The EMBO Journal vol.5 no.10 pp.2469-2476, 1986

# The mouse *Eβ2* gene: a class II MHC β gene with limited intraspecies polymorphism and an unusual pattern of transcription

**Ned S.Braunstein and Ronald N.Germain**

Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA

Communicated by B.Mach

**Analysis of genomic clones containing the *Eβ2* region from the mouse major histocompatibility complex (MHC) reveals a gene that is similar to conventional class II β genes in its overall organization and sequence but unusual in that it shows limited intraspecies allelic polymorphism. Like the conventional class II β genes, the *Eβ2* gene is transcriptionally active in a variety of B lymphoid cells. However, its transcripts consist of multiple polyadenylated species from 1.8 to 3.6 kb in size which are present at only ~1/20 the level of *Eβ1* mRNA in the same cells. In addition, unlike other class II genes, *Eβ2* transcription is not induced by γ-interferon treatment of macrophage/monocyte tumor lines. These distinctive features suggest that the Eβ2 gene product may have a unique functional role, different from that of previously studied class II proteins.**

*Key words:* class II MHC β gene/gene expression/Ia antigen/nucleotide sequence/polymorphism

## Introduction

The only well-characterized products of the murine *H-2I* and the human *HLA-D* regions of the major histocompatibility complex (MHC) are components of heterodimeric cell surface glycoproteins. These class II or Ia molecules, the murine I-A (Aβ1:Aα) and I-E (Eβ1:Eα) antigens, their human homologues DQ (DQβ:DQα) and DR (DRβ1:DRα), respectively, and the human DP antigen (DPβ:DPα), are composed of a 34-kd α chain noncovalently associated with a 28-kd β chain, and are expressed predominantly on B lymphocytes, monocytes, macrophages, dendritic cells and activated human T lymphocytes (Hood *et al.*, 1983; Kaufman *et al.*, 1984). These proteins and their respective genes show extensive intraspecies allelic polymorphism. Transfection studies have shown that the polymorphism within the amino-terminal domains of the α and β chains controls the phenomenon of MHC-restricted helper T lymphocyte recognition of antigen (Folsom *et al.*, 1985; Germain *et al.*, 1985; Lechler *et al.*, 1986). This same sequence variation also appears to control the overall immune response potential of an individual (Ir gene phenotype) (Schwartz, 1986).

Other immune phenomena that appear to be regulated by loci lying within the murine *I* region cannot be readily accounted for by the direct activity of these conventional class II molecules. Such phenomena include regulation of T cell activity by the *I-B, I-C, I-J* and *I-N* subregions (Klein *et al.*, 1983), the determination of whether an animal activates suppressor T cells to a given antigen (Is phenotype) (Benacerraf and Dorf, 1977), and the presence of Ia-like molecules on T cells (Ia-t) (Hiramatsu *et al.*, 1982). Thus, attempts have been made to identify additional class II genes in both mouse and man whose products may account for these functional observations. Hybridization studies using various class II α or β gene probes have led to the identification of several cross-hybridizing but distinct genomic or cDNA clones with significant homology to the conventional class II α and β genes, designated *Eβ2*, *Aβ2*, *Eβ3* and *Aβ3* in the mouse (Steinmetz *et al.*, 1982; Larhammar *et al.*, 1983a; Devlin *et al.*, 1984; Widera and Flavell, 1985), and *DXα*, *DZα*, *DXβ* and *DOβ* in man (Auffray *et al.*, 1984; Spielman *et al.*, 1984; Okada *et al.*, 1985; Tonnelle *et al.*, 1985).

Among these additional murine class II sequences, *Eβ2* is of particular interest due to its location between *Eβ1* and *Eα*, near to where immunogenetic, but not molecular genetic analysis maps the *I-J* locus (Steinmetz *et al.*, 1982). *Eβ2* was first identified as a DNA segment to which 3′ *DQβ* but not 5′ *Eβ1* or *DQβ* probes hybridized (Steinmetz *et al.*, 1982). To determine if this region of DNA contains a complete, potentially expressible class II β gene, and to compare the organization, sequence and mode of expression of any such gene with conventional β genes, we have analyzed genomic clones containing the *Eβ2* region. Our results reveal the presence of additional, previously unidentified exons which are similar, but not identical to those of conventional class II β genes in size and organization. The *Eβ2* gene has a low level of intraspecies allelic polymorphism, and shows an unusual pattern of transcription and expression. These findings are similar to those recently reported for two other atypical class II genes, *Aβ2* (Wake and Flavell, 1985; Larhammar *et al.*, 1985a) and *DOβ* (Tonnelle *et al.*, 1985), suggesting that these three genes are members of a new class II sub-family whose products may have unique functional roles.

## Results

### *Isolation of Eβ2 genomic clones and hybridization localization of putative Eβ2 exons*

Genomic clones containing the *Eβ2* locus from the *H-2*$^{d}$ and *H-2*$^{a/k}$ haplotypes were isolated using a cDNA probe containing *Eβ1* exon 3 and 4 sequences. Cosmid H-2$^{k}$ 7.1 was provided by M.Steinmetz (Steinmetz *et al.*, 1984). To determine if sequences homologous to the major class II β gene exons were contained in the *Eβ2* region, β exon specific probes were hybridized to Southern blots containing DNA from the three genomic clones. An *Eβ1* exon 1 probe, a *DRβ1* exon 2 probe and the *Eβ1* 3′ probe described above hybridized to restriction fragments containing putative *Eβ2* exons 1-4, respectively (Figure 1).

### *Organization and sequence of the Eβ2 gene*

Following the strategy outlined in Figure 2, the nucleotide sequence of the exons identified by cDNA cross-hybridization, their immediate flanking regions, and both the expected promoter region 5′ of exon 1, and the region downstream of exon 4 anticipated to contain the exons for the intracytoplasmic and 3′ untranslated regions, was determined (Figure 3). The borders for exons 1, 2, 3 and 4 have been chosen based on homology to other class II β genes and the occurrence of consensus splice sequences in *Eβ2* at these positions (Breathnach and Chambon,

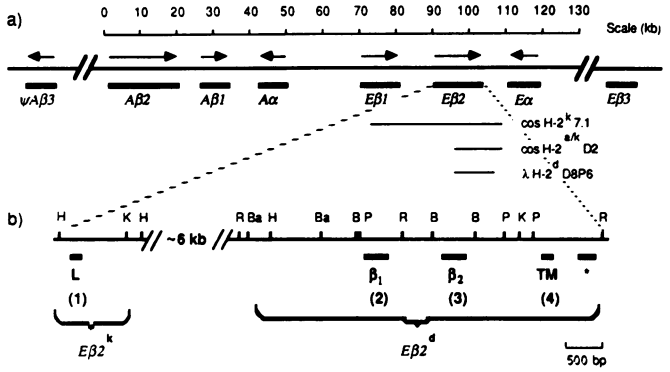


**Fig. 1.** Molecular organization of the mouse MHC and restriction map of the *Eβ2* gene. **(a)** The location of the class II genes (boxes) are shown as they lie in the mouse MHC on the short arm of chromosome 17. The centromere lies to the left of the figure. Arrows indicate direction of transcription of the genes (where known). The *Aβ3* and *Eβ3* loci have not been directly linked to the other genes, hence the discontinuities in the map. Inserts of the three clones used in this study are shown relative to their positions in this region. **(b)** Boxes identify regions having homology to other class II β genes as determined by cross-hybridization to cDNA probes. 1 refers to the portion of the gene homologous to sequences encoding the leader peptide; 2, the N-terminal external domain; 3, the second external domain; and 4, a region homologous to the transmembrane coding exon of *Eβ1*. The starred region shows sequence homology to the 3'UT region of *Eβ1*. The restriction fragment shown containing exon 1 comes from the $E\beta2^k$ allele. The rest of the map describes the $E\beta2^d$ allele. B = *Bgl*II; BA = *Bam*HI; H = *Hin*dIII; K = *Kpn*I; P = *Pst*I; R = *Eco*RI.

