The EMBO Journal vol.5 no.9 pp.2123-2129, 1986

Molecular characterization of a meiotic recombinational hotspot enhancing homologous equal crossing-over

Yasushi Uematsu, Hansruedi Kiefer, Reinhard Schulze, Kirsten Fischer-Lindahl[1] and Michael Steinmetz

Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland

[1]Present address: Howard Hughes Medical Institute, 5323 Harry Hines Blvd., Y5-310, Dallas, Texas 75235-9050, USA

Communicated by M.Steinmetz

We have cloned and sequenced a meiotic recombinational hotspot between the $A_{\beta 3}$ and $A_{\beta 2}$ genes in the major histocompatibility complex (MHC) of the mouse. This recombinational hotspot in the *Mus musculus castaneus cas3* haplotype was previously localized to a region of 9.5 kb of DNA in which five independent crossing-over events occurred at the unusually high frequency of 0.6%. Aside from *cas3*, the hotspot appears to be absent in many other MHC haplotypes. We have now confined the five recombinational breakpoints to a stretch of 3.5 kb of DNA. From the nucleotide sequence around the recombinational breakpoints, determined in the parental *cas3* and *b* haplotypes as well as for two recombinant haplotypes, we show that the two recombinant haplotypes were generated by homologous equal crossing-over and place the breakpoints within two non-overlapping stretches of 10 and 36 bp, respectively. Comparison of the DNA sequences of the hotspot-positive *cas3* and the hotspot-negative *b* haplotypes reveals a number of differences, in particular, a CAGA-repeat sequence which is present in CAS3 in six, but only four copies in C57BL/6 DNA. This repeat sequence is reminiscent of one in a previously characterized hotspot in the E_{β} gene.

Key words: major histocompatibility complex/I region/meiotic recombination/*Mus musculus castaneus*/evolution

Introduction

The major histocompatibility complex (MHC) of vertebrates is a genetic region encoding cell surface glycoproteins that present foreign antigens to the T lympocytes of the immune system. Over the last 5 years this region has been studied at the molecular level in several species. In the mouse, where our knowledge is most advanced, the MHC has been shown to contain at least 46 genes spread over at least 1600 kb of DNA (for reviews see Flavell *et al.*, 1985; Steinmetz, 1986).

Molecular maps, available for extended portions of the mouse MHC (Steinmetz *et al.*, 1982,1986; Stephan *et al.*, 1986), are not congruent with classical genetic maps, deduced from recombination frequencies between marker loci. Meiotic recombination in the mouse MHC is not random, but occurs frequently in certain areas, termed recombinational hotspots. So far, four such recombinational hotspots have been identified, located close to the $A_{\beta 3}$ gene (Steinmetz *et al.*, 1986), between the $A_{\beta 3}$ and $A_{\beta 2}$ genes (Steinmetz *et al.*, 1986), within the E_{β} gene (Steinmetz *et al.*, 1982; Kobori *et al.*, 1984; Begovich and Jones, 1985; Saha and Cullen, 1986) and between the $E_{\beta 2}$ and E_{α} genes (Lafuse *et al.*, 1986) (Figure 1). Furthermore, certain recombinational hotspots are present in some haplotypes and absent in others. Thus, the hotspots close to the $A_{\beta 3}$ gene and between the $A_{\beta 3}$ and $A_{\beta 2}$ genes have been found in the *cas4* and *cas3* haplotype MHCs, respectively, but appear to be absent in the *b, d, k, q* and *s* haplotype MHCs (Steinmetz *et al.*, 1986). Recombinational hotspots vary dramatically in activity. For instance, recombination occurs at the hotspot in the E_{β} gene with a frequency of ~0.1%, while a frequency of 1.5% is found at the hotspot close to the $A_{\beta 3}$ gene (see Steinmetz *et al.*, 1986).

