Brief Definitive Report

INTRAGENIC RECOMBINATION IN AN E_β GENE FOR A MURINE Ia ANTIGEN*

By MARIAN L. PLUNKETT,‡§ JOHN E. COLIGAN,|| CHELLA S. DAVID,¶ AND JOHN H. FREED§

From the Department of Biophysics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205; the Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20205; and the Department of Immunology, Mayo Clinic and Medical School, Rochester, Minnesota 55905

The plasma membrane-located murine Ia antigens consist of two polymorphic chains: the larger chain, designated α, is ~35,000 mol wt, whereas the smaller chain, designated β, is ~28,000 mol wt (1). Whereas the I-A antigens are encoded by genes (A_α and A_β) located within the *I-A* subregion, the I-E antigens are encoded by one gene (E_α) located in the *I-E* subregion and a second (E_β or A_e)[1] in the *I-A* subregion (2).

Previous studies from our laboratories have used F_1 hybrids to explore the serological (3, 4) and biochemical (5) basis for the complementation between the E_α and E_β genes. These studies, which established that complementation could occur between the genes involved in the *trans* configuration, provided evidence that the structural genes for the α and β chains were the complementing genetic elements. During the course of the biochemical studies (5), we compared the E_β chain derived from a (D2.GD × A.TFR5)F_1 (A^d, E^b × A^f, E^k)[2] to the E^d_β chain from B10.D2. Although the E_β chain from the F_1 animals (presumably encoded by the D2.GD parental chromosome in the F_1) was quite similar to the E^d_β chain from B10.D2, they were not identical. This result was not expected because D2.GD (*H-2*^{g2}) is a recombinant strain previously typed as $K^d\ A^d|B^b\ J^b\ E^b\ C^b\ S^b\ D^b$, and therefore it should possess the same E^d_β gene as B10.D2 (*H-2*^{d}). This difference between the E_β chain encoded by the D2.GD genome and other bona fide E^d_β chains was also noted by Jones (6), who used two-dimensional (2D) electrophoresis to document the difference.

Two obvious and mutually exclusive hypotheses may be put forward to explain the aberrant nature of the E_β chain from D2.GD. The first of these is that the recombinational event in D2.GD occurred not between the *I-A* and the *I-B* subregions, as

* Supported in part by grants CA-20820, CA-24473, and AI-14764 from the National Institutes of Health, U. S. Public Health Service.

‡ Supported by predoctoral training program grant GM-07445 from the National Institutes of Health.

§ Department of Biophysics, Johns Hopkins University School of Medicine.

|| Laboratory of Immunogenetics, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

¶ Department of Immunology, Mayo Clinic and Medical School.

[1] Jones and her co-workers originally used the designation A_e to denote the I-A subregion location of this gene. We will use the E_β notation because the gene product, the E_β chain is part of the I-E antigen.

[2] Strain D2.GD, like all *I-E*^{b} type strains, does not contain a functional E_α gene, whereas the *I-A*^{f} region of A.TFR5 does not contain a functional E_β gene. Thus, (D2.GD × A.TFR5)F_1 animals express only a single I-E antigen produced by *trans* complementation of the E^k_α and E^d_β genes.

J. Exp. Med. © The Rockefeller University Press · 0022-1007/82/03/0937/06 $1.00 937
Volume 155 March 1982 937–942

previously indicated, but rather within the E_β gene within the *I-A* subregion. Alternatively, the previous assignment of the recombinational event in D2.GD was correct, and mutations have occurred within the E_β gene of D2.GD, allowing it to drift slightly from a true E_β^d gene. Data are presented in this paper that support the first of these two alternatives.

Materials and Methods

Mice. (D2.GD × A.TFR5)F_1 mice were from the colony (CSD) at the Mayo Clinic and Medical School. Mice of the B10.D2 and B10.A(5R) strains were obtained from The Jackson Laboratory (Bar Harbor, ME).