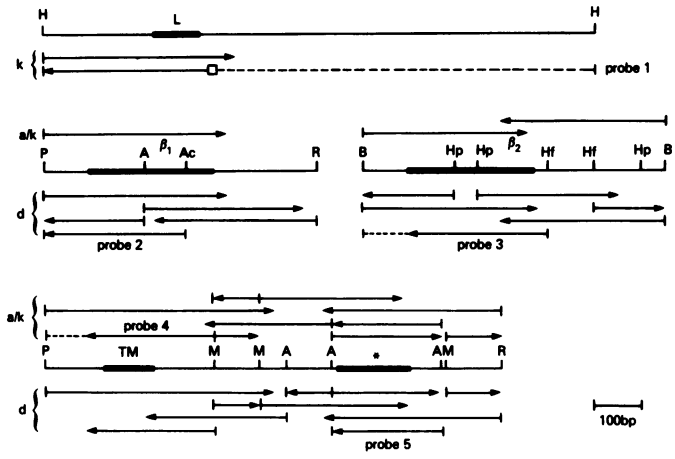


**Fig. 2.** Strategy for sequencing the *Eβ2* gene. Arrows show the direction and their length indicates the extent of sequence obtained from each clone for the *k*, *d* or *a/k* alleles. The box indicates where a synthetic oligonucleotide primer was used. Solid bars correspond to the exons described in the legend to Figure 2. The *Eβ2* fragments used as hybridization probes are numbered 1–5. Dashes indicate the full length of the restriction fragment used to make probes 1–5 where that length is different from the length sequenced. A = *Alu*I, Ac = *Acc*I, B = *Bgl*II, H = *Hin*dIII, Hf = *Hin*fI, Hp = *Hpa*II, M = *Mbo*I, P = *Pst*I, R = *Eco*RI.

1981; Cech, 1983). Analysis of *Eβ2* intron sequences demonstrates the presence of potential lariat branch points appropriately located 5′ of these consensus splice acceptor sites (Reed and Maniatis, 1985). As for other members of the immunoglobulin supergene family, all predicted splicing occurs between the first and second nucleotides of a codon (Hood *et al.*, 1983).

Several conserved sequences are found in the putative promoter region 5′ of exon 1. Two short stretches of conserved sequences found in all other class II genes analyzed to date are located 158 and 123 bp upstream of the potential initiation codon (Mathis *et al.*, 1983; Kelly and Trowsdale, 1985). A possible CAAT box is found at position −93 relative to this ATG (Figures 3 and 4). Although *Eβ2* does not contain a TATA sequence at the expected position, at least one expressed class II gene, *Aβ1*, also lacks an identifiable TATA box in this region (Malissen *et al.*, 1983; Larhammar *et al.*, 1983a).

The likely initiation codon was identified by alignment with the *Eβ1* sequence. The sequence immediately 5′ of this initiator conforms to the consensus (Kozak, 1984) and this codon is the first ATG triplet following the upstream conserved sequences. Starting with this putative initiation codon, and assuming splicing occurs as indicated, a continuous open reading frame encoding 253 amino acids is found. Based on analogy to other class II β genes, especially *Eβ1*, exon 1 would encode a hydrophobic leader peptide of 28 amino acids, followed by the first five amino acids of the mature Eβ2 protein. Exons 2, 3 and 4 would encode $\beta_1$, $\beta_2$ and TM domains of 89, 94 and 37 amino acids respectively. These are similar in size to those of other class II products and the exon 3 sequence is identical to that recently reported for $E\beta2^d$ (Denaro *et al.*, 1985). A stretch of 22 hydrophobic amino acids in exon 4 could serve as the membrane spanning segment of the protein. Immediately following the hydrophobic region, several charged amino acids are encoded. Such charged residues are thought to help anchor molecules in the membrane.

In the sequence 3′ of the transmembrane exon, several stretches having limited homology to the intracytoplasmic exon of one or another class II β gene could be identified. However, none of these could be definitively determined to be an *Eβ2* intracytoplasmic exon. This exon may be too divergent to detect by this method or may lie in a more 3′ region not sequenced. As shown in lines 23–25 of the sequence (Figure 3), a region was found which has homology to the 3′UT region of *Eβ1*. This is followed by the sequence AATAAT which, although different from the usual polyadenylation signal (Proudfoot and Brownlee, 1976), is the same as that found in a *DPβ* gene (Kelly and Trowsdale, 1985). While several possible splice acceptor sites precede this potential 3′UT sequence, sequence analysis of genomic cloned DNA alone cannot reveal which, if any, of these splice sites are used. More significantly, Northern blot analysis indicates that this sequence is not contained in the majority of *Eβ2* transcripts (see below). The precise identification of the *Eβ2* IC and 3′UT exons will therefore require isolation of an *Eβ2* cDNA.

Lastly, an inverted repeat consisting of 31 GT dinucleotides followed by eight repeats of the unit AGAC is found in the intron between exons 2 and 3. The GT-rich region is of interest in that it is typical of sequences capable of forming Z-DNA (Rich *et al.*, 1984).

### *Comparison of Eβ2 with other class II β sequences*

To determine if the potential Eβ2 protein shares features common to the other class II β sequences, the proposed Eβ2 amino acid sequence was aligned with that of other class II β polypeptides (Figure 5). The overall sizes of the various protein domains are very similar, the intradomain disulfide-linked cysteines present in the other class II $\beta_1$ and $\beta_2$ domains are conserved in Eβ2, and the hydrophobic portion of the Eβ2 protein aligns with the hydrophobic, transmembrane portions of the other molecules. Additionally, analysis of the second domains shows that those residues shared among class II $\beta_2$ domains and immunoglobulin domains are retained in Eβ2.

