinct parental peptide; \diamondsuit , peptide obtained in higher yield in the mutant. ---, pH.



FIG. 2. Comparative peptide maps of H-2D polypeptides. Tryptic peptides from ¹⁴C-labeled H-2D antigens (---) of STU vs. ³H-labeled H-2D antigens (---) of B10.SAA48 (A) or ³H-labeled H-2D antigens (---) of B10.CAS2 (B) were analyzed by ion-exchange chromatography. ---, pH.

1 to 2 million years. Hence, if one were to find identical H-2 alleles in these subspecies, one could argue that these alleles have not altered for this length of time.

Tryptic Peptide Map Comparisons of H-2K Molecules Extracted from B10.A, CBA(M523), and B10.MOL1 Mice. The B10. MOL1 line codes for a K molecule that is serologically indistinguishable from the K^k molecule of the B10.A strain. To determine whether the two molecules are also identical biochemically, we compared them by tryptic peptide mapping. In contrast to previous studies (13) in which peptides labeled in either arginine or lysine residues were analyzed separately, all products were labeled with an arginine/lysine/leucine/tyrosine mixture and compared in a single chromatography experiment. It had been shown previously that some low-yield tryptic peptides are particularly well labeled by tyrosine but not by arginine or lysine (15). In this way, the numbers of peptide maps necessary for screening wild alleles could be reduced and the vields, especially of lysine peptides, could be improved. To test the resolution of our technique, we compared peptide maps of the K^k molecule from CBA mice and the K^{km^1} molecule from CBA(M523) mice, which had previously been shown to differ from K^k by a single peptide (16). ¹⁴C-Labeled H-2K^k polypeptides from B10. A mice were mixed with ³H-labeled K^{km1} polypeptides from M523 mice and digested with trypsin, and the

resulting peptides were analyzed by ion-exchange chromatography. From the H-2K^k molecule from B10.A mice, 26 major peptide peaks were eluted from the column (Fig. 1A). This number is the sum of 15 arginine and 11 lysine peaks previously found in tryptic peptide maps of B10. A mice (13), indicating that the same number of peptides has been resolved in this map. Of course, it is possible that some of these, as well as the previously defined peaks, may contain more than one peptide. From the K^{km1} molecule of CBA(M523) mice, all of the peaks eluted at the same position, with one exception: the peak at fraction 254 (pH 4.9) was missing in the mutant. This result is consistent with the absence of one lysine peptide at pH 4.7 observed previously (16). The quantitative difference in peak height of fraction 131 is caused by the low yield of the ¹⁴C-labeled peptide in this particular chromatogram. Comparison with another chromatogram (Fig. 1B) shows that this peptide is present in the $H-2K^{k}$ molecule of B10.A mice. Differences in the first 50 fractions are neglected, because the peptides were not reproducibly eluted in this range of the gradient. In addition, there was a shoulder at fraction 165, missing in the profile from B10. A mice that might represent a difference in a low-yield peptide.

To determine the biochemical similarity of K^{k} molecules from B10.A and B10.MOL1 mice, these molecules were precipitated with antiserum K-323 and compared by tryptic pep-

Evolution: Arden and Klein

tide mapping (Fig. 1*B*). The ¹⁴C- and ³H-labeled peptides eluted identically. One minor peak (fraction 237) was not resolved in the profile from B10.A mice. However, comparison with another run of B10.A peptides (Fig. 1A) shows that this peptide is present in the H-2K^k molecule of B10.A mice. Therefore, these results indicate that the H-2K molecules from B10.MOL1 and B10.A mice are indistinguishable by comparative tryptic peptide mapping.

Tryptic Peptide Map Comparisons of H-2D Molecules Extracted from STU, B10.SAA48, and B10.CAS2 Mice. The serologically indistinguishable H-2D molecules of these three strains were extracted by using antiserum G-25, preabsorbed with STC77 and W12A cells, which carry H-2K molecules serologically and biochemically similar to STU (unpublished data), and then compared by peptide mapping. Complete identity of the peptide profiles from these strains was obtained. The comparison of separate digests (upper and lower panels of Fig. 2) attests to the reproducibility of the method used.