It is not clear at present whether recombinational hotspots are particularly abundant in the MHC as compared to other genetic regions. It has been pointed out that they could play an important role in maintaining polymorphism at certain MHC loci (Steinmetz, 1986). Meoitic recombinational hotspots have also been identified in other regions of the mammalian genome, e.g. upstream of the human β-globin gene (Orkin and Kazazian, 1984) and upstream and downstream of the human insulin gene (Lebo *et al.*, 1983). However, none of these hotspots has been characterized in the same detail as the ones in the mouse MHC.

The molecular basis of recombinational hotspots is poorly understood. DNA sequence information exists for the recombinational hotspot in the E_{β} gene. Several investigators have noticed the presence of a CAGG tetramer, repeated in tandem ~16-20 times (Saito *et al.*, 1983; Widera and Flavell, 1984) within a DNA segment containing the recombinational breakpoints (Begovich and Jones, 1985; Kobori *et al.*, 1986; Lafuse *et al.*, 1986; Steinmetz *et al.*, 1986). As discussed (Steinmetz *et al.*, 1986), this repeat sequence shows some homology to the Chi sequence (GCTGGTGG) (Smith *et al.*, 1981) of phage λ, known to be a recombinational hotspot, to the human hypervariable mini-satellite sequence (GGAGGTGGGCAGGAXG) (Jeffreys *et al.*, 1985), believed to be one, and could potentially form Z-DNA, which appears to constitute a recognition sequence (Kmiec *et al.*, 1985) for the *rec1* recombinase from lower eukaryotes. These observations have been used as arguments to suggest that the CAGG repeat forms the structural basis of the recombinational hotspot in the E_{β} gene.

We have now cloned and sequenced a region located between the $A_{\beta 3}$ and $A_{\beta 2}$ genes in which five recombinational breakpoints have been mapped. These independent recombinational events occurred at the unusually high frequency of 0.6% (Steinmetz *et al.*, 1986). The region was sequenced in the MCH haplotype *cas3* which contains this hotspot (called the $A_{\beta 3}/A_{\beta 2}$ hotspot from now on) and in the *b* haplotype, where it is absent. We report here the precise localization of two of the five recombinational breakpoints within the region sequenced and discuss the structural information with respect to the molecular basis of this and other recombinational hotspots.

Results

The $A_{\beta 3}/A_{\beta 2}$ hotspot has been cloned from the parental mouse strains, CAS3 and C57BL/6, and from two recombinant strains, CAS3(R0) and CAS3(R18)