The alignment of the amino-terminal domains of the β poly-

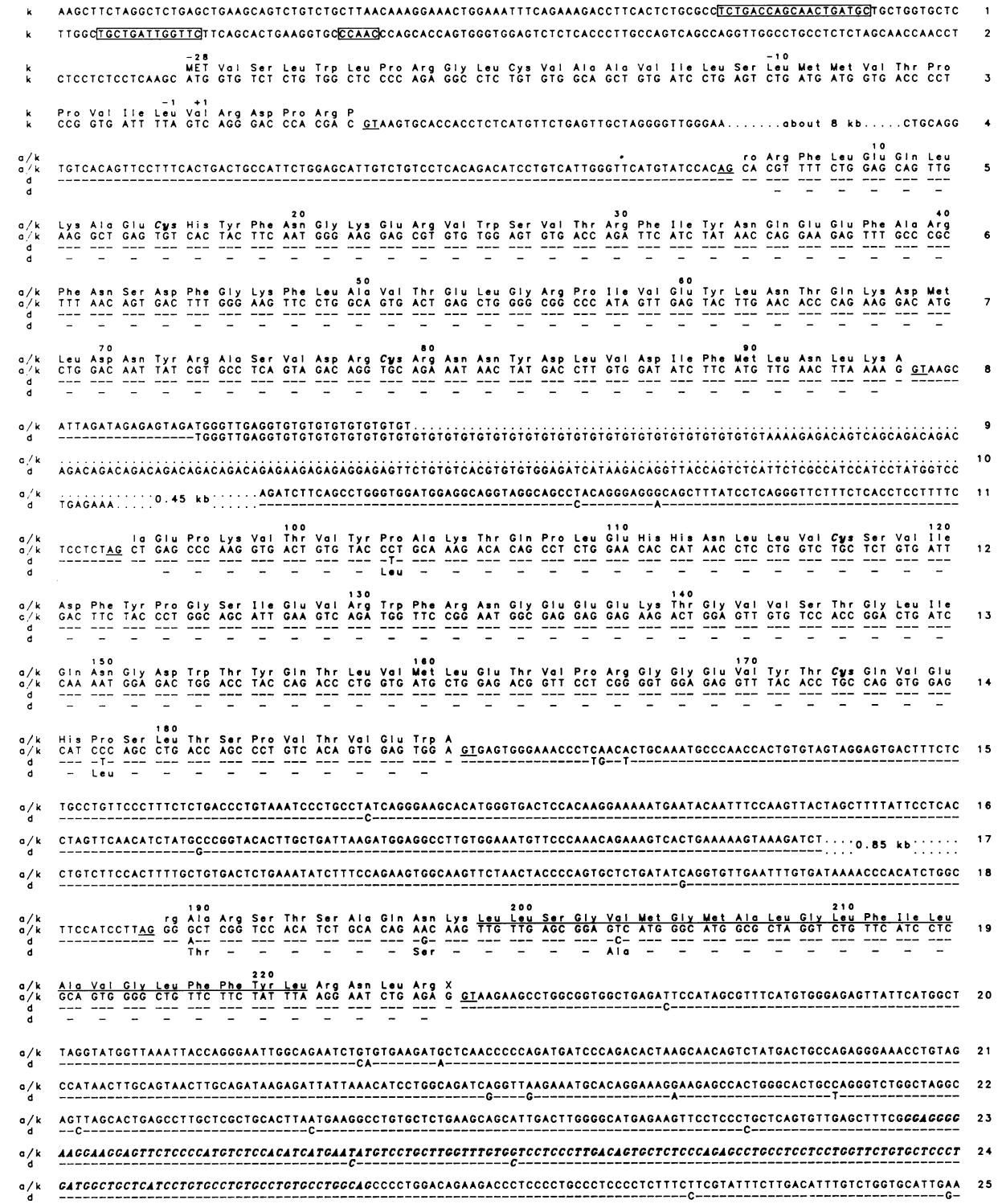

**Fig. 3.** Nucleotide sequence and predicted amino acid sequence of the $E\beta2^{a/k}$ and $E\beta2^{d}$ alleles. The exons shown encode the signal peptide, first and second external domains, and transmembrane regions of *Eβ2*. Exon boundaries were determined by alignment with other class II β genes and the conserved AG and GT dinucleotides at the proposed splice sites are underlined. Dashes represent identity between the two alleles. Dots represent regions not fully sequenced. Amino acids are numbered consecutively with no. 1 being the predicted amino terminus of the mature protein as deduced by alignment with other class II β chains. Conserved cysteines in the two external domains are italicized and the 22 hydrophobic amino acids predicted to form the transmembrane region are underlined. Boxed sequences in lines 1 and 2 are conserved residues found in the promoter regions of other class II genes. The boxed sequence in line 27 may serve as a polyadenylation signal for one of the transcripts. The parentheses in line 26 represent a deletion of four nucleotides in the *d* allele. The italicized stretch of 158 nucleotides in lines 23–25 shows homology with the *Eβ1* 3′UT region.

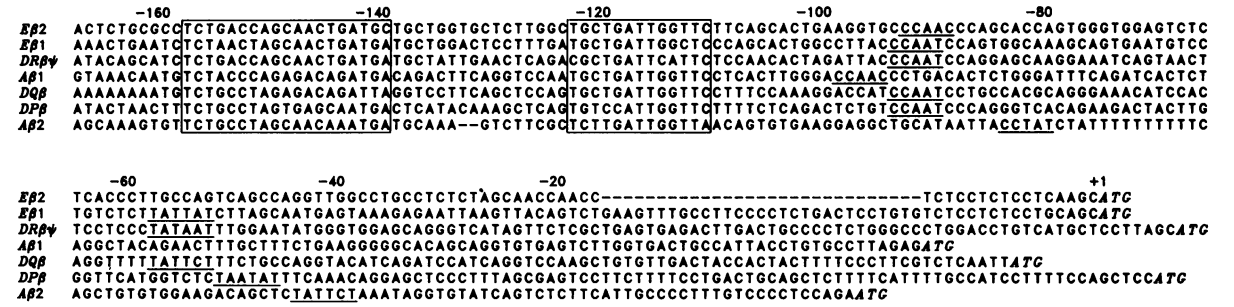

**Fig. 4.** Comparison of the promoter regions of class II genes. The conserved 19-mer and 12-mer sequences are boxed, CAAT and TATA sequences are underlined, and the initiation codons are italicized. Numbering is based on the *Eβ2* sequence. Dashes represent gaps inserted to maximize homology. Sequences were obtained from: *Eβ1*, Saito *et al.* (1983); *DRβΨ*, Larhammar *et al.* (1985b); *Aβ1*, Larhammar *et al.* (1983a); *DQβ*, Larhammar *et al.* (1983b); *DPβ*, Kelly and Trowsdale (1985); *Aβ2*, Larhammar *et al.* (1985a).

**Fig. 5.** Alignment of class II β polypeptides. The deduced amino acid sequence of *Eβ2* is shown starting with the predicted $NH_2$ terminus of the mature protein. For the other class II β chains, only those residues that differ from *Eβ2* are shown. Dashes represent gaps inserted to preserve the homology. Numbering is based on the *Eβ2* sequence. The open boxes enclose first domain residues common to all class II β chains except *Eβ2*. The hatched boxes enclose residues in common with an immunoglobulin constant region domain and the class II $\beta_2$ domains. The transmembrane regions are underlined in all the chains. Arrows represent exon boundaries. The * represents the amino acid substitution in *Eβ2* which alters the site for N-linked glycosylation. Sequences were obtained from: Eβ1, Saito *et al.* (1983); DRβ1, Long *et al.* (1983); Aβ1, Malissen *et al.* (1983); DQβ, Schenning *et al.* (1984); Aβ2, Larhammar *et al.* (1985a); DOβ, Tonnelle *et al.* (1985); DPβ, Kelly and Trowsdale (1985); γ2a, Dognin *et al.* (1981).

peptides shows that 18 out of 24 residues common to all other $\beta_1$ domains are present in Eβ2. Two of the six differences are conservative (Ala → Val and Glu → Asp), three are non-conservative (Asp → Asn, Trp → Leu and Arg → Leu), and the final difference (Thr → Lys) eliminates the single site for N-linked glycosylation found in all the other class II β chains.