Tryptic Peptide Map Comparisons of H-2K and H-2D Molecules Extracted from Other Wild-Derived and Inbred Lines. In our collection of B10.W lines (i.e., lines carrying wild-derived H-2 haplotypes on the B10 background) and in the classical inbred strains, we found several instances of serologically indistinguishable or similar K and D molecules. We compared several of these molecules by the peptide mapping method, and a summary of all the peptide-mapping comparisons, including those done previously (13), is given in Table 2. In general, there is a remarkable correlation between the serological findings and the biochemical analysis: only molecules distinguishable serologically can also be distinguished by peptide mapping. However, the main point is that H-2 molecules occurring in inbred mice occur also quite frequently in wild mice from different localities. The inbred- and wild-derived molecules are often indistinguishable serologically and biochemically.

m 11 o	^			•
Table 2	Summary of	tryptic	peptide ma	n comparisons
	Summing of	or y poro	populae ma	p comparisons

DISCUSSION

The Origin of B10. MOL1 and B10. CAS2 H-2 Alleles. Before making any farfetched conclusions about the evolution of the H-2 alleles, one has to consider the possibility that the congeneic lines B10.MOL1 and B10.CAS2 might have been genetically "contaminated" and somewhere in their development might have acquired H-2 alleles from inbred strains. This possibility would seem to apply in particular to the B10.MOL1 line, which carries part of a haplotype $(H-2^k)$ involved in the production of this line (backcrossing to the $H-2^k$ B10. BR strain). However, we have H-2 typed the original wild mouse colony of M. m. molossinus and found the $H-2^{w12}$ haplotype of B10. MOL1 already present in these mice. Contamination during backcrossing could therefore be excluded. The possibility then remains that the wild mouse colony has been contaminated by M.m. domesticus but this again is highly unlikely because the mice retained the morphology and isozyme profiles of M. m. molossinus (unpublished data). Furthermore, although one may feel apprehensive about the origin of $H-2^{w12}$ in the B10. MOL1 line, there is no possibility whatever that the $H-2^{w17}$ haplotype of B10.CAS2 mice could be of M.m. domesticus origin. This haplotype does not occur in any of the inbred strains and we have not found it in any M.m. domesticus wild mice. Furthermore, haplotypes carrying the D^{w3} allele were not involved in any way in the derivation of the B10.CAS2 line: the STU strain was brought into our colony only after our move from the United States to Germany and has not been available outside of Tübingen at any time and the B10.SAA48 strain has never been in contact with the wild colony of M.m. castaneus. We conclude therefore that the probability of the B10.MOL1 and B10.CAS2 lines carrying H-2 alleles of M. m. domesticus is virtually zero.

Strain	<i>H-2</i> haplotype	Origin of <i>H-2</i> region*						Peptides, no.		Variance †	Serological
		K	A _a	A _β	E_{β}	E _a	D	Detected	Different	%	difference
B10.S(9R)	t4	8	s	8	s	7	d	28			
B10.GAA37‡	w21	8	s	8	s	w21	w16	28	0	0	No
B10.KPB128‡	w19	8	8	8	8	w19	Ь	28	0	0	No
B10.A	a	k	k	k	k	7	d	26			
CBA(M523)	km1	km1	k	k	k	7	k	26	2	3	Yes
B10.MOL1	w12	k	k	k	_	7	w12	26	0	0	No
B10.Q	a	a	a	a	a	0	a	31			
B10.CHA2	w26	q	k	k	k	7	w26	31	0	0	No
B10.PL	u	u	u	u	u	7	u	28			
B10.NZW	2	u	u	u	u	7	z	28	0	0	No
B10.NZW	z	u	u	u	u	7	z	31			
B10.STC77	w14	dv2	р	р	р	р	Z	31	0	0	No
STU	w34	dv4	k	k	k	7	w3	32			
B10.CAS2	w17	w17	w17	w17	_	0	w3	32	0	0	No
B10.SAA48	w3	w3	w3	<i>w</i> 3	w3	7	w 3	32	0	0	No
TBR3	at3	s	_	_		0	Ь	29			
B10.KPB128	w19	8	8	\$	s	w19	b	29	0	0	No
B10.WB	j	j	j	j	j	7	b	29	0	0	No

-, Not known. Symbols in bold face indicate regions for which the identity of gene products has been established by peptide mapping.