Isolation of I-E Antigens. The α and β chains of I-E antigens were isolated as previously described (5). Briefly, leukocytes from normal spleens were metabolically radiolabeled with ^{3}H-labeled or ^{14}C-labeled amino acids by incubation for 21 h, and, after labeling, Nonidet P-40 (NP-40) lysates of the cells were prepared. The I-E antigens were isolated by lentil lectin chromatography and indirect immunoprecipitation, using an (A.BY × B10.A(18R))F_1 anti-B10.A(3R) alloantiserum and *Staphylococcus aureus* Cowan I strain. Alternatively, the immunoglobulin-depleted NP-40 lysate was applied directly to an affinity chromatography column, prepared by coupling purified anti-Ia.7 monoclonal antibody (14-4-4S, produced by hybridoma cells generously provided by Dr. David H. Sachs, National Institutes of Health, Bethesda, MD) to Sepharose, followed by elution of the bound I-E antigen from the column with glycine-HCl (pH 2.1).[3] The E_α and E_β chains from the isolated I-E antigens were separated without reduction by preparative polyacrylamide gel electrophoresis in sodium dodecylsulfate-containing buffers (SDS-PAGE).

Tryptic Digestion and Analysis of Peptides. Purified E_α and E_β chains obtained from SDS-PAGE were reduced and alkylated and digested with trypsin, as previously described (5). Analysis of tryptic peptides was by medium pressure ion exchange chromatography on a DC-6A cation exchange resin (Dionex Corp., Sunnyvale, CA) eluted with a pH 3.13 to pH 5.05 pyridine/acetic acid gradient (5).

Amino-Terminal Amino Acid Sequence Analysis. The E_α and E_β chains to be used for amino acid sequence studies were extracted from the appropriate SDS-PAGE slices with 0.01% SDS and then lyophilized. The lyophilized material was redissolved in 1 ml water, dialyzed twice against 1 liter of 0.001% SDS, and re-lyophilized. The sample was dissolved in 2 N formic acid and applied to a Beckman 890C sequencer (Beckman Instruments Inc., Fullerton, CA) for amino acid sequence analysis (7).

Results

Tryptic Peptide Comparison of the E_β^d Chain with the E_β^{g2} Chain from (D2.GD × A.TFR5)F_1 Hybrids. Previous studies of the I-E antigen obtained from (D2.GD × A.TFR5)F_1 hybrids have shown that the E_α chain contributed by the A.TFR5 parent is identical with the E_α^k chain from B10.BR(5).

However, the E_β^{g2} chain contributed by the D2.GD parent differed somewhat from the E_β^d chain from B10.D2(5). Fig. 1 shows the comparison of [^{14}C]arginine and lysine-labeled tryptic peptides from B10.D2 with the [^{3}H]arginine and lysine-labeled tryptic peptides derived from the E_β^{g2} chain from the D2.GD parent in the F_1 hybrid. The use of arginine and lysine labels allows visualization of all tryptic peptides in the molecules, except the carboxy-terminal peptide. Although the two profiles are quite similar, sharing fourteen peptides between them, there are three peptides unique to the E_β^{g2} chain (marked by the open arrows) and one unique to the E_β^d chain (marked by the filled arrow).

Tryptic Peptide Comparison of the E_β^b Chain with the E_β^{g2} Chain from (D2.GD × A.TFR5)F_1

[3] Kupinski, J. M., M. L. Plunkett, and J. H. Freed. Manuscript in preparation.

Hybrids. We next compared the tryptic peptide profile from the E_β^{g2} chain with that from the E_β^b chain from B10.A(5R). The rationale for this experiment was as follows: if the E_β chain from the *g2* haplotype arises from an $H\text{-}2^d|H\text{-}2^b$ intragenic recombinational event, then those tryptic peptides that are unique to E_β^{g2} as compared with E_β^d should correspond to some of the peptides produced by tryptic digestion of the E_β^b chain. Fig. 2 shows the comparison of the arginine and lysine-labeled tryptic peptides from the E_β^{g2} and E_β^b chains. As may be seen from the figure, the E_β chains from these two haplotypes share a number of peptides (peptide homology of ~55%) but, significantly, those three peptides by which E_β^{g2} differs from E_β^d are shared between E_β^{g2} and E_β^b. These peptides are indicated by arrows in Fig. 2.