Overall nucleic acid and deduced amino acid sequence homologies between the *Eβ2* exons/domains and other class II β exons/domains are given in Table I. At both the nucleotide and amino acid levels, exon 3 (second domain) is clearly most homologous to *Eβ1* and shows the same degree of homology to *DRβ* as is found between the other interspecies class II β homologues, i.e. *Eβ1* and *DRβ1* or *Aβ1* and *DQβ*. In this respect *Eβ1* and *Eβ2* seem to be like the non-allelic *DRβ* loci in comprising an isotypic family. However, the *Eβ2* second exon (first domain) has no clear homologue: it is no more like *Eβ1*, *DRβ1* or *DRβ3* than *Eβ1* is like *Aβ1* or *DPβ*. This divergence is much greater than that seen between the human non-allelic but isotype-related *DRβ1* and *DRβ3* loci. The transmembrane region is similarly divergent.

*Relative lack of polymorphism among Eβ2 alleles*

The sequences of the *d* and *a/k* alleles of *Eβ2* (Figure 3) show that their second exons are identical. In the third exon there are two nucleotide changes and in the transmembrane coding exon, three changes. All five changes alter the amino acid at their respective positions; at each of these positions, the $E\beta2^{a/k}$ amino acid sequence is the same as $E\beta1^{d}$ (Saito *et al.*, 1983) while the $E\beta2^{d}$ sequence is different. Note that while the exon 2 sequences shown are identical, allelic differences in the intervening sequence (not shown) confirm that the two sequences derive from different haplotypes. *Eβ2* is therefore less polymorphic than other β genes and, in particular, shows fewer exon 2 polymorphisms between alleles than does either *Eβ1* (Saito *et al.*, 1983; Mengle-Gaw and McDevitt, 1983; Denaro *et al.*, 1984) or *Aβ1* (Choi *et al.*, 1983). Whether this relative lack of *Eβ2* polymorphism is a consequence of selection for function or results from *Eβ2* lying in a region of the MHC having low sequence divergence (Steinmetz *et al.*, 1984) remains to be determined.

To examine further the extent of polymorphism at this locus, DNA from mice having the *p, s, r, b, d, a* and *k H-2* genotypes

**Table I.** Percent homology between class II β exon nucleotide and deduced amino acid sequences[a]

A. Exon 2 ($\beta_1$ domain)

| | Eβ2 | Eβ1 | Aβ1 | Aβ2 | DRβ1 | DQβ | DPβ | DOβ | DRβ3 | |
|---|---|---|---|---|---|---|---|---|---|---|
| Eβ2 | – | 67 | 63 | 58 | 70 | 65 | 64 | 61 | 73 | |
| Eβ1 | 53 | – | 75 | 60 | 78 | 70 | 70 | 63 | 75 | |
| Aβ1 | 45 | 60 | – | 63 | 74 | 77 | 73 | 63 | 76 | |
| Aβ2 | 41 | 45 | 45 | – | 61 | 64 | 64 | 83 | 63 | % Nucleotide |
| DRβ1 | 55 | 68 | 55 | 46 | – | 72 | 77 | 63 | 89 | homology |
| DQβ | 50 | 57 | 66 | 49 | 57 | – | 74 | 67 | 73 | |
| DPβ | 54 | 61 | 63 | 50 | 66 | 61 | – | 64 | 74 | |
| DOβ | 45 | 48 | 52 | 78 | 52 | 57 | 51 | – | 65 | |
| DRβ3 | 60 | 63 | 64 | 52 | 76 | 57 | 62 | 54 | – | |
| | | | | % Amino acid homology | | | | | | |

B. Exon 3 ($\beta_2$ domain)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Eβ2 | – | 90 | 71 | 70 | 85 | 73 | 77 | 71 | – | |
| Eβ1 | 85 | – | 70 | 70 | 83 | 73 | 77 | 70 | – | |
| Aβ1 | 67 | 66 | – | 66 | 69 | 82 | 76 | 68 | – | |
| Aβ2 | 62 | 65 | 59 | – | 73 | 70 | 71 | 83 | – | % Nucletoide |
| DRβ1 | 85 | 82 | 68 | 65 | – | 73 | 77 | 73 | – | homology |
| DQβ | 68 | 70 | 83 | 63 | 69 | – | 79 | 73 | – | |
| DPβ | 71 | 70 | 72 | 62 | 72 | 76 | – | 72 | – | |
| DOβ | 63 | 63 | 63 | 80 | 66 | 66 | 66 | – | – | |
| | | | | % Amino acid homology | | | | | | |

C. Exon 4 (TM domain)

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Eβ2 | – | 77 | 67 | 57 | 75 | 67 | 59 | 57 | – | |
| Eβ1 | 62 | – | 72 | 62 | 87 | 73 | 71 | 64 | – | |
| Aβ1 | 46 | 70 | – | 65 | 76 | 85 | 70 | 68 | – | |
| Aβ2 | 35 | 41 | 41 | – | 63 | 69 | 63 | 79 | – | % Nucleotide |
| DRβ1 | 59 | 89 | 73 | 41 | – | 78 | 68 | 68 | – | homology |
| DQβ | 51 | 73 | 81 | 46 | 76 | – | 71 | 75 | – | |
| DPβ | 43 | 57 | 59 | 43 | 54 | 59 | – | 63 | – | |
| DOβ | 41 | 51 | 57 | 73 | 51 | 57 | 43 | – | – | |
| | | | | % Amino acid homology | | | | | | |

[a]For source of sequences, see legend to Figure 6. DRβ3 sequence was from Spies *et al.* (1985).

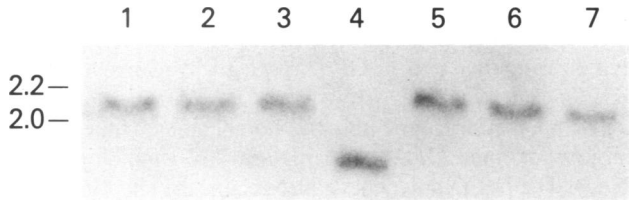


**Fig. 6.** *Eβ2* restriction fragment polymorphism. Lanes contain 10 μg of *Pst*I-digested liver DNA from the following mouse strains: **(1)** B10.P (*H-2*$^p$); **(2)** B10.S (*H-2*$^s$); **(3)** B10.RIII (*H-2*$^r$); **(4)** C57B1/6 (*H-2*$^b$); (5) BALB/c (*H-2*$^d$); (6) B10.A (*H-2*$^a$); **(7)** C3H (*H-2*$^k$). Hybridization probe 2 (Figure 3) was used for this Southern blot. DNA size markers correspond to *Hin*dIII digested λ DNA.