Genomic DNA libraries were constructed in the phage λ vector

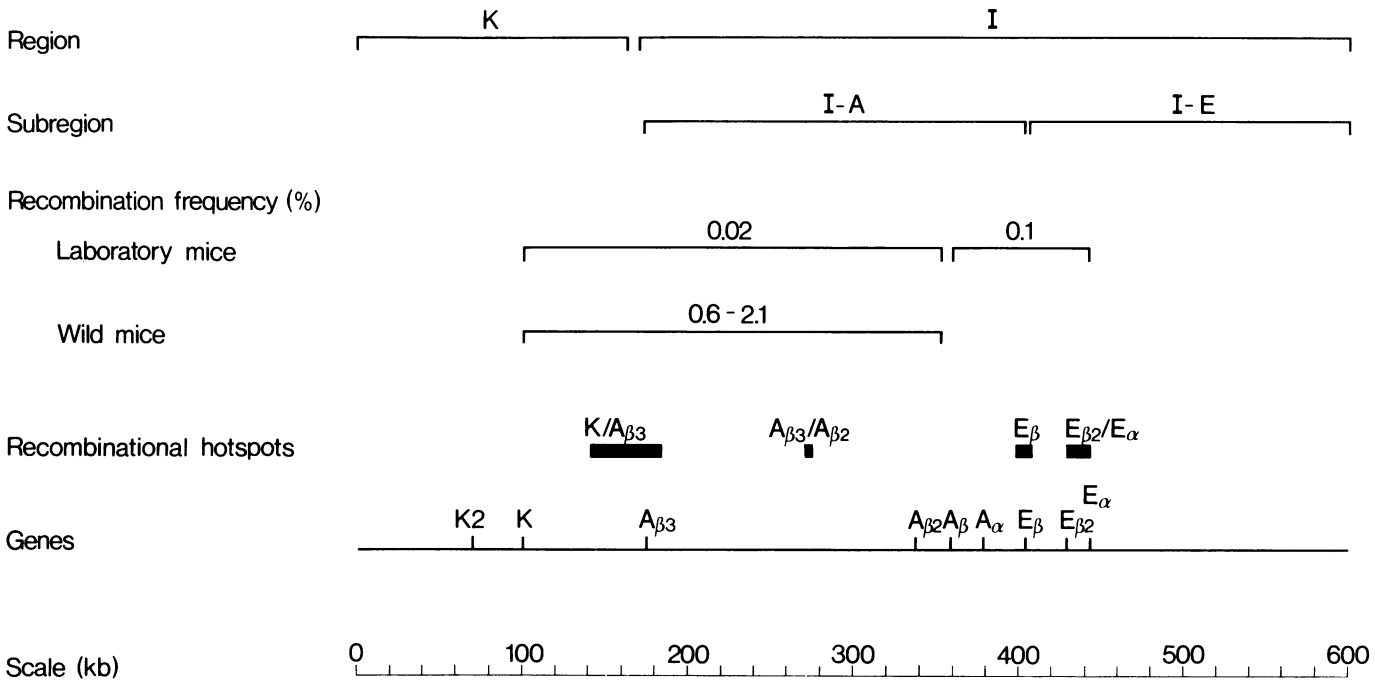

Fig. 1. Map of the proximal part of the MHC of the BALB/c mouse. Borderlines between genetic regions and subregions are defined by recombinational hotspots. The K/$A_{\beta 3}$ separates the *K* from the *I* region while the E_β hotspot separates the *I-A* from the *I-E* subregion. Recombination frequencies between marker loci were taken from Klein (1975), Shiroishi *et al.* (1982) and Steinmetz *et al.* (1986). Genes, indicated by vertical lines, are not drawn to scale. Recombinational hotspots so far identified are named according to the genes between which they map. Their occurrence is haplotype dependent; the black bars indicate the length of the DNA segments to which they could be confined (see text).

EMBL3 using liver DNAs from mouse strains CAS3, C57BL/6, CAS3(R0) and CAS3(R18). The libraries were screened with three single copy restriction fragments from a BALB/c cosmid clone spanning the $A_{\beta 3}/A_{\beta 2}$ recombinational hotspot (Steinmetz *et al.*, 1986) (see Figure 2). Four λ clones were isolated from the CAS3, two from the C57BL/6 and one each from the CAS3(R0) and CAS3(R18) libraries. These clones were then mapped with 13 restriction enzymes (not shown) and by comparison with known restriction maps (Steinmetz *et al.*, 1986) for BALB/c, C57BL/6 and CAS3 DNAs across the recombinational hotspot, they could be aligned as shown in Figure 2.

The cloned DNA fragments confirm the polymorphic restriction sites previously determined by Southern blot analysis and used to localize the recombinational breakpoints in the five CAS3 recombinants to a region of 9.5 kb of DNA (Steinmetz *et al.*, 1986). They also confirm that the sequence organization across the $A_{\beta 2}/A_{\beta 2}$ recombinational hotspot is not grossly different in BALB/c, C57BL/6 and CAS3 DNA. One clone, however, isolated from the CAS3 library, contains a DNA segment that overlaps only partially with the previously determined restriction map. We think that this clone, CAS3.12, which actually spans the gap between λ clones CAS3.31 and CAS3.32, was derived by ligation of two non-contiguous DNA fragments during construction of the library and therefore represents a cloning artefact.