Sequence Analysis of the Amino-Terminal Portion of E_β Chains. Radiochemical sequence analysis was carried out on E_β chains metabolically radiolabeled with [^{3}H]phenylalanine. The results of the sequencer runs are presented in Fig. 3 and are summarized in Table I. These data demonstrate that the amino-terminal [^{3}H]phenylalanine profiles are identical for the E_β^d and E_β^{g2} chains, and these differ from the profile for E_β^b. This suggests that the amino-terminal portion of the E_β^{g2} chain is derived from the *d* haplotypes as opposed to the *b* haplotype. The E_α^k sequence (Fig. 3, panel D) from B10.A(5R) is included because the E_β chains that were sequenced contained E_α chains at the level of ~10% contamination. (In panels A–C, compare the size of the peaks at step 8 with the heights of the peaks at step 12.) All assigned phenylalanine positions are from peaks that fit a single repetitive yield line (~92%). These data support the postulate that each sample sequenced was a single species that, except for the limited

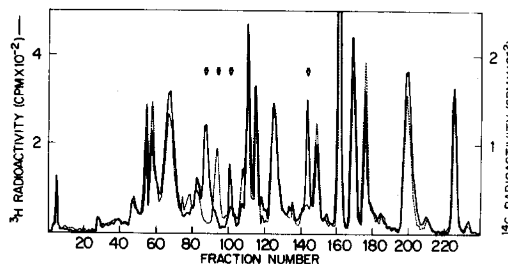

FIG. 1. Ion exchange chromatography of arginine- and lysine-labeled E_β^d chain from B10.D2 (---) and from the E_β^{g2} chain from (D2.GD × A.TFR5)F$_1$ (——). The open arrows mark peptides unique to the E_β^{g2} chain; the filled arrow marks a peptide unique to the E_β^d chain. See Materials and Methods and reference 5 for details of the method.

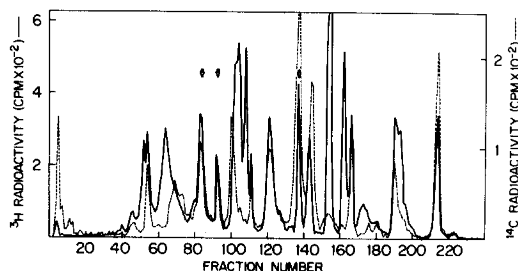

FIG. 2. Ion exchange chromatography of arginine- and lysine-labled tryptic peptides from the E_β^b chain from B10.A(5R) (---) and from the E_β^{g2} chain from (D2.GD × A.TFR5)F$_1$ (——). The arrows mark the three peptides by which the E_β^{g2} chain differs from the E_β^d chain.

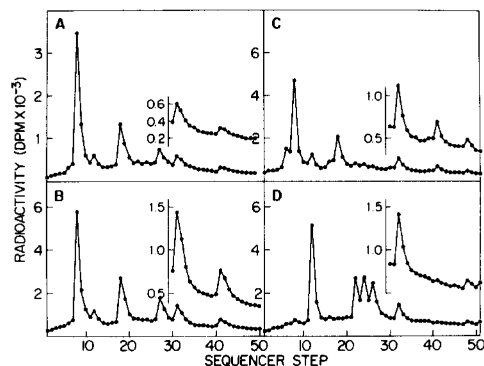

Fig. 3. Radiochemical sequence analysis of the amino-terminal portions of [^{3}H]phenylalanine-labeled E_β^{g2} chain from (D2.GD × A.TFR5)F$_1$ (A); E_β^d chain from B10.D2 (B); and the E_β^b chain (C) and the E_α^k chain (D) from B10.A(5R). Radioactivity found at each step is plotted against the sequencer step. See Materials and Methods and reference 7 for details of the method.

Table I
Amino-Terminal Sequence Analysis of E_β Chains

Strain	Chain	Sequencer steps containing phenylalanine										
(D2.GD × A.TFR5)F$_1$	E_β^{g2}	8	—*	18	—	—	—	27	31	—	41	—
B10.D2	E_β^d	8	—	18	—	—	—	27	31	—	41	—
B10.A(5R)	E_β^b	8	—	18	—	—	—	—	—	32	41	48
B10.A(5R)	E_α^k	—	12	—	22	24	26	—	—	32	—	48

* Indicates phenylalanine is absent at this position.

contamination by the E_α chain, was homogeneous E_β chain. The phenylalanine residues identified within positions 1–32 agree with data previously published for E_β^b, E_β^d, and E_α^k chains (8–11), whereas those identified at positions 41 and 48 (see Table I) have not been reported previously.