was analyzed by Southern blotting. One such blot of *Pst*I-digested DNA hybridized to *Eβ2*-specific probe 2 (see Figure 2) is shown in Figure 6. As can be seen, the only polymorphism is in the *b* allele. As *Pst*I sites are not present in the exons of the *d* or *a/k* allele (see Figure 1), this polymorphism does not necessarily reflect a difference in the *Eβ2*$^b$ coding sequence. Similarly, *Eco*RV-digested DNA showed *r* and *b* to be distinct from each

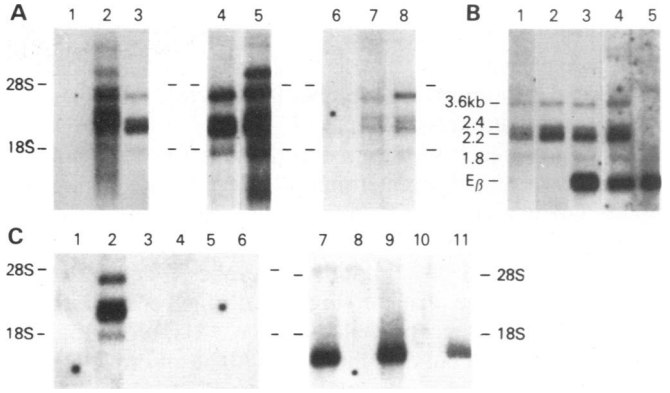


**Fig. 7.** *Eβ2* transcripts in various cells. **(A)** 10 μg of poly(A)$^+$ cytoplasmic RNA **(lanes 1–3** and **5)**, poly(A)$^+$ whole cellular RNA **(lane 4)**, or unselected total cellular RNA **(lanes 6–8)** from untransfected L cells **(lane 1)**, L-cell transfectant NB4.1 **(lane 2)**, CH-1 **(lane 3)**, MRL *lpr/lpr* spleen **(lane 4)**, LK35-2 **(lane 5)**, BALB/c non-adherent splenocytes **(lane 6)**, or LPS-stimulated BALB/c **(lane 7)** or C57B1/6 **(lane 8)** splenocytes were applied in each lane and hybridized to *Eβ2* probe 2 (Figure 3) following transfer onto nitrocellulose. **(B)** Each lane contains 5 μg of poly(A)$^+$ cytoplasmic RNA from CH-1. The transferred RNAs from **lanes 1–5** were hybridized to *Eβ2* probes 1–5 (Figure 3) respectively. **(C)** 1 μg **(lanes 7–11)**, 5 μg **(lane 2)** or 10 μg **(lanes 1** and **3–6)** of poly(A)$^+$ RNA were analyzed per lane. The RNAs were from an L-cell control **(lane 1)**, CH-1 **(lanes 2** and **7)**, unstimulated P388D$_1$ **(lanes 3** and **8)** or WEHI-3 **(lanes 5** and **10)**, and γ-interferon stimulated P388D$_1$ **(lanes 4** and **9)** or WEHI-3 **(lanes 6** and **11)**. The transferred RNAs from **lanes 1–6** were hybridized to *Eβ2* probe 2 and those from **lanes 7–11** to the *Eβ1* 3′ cDNA probe.

other and from the other five alleles, which showed a common pattern (not shown). *Eco*RI digests revealed no polymorphisms between the *b, d* and *k* alleles (not shown). Thus, consistent with the DNA sequence anlaysis, *Eβ2* shows fewer restriction fragment length polymorphisms for these *H-2* haplotypes than does *Eβ1* or *Aβ1* (Robinson *et al.*, 1983; Bukara *et al.*, 1985).

*Expression of Eβ2 specific transcripts*

To examine the transcriptional activity of the *Eβ2* gene, a probe from exon 2 of *Eβ2* (probe 2 in Figure 2) was utilized because, as predicted by sequence analysis, it did not cross-hybridize under high stringency to Southern blots of DNA from cosmids spanning all of the other seven known murine class II loci (not shown).

Northern blots hybridized to probe 2 demonstrate the presence of four transcripts of 3.6, 2.4, 2.2 and 1.8 kb in RNA from the B cell tumor CH-1, MRL *lpr/lpr* spleen cells, and in LPS-stimulated BALB/c and C57B1/6 spleen cells. These transcripts plus two others (4.1 and 1.0 kb) were found in RNA from the B cell hybridoma LK-35.2 (Figure 7a). The sizes of these transcripts differ from the 1.3-kb mature *Aβ1* and *Eβ1* mRNA and the steady-state level of the predominant species is about 1/20th that of the *Eβ1* mRNA when directly compared using a different probe that hybridizes to both *Eβ1* and *Eβ2* (see Figure 7b, lanes 3 and 4). The pattern of transcript sizes is neither allele specific nor tissue specific as similar sized transcripts were found in *H-2*$^b$ and *H-2*$^d$ spleen, the *H-2*$^k$ B-cell tumor CH-1, and in an L-cell transfected with cosH-2$^k$ 7.1 (NB4.1). As probe 2 clearly does not hybridize to *Eβ1* mRNA, its detection of RNA species in NB4.1 which co-migrate with the unique RNA species detected in the other cells establishes that these transcripts are encoded by the *Eβ2* gene. All of the *Eβ2* transcripts appear to be typical of mature mRNA molecules by several criteria. They are present in cytoplasmic mRNA preparations (Figure 7a, lanes 2, 3 and 5), and are specifically enriched by oligo(dT) purification to the

same extent as *Eβ1* transcripts (data not shown). Finally, the four *Eβ2* transcripts detected in the CH-1 cell line do not appear to contain sequences from introns 1, 2 and 3 as determined by Northern blot hybridization to probes from each of these predicted *Eβ2* introns (data not shown).

Probes from each of the predicted exons (probes 1–4, see Figure 2) were used to determine if each of these regions are represented in all of the *Eβ2* transcripts. Each probe detects the same four *Eβ2*-related mRNA species in the B cell tumor CH-1 (Figure 7b, lanes 1–4). In addition, probes 3 and 4 detect a typical *Eβ1* message. Probe 5 (Figure 2), which comes from the region having sequence homology to the 3′UT region of *Eβ1*, detects *Eβ1* mRNA but only the 1.8-kb *Eβ2* transcript. The size of this latter transcript is consistent with a mRNA species containing properly spliced sequences from exons 1–4, but continuing past the predicted 3′ end of exon 4 and ending with a poly(A) tract directed by the possible polyadenylation signal (AATAAT) noted above (Figure 3, line 27). That probe 5 does not detect the other three transcripts suggests that they represent splicing to an unidentified 3′UT region. Such alternative splicing of the 3′ end of other class II genes has been previously reported (Kappes *et al.*, 1984). Thus, the heterogeneity of *Eβ2* transcripts reflects, at least in part, alternative splicing, and may additionally be due to the use of alternative polyadenylation signals for some of the mRNA species.

Expression of mRNA for the conventional class II genes appears to be coordinately regulated and can be induced in several macrophage or monocyte tumors by γ-interferon (Steeg *et al.*, 1982; Paulhock-King *et al.*, 1985). In Figure 8c, RNA from two such lines, P388D1 and WEHI-3, clearly show induction of *Eβ1* transcription after culture with γ-interferon-containing conditioned medium. In striking contrast, no *Eβ2* transcripts were found in the RNA from these same Ia positive γ-interferon-treated cells.