* See ref. 1.

[†] Determined by using the formula (no. of different peptides \times 100)/no. of ³H- and ¹⁴C-labeled peptides.

[‡]See ref. 13.



FIG. 3. Orthodox (Left) and unorthodox (Right) interpretation of Mhc allele evolution in mice and humans. Hypothetical evolutionary trees of two HLA and two H-2 alleles are shown.

Implications of the Findings for the Evolution of the H-2Genes. If the occurrence of identical H-2 molecules in different mouse subspecies is not the result of contamination and if the molecules are truly identical (aside from resolution problems possibly overestimating similarity, peptides that flow through the column, are insoluble, or are present in very low yields are not compared), how then can one reconcile this finding with the extraordinary H-2 polymorphism? One possibility is that the identity is the result of convergent evolution: one could argue. for example, that the K^k alleles in $M \cdot m$. domesticus and $M \cdot m$. molossinus arose independently under similar selective pressures. If the result obtained with the B10. MOL1 and B10. CAS2 lines were the only available data pertaining to this question, one could perhaps entertain this possibility seriously but, since we find identical H-2 alleles all over the world (Table 2), we consider this explanation extremely unlikely. It seems to us a physical impossibility to develop within a relatively short time (1 to 2 million years) so many identical alleles in so many different places. Also, we should have been able to see this extremely rapid diversification of alleles in action by the occurrence of many intermediate forms, but we found no evidence for an abundance of such forms. We therefore consider the convergent evolution argument highly unlikely and so we are left with only two possibilities: either the alleles are in fact genes at different loci or the H-2 polymorphism is older than the house mouse. The former explanation requires the existence of at least 100 K and 100 D functional genes in one H-2 chromosome. However, recent DNA cloning studies do not provide any evidence for such a multitude of H-2 genes (17–20). The maximum estimate is 20-30 genes, many of which (perhaps the majority) are nonfunctional pseudogenes. These findings argue against the pseudoallele hypothesis and leave us with the possibility that H-2 genes evolve differently from other alleles. We propose that the evolution of H-2 alleles is trans-specific and started long before the house mouse became established as a species (ref. 9; Fig. 3). One can envision that the individual Mhc alleles were already beginning to diversify at the time when the complex assumed its present-day function-perhaps in the vertebrate ancestors. Then, during vertebrate evolution, mutations gradually accumulated, diversifying more and more the initial group of alleles, in the same way that individual genes diversify in the different species. Each species thus always starts with a certain set of already diversified alleles, diversifies them further, and passes them on to a newly evolving species.

This hypothesis explains the following findings, which are difficult to explain in any other way. First, it explains the occurrence of allotypic determinants in different species: for example, Iha et al. (21) found certain HLA alloantigens to be shared by some humans with some cows; Albert et al. (22) and Ivašková et al. (23) made a similar observation while comparing humans, mice, and rabbits. Second, the hypothesis explains the observation that the degree of homology of two K alleles or of two D alleles is almost the same as that of the K and D genes (2-5). And third, it explains the fact that the degree of homology of human and mouse Mhc genes is almost the same as that of the H-2 genes K and D or the HLA genes A and B (2-5). As one can see from the hypothetical pedigree (Fig. 3), these observations are exactly as expected under the assumption of trans-specific evolution of Mhc alleles. We conclude therefore that the *trans*-specific model of evolution is the only explanation that makes sense of the Mhc polymorphism and of the remarkable variation of the Mhc molecules at the level of the primary structure.

We thank Dr. D. Götze for antiserum G-25, Mr. Alfons Grammer, Physikalisches Institut, Tübingen, for help in writing the FORTRAN programs, Dr. Hildegard Kraut for the supply of Lens culinaris lectin, Ms. Rosemary Franklin for editorial assistance, and Ms. K. Masur for secretarial help. This work was supported in part by Grant AI 14736 from the National Institutes of Health and by a grant from the Volkswagen Foundation.