Discussion

Because the D2.GD (12) and B10.D2 (13) strains were each derived from crosses involving DBA/2 (*H-2^d*), any genetic elements with the *d* haplotype in these strains should be identical because they were derived from the same donor strain. For the A_α^d and A_β^d chains this appears to be true, as Jones (6) was able to demonstrate identical two dimensional gel patterns for these chains from all three strains. The analysis of the E_β chains is complicated somewhat by the fact that D2.GD will not express normal amounts of the E_β^{g2} chain, and an F$_1$ hybrid between D2.GD and A.TFR5 ($A^{\text{“}d\text{”}}$ E^b × A^f E^k) must be used to obtain usable amounts of the E_β^{g2} chain. Again, Jones' two dimensional gel analysis (6) suggested that there was a difference in electrophoretic mobility among the E_β chains from the three strains and that the E_β chain derived from the D2.GD genome in the F$_1$ hybrid was clearly different from the E_β^d chains from B10.D2 and DBA/2. Furthermore, previous observations from one of our laboratories (14) have identified a serological difference between these E_β chains: bona fide E_β^d chains (such as that from B10.D2) express a specificity, Ia.50, absent from the E_β^{g2} chain.

In our previous analysis (5) of the E_α and E_β chains from (D2.GD × A.TFR5)F$_1$ animals, we made two observations: (*a*) the E_α^k chain, putatively contributed by the

A.TFR5 parental chromosome, was identical to authentic E^k_α chains; and (*b*) the E^{g2}_β chain, putatively contributed by the D2.GD parental chromosome, differed slightly from authentic E^d_β chains. As the results presented in Figs. 1 and 2 amply demonstrate, the three peptides unique to the E^{g2}_β chain, when compared with the E^d_β chain, do co-elute with tryptic peptides from the E^b_β chain.

There are two likely explanations for the "aberrant" E_β chain in D2.GD. The first is that a series of mutations have occurred within the E^{g2}_β gene. This would make the E^{g2}_β vs. E^d_β system analogous to the system of mutations observed in the H-2K^b antigen (15). As few as two amino acid substitutions, both of the X → lysine or X → arginine type, could explain the differences observed between the E^{g2}_β and E^d_β profiles. The second explanation is that a recombinational event occurred within the gene for the E^{g2}_β chain. Hence, the haplotype for D2.GD should be: $H\text{-}2^{g2} = K^d, A^{d|b}, B^b, J^b, E^b, C^b, S^b, D^b$. We feel that the results presented in Fig. 2 strongly support this latter model because it is highly unlikely that two-point mutations in the E^d_β gene would give rise to tryptic peptides that co-elute with peptides derived from the E^b_β gene.

Current understanding of segmented genes for major histocompatibility complex molecules, e.g., the eight gene segments in a pseudogene for the *H-2*-linked Qa-2 antigen (16), does not permit one to say whether the crossover event in D2.GD occurred within a coding sequence (exon) or within an intervening sequence (intron) between two exons. Either of these possibilities (both of which would be a classical "intragenic" recombination) would be consistent with our results, and distinguishing between them will require full knowledge of the genomic DNA in this region of the 17th chromosome of D2.GD.

We explored the genetic origin of the E^{g2}_β chain further by comparing its amino-terminal sequence (phenylalanines only) to that of the E^d_β and E^b_β chains. The choice of labeled phenylalanine was based on previously published (9, 10) data demonstrating that the positions of phenylalanines in E^d_β and E^b_β chains were different. The results presented in Fig. 3 and Table I suggest that the amino-terminal portion of the E^{g2}_β chain, at least up to position 48, is derived from the *d* haplotype. This observation, when coupled with the information available on the genetic origin of the other *H-2* regions of D2.GD, suggests that the orientation of the coding strand for the E_β chain (at least for D2.GD) is centromeric → telomeric for the 5′ → 3′ reading direction.

Recently, Rose and Cullen (17) have published a similar study using the recombinant haplotype $H\text{-}2^{t1}$ to suggest that the A_α gene most likely lies to the centromeric side of the A_β and E_β genes. These results combined with those presented here suggest that the gene order for the *I-A* subregion of the *H-2* complex should be (*H-2K*), A_α, A_β, E_β ... (*H-2D*).

Summary

The recombinant strain D2.GD was originally typed as $I\text{-}A^d$ by serological methods. Indeed, the A_α and A_β chains of the I-A antigens appear to exhibit normal behavior by the criteria of serology and two dimensional gel analysis. However, the E_β chain encoded by the *I-A* subregion of this strain, one of the two components of the plasma membrane located I-E antigens produced in (D2.GD × A.TFR5)F_1 animals, has been demonstrated to be the product of an intragenic recombinational event between E_β genes from the *d* and *b* haplotypes.