## Discussion

The data presented here show the *Eβ2* gene to be similar in many regards to conventional class II β genes. Like these genes, the leader peptide, two external domains and the transmembrane region are encoded by separate exons. Each of these exons is appropriately flanked by splice sequences conforming to the AG/GT rule, potential lariat branch points are present, and the predicted fully spliced mRNA should have an open reading frame allowing its translation into a protein product. The predicted protein shows homology to other class II β chains, contains the functionally important hydrophobic leader and transmembrane sequences, has appropriately placed cysteines to allow for intrachain disulfide bonding necessary to define the two external β domains, and shows homology to immunoglobulin C-regions in its second domain. Sequences found in the promoter regions of other class II genes are conserved in the homologous *Eβ2* region and polyadenylated *Eβ2* specific transcripts containing sequences from the predicted exons can be detected in cytoplasmic RNA from several sources. As these transcripts do not contain intron sequences, it is most likely that they represent fully processed mRNA. Thus *Eβ2* does not appear to be a pseudogene and its transcripts should be translated into a class II β-like protein in those cells expressing *Eβ2* mRNA.

With respect to the potential for Eβ2 polypeptide expression on the cell membrane, several issues need to be considered. First, the exon(s) encoding the C terminus of *Eβ2* have not yet been identified. However, the sequence in the transmembrane exon following the region encoding the hydrophobic membrane spanning segment encodes a typical charged intracytoplasmic segment. Thus, although the full extent of the intracytoplasmic portion of Eβ2 is unknown, at a minimum, those residues necessary to anchor the Eβ2 protein in the cell membrane are present. Furthermore, since the *Eβ2* transcripts detected are polyadenylated, there must be at least one additional 3′ *Eβ2* exon which contains the 3′UT region and a signal for polyadenylation. Second, an unusual feature of the predicted Eβ2 protein chain is the lack of a site for N-linked glycosylation. However, other proteins which lack N-linked sugars are expressed on the cell membrane. These include the δ chain of the human T cell T3 complex (Van den Elsen *et al.*, 1984) and a truncated version of a mouse class I molecule having only the C-2 external domain, transmembrane and intracytoplasmic regions (McCluskey *et al.*, 1986). The absence of N-linked sugars on the Eβ2 protein should not, therefore, preclude its surface expression. Third, it is likely that any Eβ2 protein, like all other β polypeptides, requires a class II α chain for membrane expression (Germain and Malissen, 1986). Eβ2-like molecules have not been detected to date in immunoprecipitates obtained with anti-Aα or Eα specific monoclonal antibodies. We have also been unable as yet to detect cell surface Eα chain expression on L-cells co-transfected with the *Eα* and *Eβ2* genes using an Eα chain specific monoclonal antibody (14.4.4S). This was despite finding *Eβ2* and *Eα* mRNA in the primary transfectant pool (unpublished observations). Aside from the possibility that *Eβ2* does not encode an expressed protein product, the failure to detect Eβ2 protein in immunoprecipitation experiments or Eα protein in the transfection experiments may have been due to: (i) the low level of *Eβ2* mRNA expression; (ii) the loss of the relevant α chain epitopes following pairing with the Eβ2 chain; or (iii) a requirement for Eβ2 to be expressed with some as yet unidentified Eα2. In the latter case, the identification of a putative α chain partner for Eβ2 may require biochemical isolation of the Eβ2:α chain protein dimer.

Homologues for all previously studied murine class II genes have been found in man (Kaufman *et al.*, 1984; Widera and Flavell, 1985; Tonnelle *et al.*, 1985), suggesting the existence of a human *Eβ2* equivalent. Of all the identified human class II β genes, the best candidate is *DRβ3*. Both *Eβ2* and *DRβ3* lie in similar locations in their respective MHCs, have a high level of second domain homology to *Eβ1* and *DRβ1* (Table I and Denaro *et al.*, 1985), and show significantly lower levels of intraspecies polymorphism than *Eβ1* or *DRβ1* respectively (Sorrentino *et al.*, 1985; Spies *et al.*, 1985). However, the $\beta_1$ domains of the *Eβ2* and *Eβ1* polypeptides are only 50% homologous, while the comparable *DRβ3* and *DRβ1* domains are 76% homologous (see Table I). The DRβ3 polypeptide also retains the site for N-linked glycosylation seen in other β chains. Additional insight into the possible relationship between *Eβ2* and *DRβ3* should come from analysis of the tissue distribution and interferon sensitivity of *DRβ3* expression.

*Eβ2* differs from the conventional class II β genes in two important ways and these differences must be taken into account in considering possible functions of an *Eβ2*-containing molecule. First, *Eβ2* is relatively non-polymorphic. Second, *Eβ2* is expressed only at low levels in B cells and is not induced in macrophage or monocyte tumor lines by γ-interferon. Both of these features are shared by the newly described murine *Aβ2* gene and its human homologue *DOβ*, and are in marked contrast to *Aβ1*, *Eβ1*, *DQβ* and *DRβ1*. Thus, if Eβ2 or Aβ2 polypeptides were to pair with a polymorphic α chain, the low level of expression and restricted tissue distribution of these β chains suggest that the resultant class II molecule could have a distinct role even

as a participant in the conventional antigen-presentation pathway involving Ia. Perhaps of greater interest is the possibility that these atypical $\beta$ chains might pair with non-polymorphic $\alpha$ chains, such as $E\alpha$. In this case, one might speculate that such conserved molecules have a unique role in presentation of a specific subset of antigens or are molecules with a distinct role in differentiation or regulatory interactions of B lymphocytes.

The present study was undertaken to evaluate the potential role of $E\beta 2$ in certain immunological phenomena not attributable to conventional class II molecules. We conclude that $E\beta 2$, along with $A\beta 2$ and $DO\beta$, comprise a new and distinct family of MHC genes whose products and pattern of expression are sufficiently different from the prototypic class II genes to suggest that they play unique roles in the functioning of the immune system.

## Materials and methods

*Reagents*

Calf intestinal phosphatase and DNA polymerase I (whole enzyme) were obtained from Boehringer-Mannheim, and DNase I from Worthington. All other DNA-modifying enzymes and all restriction endonucleases were purchased from either Bethesda Research Laboratories or New England Biolabs. Deoxy- and dideoxy-nucleotides, and oligo(dT) were from Pharmacia P-L Biochemicals. [$\alpha$-$^{32}$P]dATP and [$\alpha$-$^{32}$P]dCTP were from Amersham. The 17-bp universal sequencing primer was purchased from Collaborative Research. Fetal calf serum (FCS) was obtained from Hyclone. All other tissue culture media were from M.A.Bioproducts or Gibco.