Southern blot analyses of polymorphic restriction sites confine all five recombinational events to 3.5 kb of DNA

To map the five recombinational breakpoints more precisely, we analysed additional polymorphic restriction sites identified by restriction mapping and DNA sequence determination (see below) of the C57BL/6 and CAS3 clones. This analysis localized the recombinational breakpoint in CAS3(R25) to a region of 0.7 kb of DNA defined by polymorphic *Taq*I and *Msp*I sites (Figure 3 and Table I). Interestingly, the other four recombinational breakpoints were not within this stretch of DNA but adjacent to it within two overlapping regions of 0.5 kb and 2.8 kb of DNA, respectively. As evident from Figure 3 and Table I, the recombinational breakpoints in CAS3(R0), CAS3(R11) and CAS3(R18) are between the polymorphic *Msp*I and *Sph*I sites, 0.5 kb apart, while the recombinational breakpoint in CAS3(R23) is between the same polymorphic *Msp*I site and a polymorphic *Bgl*I site, 2.8 kb further downstream. As shown in Figure 3, hybridization with probe B identified a 2.1- and a 1.6-kb *Msp*I fragment in CAS3(R18) DNA. Also, a slight size difference between the 2.1 kb *Msp*I fragments of CAS3(R18) and CAS3(R25) is seen as compared to the one in C57BL/6 DNA. We think that *Msp*I identifies a heterozygosity in CAS3(R18) and that the size difference can be used to confine further the breakpoint in CAS3(R25) as discussed in the legend to Figure 3.

DNA sequence determination of the recombinational hotspot reveals multiple differences between C57BL/6 and CAS3 DNA

To search for unusual sequence elements, present in CAS3 but not in C57BL/6 DNA, which could potentially form the structural basis of the $A_{\beta 3}/A_{\beta 2}$ recombinational hotspot, we subcloned the *Sal*I – *Sac*I restriction fragments from phage clones B6.11 and CAS3.31 (Figure 2) into M13 vector DNAs and determined their sequences using the dideoxy-chain termination method. As shown in Figure 4, multiple differences, including nucleotide substitutions and insertions or deletions, distinguish the *cas3* from the *b* haplotype DNA at the recombinational hotspot. Some of the sequence differences identified the polymorhic restriction sites used to confine the recombinational breakpoints (see above).

The region sequenced was then used to screen a DNA databank (Microgenie, Beckman Instruments) for homologous sequences and searched by computer for the presence of specific sequence elements. The homology search revealed a mouse B1 repeat