Sequence analysis suggests that the amino-terminal portion of the E^{g2}_β chain is

derived from the *d* haplotype and, therefore, that the coding strand for this gene is oriented centromeric → telomeric (5′ to 3′ direction). Finally, these data combined with the data of Rose and Cullen (17) allow the ordering of the genes within the *I-A* subregion as (H-$2K$), A_α, A_β, E_β ... (H-$2D$).

Received for publication 14 December 1981

References

1. Freed, J. H. Biochemistry of the murine Ia Antigens. *In* Ia Antigens and Their Analogs in Man and Other Animals. S. Ferrone and C. David, editors. CRC Press, Boca Raton, Florida. In press.
2. Jones, P. P., D. B. Murphy, and H. O. McDevitt. 1978. Two-gene control of the expression of a murine Ia antigen. *J. Exp. Med.* **148:**925.
3. Lafuse, W. P., J. F. McCormick, and C. S. David. 1980. Serological and biochemical identification of hybrid Ia antigens. *J. Exp. Med.* **151:**709.
4. Lafuse, W. P., J. F. McCormick, and C. S. David. 1980. Lack of gene complementation prevents expression of I-E^k in recombinant A.TFR5. *J. Immunol.* **124:**2511.
5. Plunkett, M. L., C. S. David, and J. H. Freed. 1981. Biochemical evidence for *trans* complementation of structural genes in the expression of I-E antigens in F_1 hybrids. *J. Immunol.* **127:**1679.
6. Jones, P. P. 1980. Aberrant A_e (E_β) Ia polypeptide chain in H-2^{g2} haplotype mice. Possible result of an intragenic recombination or mutation. *J. Exp. Med.* **152:**1453.
7. Coligan, J. E., and T. J. Kindt. 1981. Determination of protein primary structure by radiochemical techniques. *J. Immunol. Methods.* **47:**1.
8. McMillan, M., J. M. Cecka, D. B. Murphy, H. O. McDevitt, and L. Hood. 1977. Structure of murine Ia antigens: partial NH_2-terminal amino acid sequences of products of the *I-E* or *I-C* subregion. *Proc. Natl. Acad. Sci. U. S. A.* **74:**5135.
9. McMillan, M., J. M. Cecka, D. B. Murphy, H. O. McDevitt, and L. Hood. 1978. Partial amino acid sequences of murine Ia antigens of the I-E/C^d subregion. *Immunogenetics.* **6:**137.
10. Silver, J., W. A. Russell, B. L. Reis, and J. A. Frelinger. 1977. Chemical characterization of murine Ia alloantigens determined by the *I-E/I-C* subregions of the *H-2* complex. *Proc. Natl. Acad. Sci. U. S. A.* **74:**5131.
11. Cook, R. G., M. H. Seligman, J. D. Capra, J. W. Uhr, and E. S. Vitetta. 1979. Structural studies on the murine Ia alloantigens. IV. NH_2-Terminal sequence analysis of allelic products of the *I-A* and *I-E/C* subregions. *J. Immunol.* **122:**2232.
12. Lilly, F., and J. Klein. 1973. An H-2^g-like recombinant in the mouse. *Transplantation.* **16:** 530.
13. Snell, G. D. 1958. Histocompatibility genes of the mouse. II. Production and analysis of isogenic resistant lines. *J. Natl. Cancer Inst.* **21:**843.
14. Lafuse, W. P., H. I. Hendrickson, P. S. Corser, and C. S. David. 1981. Serological and biochemical analysis of a varient A_e (E_β) Ia polypeptide chain in D2.GD (H-2^{g2}). *J. Immunol.* **127:**811
15. Nairn, R., K. Yamaga, and S. G. Nathenson. 1980. Biochemistry of the gene products from murine MHC mutants. *Annu. Rev. Genet.* **14:**241.
16. Steinmetz, M., K. W. Moore, J. G. Frelinger, B. T. Sher, F.-W. Shen, E. A. Boyse, and L. Hood. 1981. A pseudogene homologous to mouse transplantation antigens: transplantation antigens are encoded by eight axons that correlate with protein domains. *Cell.* **25:**683.
17. Rose, S. M., and S. E. Cullen. 1981. A varient α-chain in an Ia molecule from the I-A subregion of the mouse major histocompatibility complex: a possible intragenic recombination. *J. Immunol.* **127:**1472.