*Gene cloning and sequencing*

A BALB/c embryo ($H\text{-}2^d$) genomic DNA library was prepared by cloning *Mbo*I partially digested DNA into the *Bam*HI site of λ Charon 28 and was provided by J.Seidman and P.Leder. A cosmid library was similarly constructed in the *Bam*HI site of pOPF using DNA from the [B10.A × BW5147] ($H\text{-}2^a \times H\text{-}2^k$) T cell hybridoma 2B4 (Hedrick *et al.*, 1982) and was provided by D.Margulies. These libraries were screened with nick-translated cDNA probes as described (Maniatis *et al.*, 1982). Relevant restriction fragments were subcloned into pUC-9 or M13mp8 or 9. Sequencing of M13 subcloned fragments was done as described (Miller *et al.*, 1985). A synthetic oligonucleotide primer (5′-TAGCAACTCAGA-ACATG-3′) was made on an Applied Biosystem 380A DNA synthesizer by solid phase chemistry and was used for one template. The sequence of exons 2 and 3 of $E\beta 2^{a/k}$ was determined by sequencing homologous restriction fragments from $E\beta 2^d$ and $E\beta 2^{a/k}$, running like nucleotide lanes together (i.e. AAGGCCTT), and analyzing for any differences between adjacent lanes.

*Cell lines*

DAP.3 is a subclone of a thymidine kinase (tk) negative L-cell line (Margulies *et al.*, 1983), LK35-2 is a mouse B cell hybridoma ($H\text{-}2^d \times H\text{-}2^k$) (Kappler *et al.*, 1982), and CH-1 is a mouse B cell lymphoma ($H\text{-}2^a$) (Lynes *et al.*, 1978). P388$D_1$ is a BALB/c derived monocyte–macrophage line (Koren *et al.*, 1975), and WEHI-3 a BALB/c myelomonocyte line (Warner *et al.*, 1969).

*Cell culture and cell separation*

DAP.3 was maintained in DMEM supplemented with 10% FCS, 2 mM L-glutamine, 0.1 mM non-essential amino acids, and 10 μg/ml gentamicin (DMEM-10). P388$D_1$ and WEHI-3 were grown in the same media additionally supplemented with 130 μg/ml oxaloacetic acid, 80 μg/ml insulin, 55 μg/ml sodium pyruvate and NCTC-109 medium (1:10). For stimulation with γ-interferon, $50 \times 10^6$ P388$D_1$ or WEHI-3 were cultured for 72 h at 10% $CO_2$, 37°C in 200 ml medium supplemented with 10% v/v supernatant from 72 h Con-A stimulated rat spleen cells. LK35-2 and CH-1 were grown in RPMI supplemented with 10% FCS, 2 mM L-glutamine, 0.1 mM non-essential amino acids, $5 \times 10^{-5}$ M 2ME, and 10 μg/ ml gentamicin (RPMI-10).

Non-adherent BALB/c and C57B1/6 spleen cells were prepared as described (Mosier, 1967). $4 \times 10^7$ of each BALB/c or C57B1/6 non-adherent splenocytes were cultured for 60 h at 37°C, 5% $CO_2$ in 100 ml RPMI-10 containing 50 μg/ml LPS.

*DNA-mediated gene transfer*

DAP.3 was co-transfected with the indicated gene(s) plus the *tk* gene using the standard calcium phosphate precipitation technique as described (Margulies *et al.*, 1983). Selection was in DMEM-10 supplemented with hypoxanthine (15 μg/ml), aminopterin (0.2 μg/ml) and thymidine (5 μg/ml).

*RNA preparation and Northern blots*

Cytoplasmic RNA was extracted from CH-1, LK35-2, DAP.3 and DAP.3 transfectants as described (Mushinski *et al.*, 1980). Total cellular RNA from whole MRL *lpr/lpr* spleen, stimulated and fresh non-adherent splenocytes, and γ-interferon induced or non-induced P388$D_1$ and WEHI-3 was prepared by the guanidine isothiocyanate method (Chirgwin *et al.*, 1979). RNAs were fractionated in 1% agarose formaldehyde gels and blotted as described, except that the formaldehyde concentration in the gel was only 0.22 M (Maniatis *et al.*, 1982).

*DNA preparation and Southern blots*

Genomic DNA was prepared, gel electrophoresed, and blotted as described previously (Southern, 1975; Wilson *et al.*, 1979).

*Probes and hybridization*

cDNA probes were prepared by nick-translation of appropriate gel-purified restriction fragments. $E\beta 2$ gene probes (1–5) were prepared from M13 subclones by synthesizing a radiolabelled second DNA strand using the sequencing primer and limiting amounts of [$\alpha$-$^{32}$P]dCTP. This was followed by a chase with excess cold dCTP. The labelled insert was prepared by digestion with the appropriate restriction enzyme(s), polyacrylamide gel electrophoresis, and electroelution.

All hybridizations were carried out at 40°C in 10% dextran sulfate, 40% formamide, 4 × SSC, 1 × Denhardt's solution, 50 μg/ml salmon testis DNA, and $1-2 \times 10^6$ c.p.m./ml of radiolabelled probe. Blots were washed in three changes of 2 × SSC, 0.1% SDS at room temperature followed by three changes of 0.1 × SSC, 0.1% SDS at 60°C.

*Computer graphics*

Figures 4, 5 and 6 were drawn with the help of the DNA:DRAW program as described by Shapiro and Senapathy (1986).

## Acknowledgements

We would like to thank Eric Long, David Margulies, Michael Steinmetz and Georg Widera for providing valuable reagents and clones, and Eric Long, David Margulies, Jim Miller, Ronald Schwartz and Dinah Singer for reviewing the manuscript. We are indebted to Shirley Starnes for her excellent typographical assistance.

## References

Auffray,C., Lillie,J.W., Arnot,D., Grossberger,D., Kappes,D. and Strominger, J.L. (1984) *Nature*, **308**, 327–333.

Benacerraf,B. and Dorf,M.E. (1977) *Cold Spring Harbor Symp. Quant. Biol.*, **XLI**, 465–475.

Breathnach,R. and Chambon,P. (1981) *Annu. Rev. Biochem.*, **50**, 349–383.

Bukara,M., Vincek,V., Figueroa,F. and Klein,J. (1985) *Immunogenetics*, **21**, 569–579.

Cech,T.R. (1983) *Cell*, **34**, 713–716.

Chirgwin,J.A., Przybyla,A., MacDonald,R. and Rutter,W. (1979) *Biochemistry*, **18**, 5294–5299.

Choi,E., McIntyre,K., Germain,R.N. and Seidman,J.G. (1983) *Science*, **221**, 283–286.

Denaro,M., Gustafsson,K., Larhammar,D., Steinmetz,M., Peterson,P.A. and Rask,L. (1985) *Immunogenetics*, **21**, 613–616.

Denaro,M., Hammerling,U. and Peterson,P.A. (1984) *EMBO J.*, **3**, 2029–2032.

Devlin,J.J., Wake,C.T., Allen,H., Widera,G., Mellor,A., Fahrner,K., Weiss,E. and Flavell,R.A. (1984) In Cantor,H., Chess,L. and Sercarz,E. (eds), *Regulation of the Immune System*. Alan R.Liss, New York, pp. 57–66.

Dognin,M.J., Lauwereys,M. and Strossberg,A.D. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4031–4035.

Folsom,V., Gay,D. and Tonegawa,S. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1678–1682.