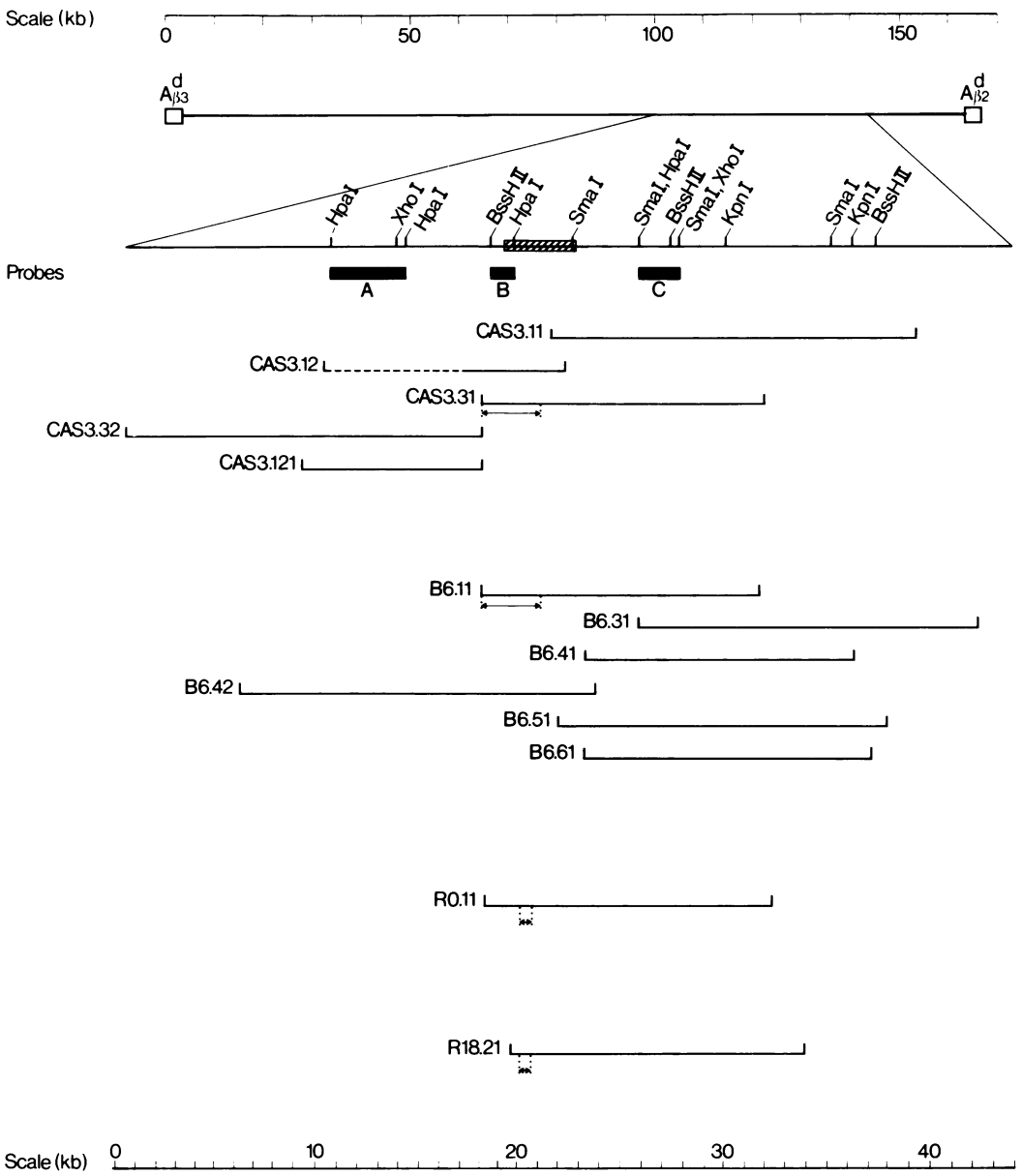

Fig. 2. Genomic clones, isolated from CAS3, C57BL/6, CAS3(R0) and CAS3(R18) DNA, span the $A_{\beta 3}/A_{\beta 2}$ recombinational hotspot. The $A_{\beta 3}$ and $A_{\beta 2}$ genes identified in the BALB/c mouse (*d* haplotype) are indicated by open boxes. A blow-up shows the portion of the DNA which is represented in the phage λ clones (brackets), isolated from the four different mice [CAS3, B6 for C57BL/6, R0 for CAS3(R0) and R18 for CAS3(R18)] and contains the recombinational hotspot (hatched box) confined by polymorphic restriction sites (see text). Some restriction sites, non-polymorphic in the five strains shown, are indicated. Hybridization probes are: **A**, 3.5-kb *Hpa*I fragment; **B**, 1.5-kb *Bss*HII-*Hpa*I fragment; **C**, 2-kb *Sma*I fragment. All probes were isolated from cosmid cosH-2^dII 5.10 (Steinmetz *et al.*, 1986). The stippled portion of clone CAS3.12 is probably a cloning artefact (see text). Arrows indicate the region sequenced in the respective clones.

(Krayev *et al.*, 1980) located between positions 149 and 307 (Figure 4) and a mouse B2 repeat (Krayev *et al.*, 1982) between positions 37 and 141 (Figure 4). Furthermore, 72% homology was found between a 36-bp stretch (positions 2335 – 2370), which contains the breakpoint in the recombinant strain CAS3(R18) (see below) and a portion of xenotropic solitary LTR sequences present in the mouse genome in ~500 copies (Wirth *et al.*, 1983).