- Klein, J. & Figueroa, F. (1981) Immunol. Rev. 60, 23-57. 1.
- 2. Nathenson, S. G., Ewenstein, B. M., Uehara, H., Martinko, J. M., Coligan, J. E. & Kindt, T. J. (1981) in Current Trends in Histocompatibility, eds. Reisfeld, R. A. & Ferrone, S. (Plenum, New York), Vol. 1, pp. 415-436.
- Silver, J. & Hood, L. (1976) Contemp. Top. Mol. Immunol. 5, 3. 35-68.
- Vitetta, E. S. & Capra, J. D. (1978) Adv. Immunol. 26, 147-193.
- Strominger, J. L., Engelhard, V. H., Fuks, A., Guild, B. C., Hyafil, F., Kaufman, J. F., Korman, A. J., Kostyk, T. G., Kran-gel, M. S., Lancet, D., Lopez de Castro, J. A., Mann, D. L., Orr, H. T., Parham, P., Parker, K. C., Ploegh, H. L., Pober, J S., Robb, R. J. & Shackelford, D. A. (1981) in The Role of the Maior Histocompatibility Complex in Immunobiology, eds. Benacerraf, B. & Dorf, M. E. (Garland, New York), pp. 115-171.
- Strosberg, A. D. (1977) Immunogenetics 4, 499-513.
- Ohta, T. (1981) Genet. Res. 37, 133-149. 7.
- Bodmer, W. F. (1973) Transplant. Proc. 5, 1471-1475. 8.
- Klein, J. (1980) in Immunology 80, eds. Fougereau, M. & Daus-9. set, J. (Academic, London), pp. 239–253. Zaleska-Rutczynska, Z. & Klein, J. (1977) J. Immunol. 119,
- 10. 1903-1911.
- Schäfer, W. (1979) Z. Naturforsch. 34c, 306-309. 11.
- Duncan, W. R., Wakeland, E. K. & Klein, J. (1979) Immunoge-12. netics 9, 261-272.
- 13. Arden, B., Wakeland, E. K. & Klein, J. (1980) J. Immunol. 125, 2424 - 2428
- 14. Yonekawa, H., Moriwaki, K., Gotoh, O., Watanabe, J., Hayashi, J.-I., Miyashita, N., Petras, M. L. & Tagashira, Y. (1980) Jpn. J. Genet. 55, 289-296.
- Uehara, H., Ewenstein, B. M., Martinko, J. M., Nathenson, S. 15. G., Coligan, J. E. & Kindt, T. J. (1980) Biochemistry 19, 306-315.
- Ewald, S. J., Klein, J. & Hood, L. E. (1979) Immunogenetics 8, 16. 551-559.
- 17. Cami, B., Brégégère, F., Abastado, J. P. & Kourilsky, P. (1981) Nature (London) 291, 673-675.
- Steinmetz, M., Frelinger, J. G., Fisher, D., Hunkapiller, T., Pereira, D., Weissman, S. M., Uehara, H., Nathenson, S. & 18 Hood, L. (1981) Cell 24, 125-134.
- Kvist, S., Brégégère, F., Rask, L., Cami, B., Garoff, H., Daniel, 19 F., Wiman, K., Larhammar, D., Abastado, J. P., Gachelin, G., Peterson, P. A., Dobberstein, B. & Kourilsky, P. (1981) Proc. Natl. Acad. Sci. USA 78, 2772-2776
- Ploegh, H. L., Orr, H. T. & Strominger, J. L. (1980) Proc. Natl. 20. Acad. Sci. USA 77, 6081-6085
- Iha, T. H., Gebrandt, G., Bodmer, W. F., McGary, D. & Stone, 21. W. H. (1973) Tissue Antigens 3, 291-302.
- Albert, E., Kano, K., Abeyounis, C. J. & Milgrom, F. (1969) 22 Transplantation 8, 466-471
- 23. Ivašková, E., Dausset, J. & Iványi, P. (1972) Folia Biol. (Praha) 18, 194-197.