Germain,R.N. and Malissen,B. (1986) In Paul,W.E., Fathmas,C.G. and Metzger,H. (eds) *Annual Review of Immunology*. Annual Reviews Inc., Palo Alto, CA, pp. 281–315.

Germain,R.N., Ashwell,J.D., Lechler,R.I., Margulies,D.H., Nickerson,K.M., Suzuki,G. and Tou,J.Y.L. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2940–2944.

Hedrick,S.M., Matis,L.A., Hecht,T.T., Samelson,L.E., Longo,D.L., Heber-Katz,E. and Schwartz,R.H. (1982) *Cell*, **30**, 141–152.

Hiramatsu,K., Ochi,A., Miyatani,S., Segawa,A. and Tada,T. (1982) *Nature*, **296**, 666–668.

Hood,L., Steinmetz,M. and Malissen,B. (1983) *Annu. Rev. Immunol.*, **1**, 529–568.

Kappes,D.J., Arnot,D., Okada,K. and Strominger,J. (1984) *EMBO J.*, **3**, 2985–2993.

Kappler,J., White,J., Wegmann,D., Mustain,E. and Marrack,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 3504–3607.

Kaufman,J.F., Auffray,C., Korman,A.J., Shackleford,D.A. and Strominger,J. (1984) *Cell*, **36**, 1–13.

Kelly,A. and Trowsdale,J. (1985) *Nucleic Acids Res.*, **13**, 1607–1621.

Klein,J., Figueroa,F. and David,C.S. (1983) *Immunogenetics*, **17**, 553–596.

Koren,H.S., Handwerger,B.S. and Wunderlich,J.R. (1975) *J. Immunol.*, **114**, 894–897.

Kozak,M. (1984) *Nucleic Acids Res.*, **12**, 857–872.
Larhammar,D., Hammerling,U., Denaro,M., Lund,T., Flavell,R.A., Rask,L. and Peterson,P.A. (1983a) *Cell*, **34**, 179–188.
Larhammar,D., Hyldig-Nielson,J.J., Servenius,B., Andersson,G., Rask,L. and Peterson,P.A. (1983b) *Proc. Natl. Acad. Sci. USA*, **80**, 7313–7317.
Larhammar,D., Hammerling,U., Rask,L. and Peterson,P.A. (1985a) *J. Biol. Chem.*, **260**, 14111–14119.
Larhammar,D., Servenius,B., Rask,L. and Peterson,P.A. (1985b) *Proc. Natl. Acad. Sci. USA*, **82**, 1475–1479.
Lechler,R.I., Ronchese,F., Braunstein,N.S. and Germain,R.N. (1986) *J. Exp. Med.*, **163**, 678–696.
Long,E.O., Wake,C.T., Gorski,J. and Mach,B. (1983) *EMBO J.*, **2**, 389–394.
Lynes,M.A., Lanier,L.L., Babcock,G.F., Wettstein,P.J. and Houghton,G. (1978) *J. Immunol.*, **121**, 2352–2357.
Malissen,M., Hunkapiller,T. and Hood,L. (1983) *Science*, **221**, 750–754.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, New York.
Margulies,D.H., Evans,G.A., Ozato,K., Camirini-Otero,R.D., Tanaka,K., Appella,E. and Seidman,J.G. (1983) *J. Immunol.*, **130**, 463–470.
Mathis,D.J., Benoist,C.O., Williams,V.E. II, Kanter,M.R. and McDevitt,H.O. (1983) *Cell*, **32**, 745–754.
Mengle-Gaw,L. and McDevitt,H.O. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7621–7625.
McCluskey,J., Bluestone,J.A., Coligan,J.E., Maloy,W.L. and Margulies,D.H. (1986) *J. Immunol.*, **136**, 1–10.
Miller,J., Malek,T.R., Leonard,W.J., Greene,W.C., Shevach,E.M. and Germain, R.N. (1985) *J. Immunol.*, **134**, 4212–4217.
Mosier,D.E. (1967) *Science*, **158**, 1573–1575.
Mushinski,J.F., Blattner,F.R., Owens,J.D., Finkelman,F.D., Kessler,S.V., Fitzmaurice,L., Potter,M. and Tucker,P.W. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 7405–7409.
Okada,K., Boss,J.M., Prentice,H., Spies,T., Mengler,R., Auffray,C., Lillie,J., Grossberger,D. and Strominger,J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3410–3414.
Paulhock-King,D., Sizer,K.C., Freund,Y.R., Jones,P.P. and Parnes,J.R. (1985) *J. Immunol.*, **135**, 632–636.
Proudfoot,N.J. and Brownlee,G.G. (1976) *Nature*, **203**, 211–214.
Reed,R. and Maniatis,T. (1985) *Cell*, **41**, 95–105.
Rich,A., Nordheim,A. and Wang,A.H.-J. (1984) *Annu. Rev. Biochem.*, **53**, 791–846.
Robinson,R.R., Germain,R.N., McKean,D.J., Mescher,M. and Seidmann,J. (1983) *J. Immunol.*, **131**, 2025–2031.
Saito,H., Maki,R.A., Clayton,L.K. and Tonegawa,S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5520–5524.
Schenning,L., Larhammar,D., Bill,P., Wiman,K., Jonsson,A., Rask,L. and Peterson,P.A. (1984) *EMBO J.*, **3**, 447–452.
Schwartz,R.H. (1986) *Adv. Immunol.*, **38**, 31–201.
Shapiro,M.B. and Senapathy,P. (1986) *Nucleic Acids Res.*, **14**, 65–73.
Sorrentino,R., Lillie,J. and Strominger,J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3794–3798.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
Spielman,S.R., Lee,J., Bodmer,W.F., Bodmer,J.G. and Trowsdale,J. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 3461–3465.
Spies,T., Sorrentino,R., Boss,J.M., Okada,K. and Strominger,J.L. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5165–5169.
Steeg,P.S., Moore,R.N., Johnson,H.M. and Oppenheim,J.J. (1982) *J. Exp. Med.*, **156**, 1780–1793.
Steinmetz,M., Malissen,M., Hood,L., Orn,A., Maki,R.A., Dastoornikoo,G.R., Stephan,D., Gibb,E. and Romaniuk,R. (1984) *EMBO J.*, **3**, 2995–3003.
Steinmetz,M., Minard,K., Horvath,S., McNicholas,J., Frelinger,J., Wake,C., Long,E., Mach,B. and Hood,L. (1982) *Nature*, **300**, 35–42.
Tonnelle,C., DeMars,R. and Long,E.O. (1985) *EMBO J.*, **4**, 2839–2847.
Van den Elsen,P., Shipley,B., Borst,J., Coligan,J.E., Markham,A.F., Orkin,S. and Terhorst,C. (1984) *Nature*, **312**, 413–418.
Wake,C.T. and Flavell,R.A. (1985) *Cell*, **42**, 623–628.
Warner,N.L., Moore,M.A.S. and Metcalfe,D.J. (1969) *J. Natl. Cancer Inst.*, **43**, 963–982.
Widera,G. and Flavell,R.A. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 5500–5504.
Wilson,R., Miller,J. and Storb,U. (1979) *Biochemistry*, **18**, 5013–5021.

*Received on 2 June 1986; revised on 4 August 1986*