The search for specific DNA sequence elements revealed no palindromes longer than eight nucleotides, no unusually long AT- or GC-rich regions and no stretches of alternating purine-pyrimidine residues longer than 13 nucleotides. However, two direct repeat sequences were identified: a 57-bp sequence present once in CAS3 and in two tandemly arranged copies in C57BL/6 DNA (starting at position 1495; Figure 4) and a TCTG tetramer repeated six times in CAS3 (starting at position 1921; Figure 4) and four times in C57BL/6 DNA. Interestingly, the complementary strand of this repeat, $(CAGA)_{4-6}$, is reminiscent of the CAGG repeat sequence in the E_β recombinational hotspot. No significant homology of this repeat sequence to *Chi* and the core sequence of the human hypervariable mini-satellite sequence exists. Below we will discuss the possible significance of the CAGA repeat, the 57-bp repeat and the LTR-like core sequence for the $A_{\beta 3}/A^{\beta}{}_2$ recombinational hotspot.

Sequence comparisons map two meiotic recombinational breakpoints within two non-overlapping regions of 10 and 36 bp, respectively

To confirm the localization of the recombinational breakpoints in CAS3(R0) and CAS3(R18) between the polymorphic *Msp*I

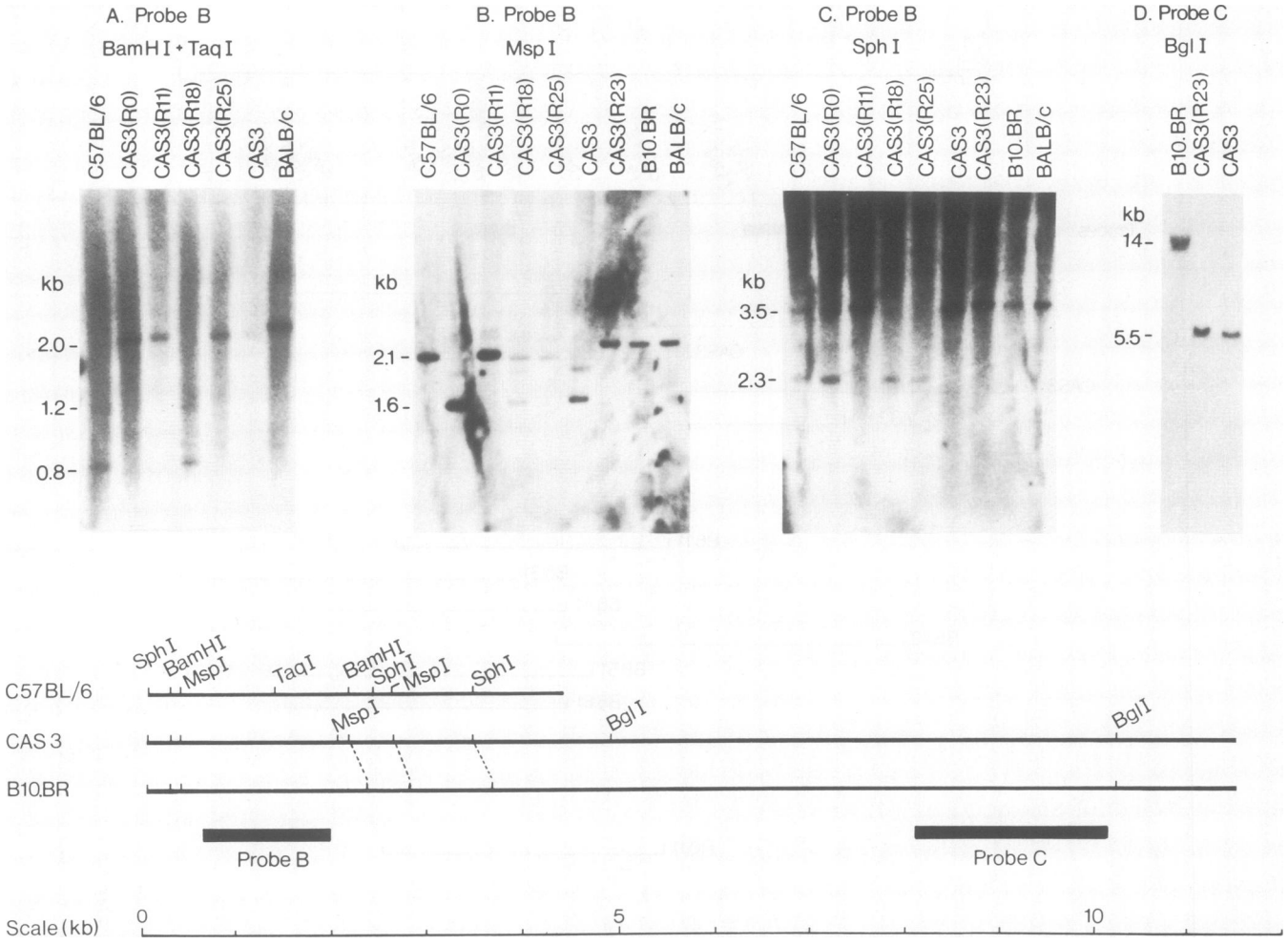

Fig. 3. Southern blot analyses of polymorphic restriction sites map the recombination breakpoints in CAS3(R0), CAS3(R11), CAS3(R18), CAS3(R23) and CAS3(R25) to a stretch of 3.5 kb of DNA. Hybridization probes and enzymes used for DNA digestions are indicated. Sizes of restriction fragments identified by hybridization are given in kbp. At the bottom, restriction maps of CAS3, C57BL/6 and B10.BR DNAs are given for the enzymes used and the locations of the hybridization probes are indicated. The analysis of the *Msp*I-digest of CAS3(R18) DNA with probe B reveals a CAS3-specific *Msp*I fragment and a fragment which co-migrates with the CAS3(R25) *Msp*I fragment indicating that CAS3(R18) is heterozygous. The sequence shown in Figure 4 for CAS3(R18) DNA contains a T at position 1959 like C57BL/6 DNA and for this reason is not cut at this position by *Msp*I while CAS3 DNA is. The CAS3(R18) chromosome sequenced is therefore identified by the larger *Msp*I fragment in the Southern blot. The smaller *Msp*I fragment presumably identifies the homologous chromosome containing an *Msp*I site at position 1959. The origin of the T nucleotide in the CAS3(R18) chromosome sequenced, which shows CAS3-specific nucleotides upstream and downstream of this position, is unknown. It could have been generated by a triple recombination event by point mutation, or by gene conversion when the (R18 × C57BL/6)F_1 was inter-crossed. The fact that one of the two CAS3(R18) and the CAS3(R25) *Msp*I fragments are slightly smaller than the C57BL/6 *Msp*I fragment is probably due to the presence of a single 57-bp repeat at position 1495 (see Figure 4 and text) as in CAS3 DNA. This is in agreement with the placement of the recombination breakpoint in CAS3(R18) (see text). The presence of a single 57-bp repeat in CAS3(R25) would further confine the recombination breakpoint in between the 57-bp repeat and the missing *Msp*I site at position 1959.

and *Sph*I sites and to map them as precisely as possible we subcloned and sequenced this region from the phage clones B6.11 and CAS3.31 (see Figure 2). As shown in Figure 4, the nucleotide sequence of CAS3(R0) DNA is identical to CAS3 DNA up to position 2171 and identical to C57BL/6 DNA from position 2162 onwards, in agreement with the Southern blot hybridization analysis. This sequence comparison localizes the recombinational breakpoint in CAS3(R0) to within 10 bp. A similar comparison using the CAS3(R18) nucleotide sequence places the recombinational breakpoint within a stretch of 36 bp, ~150 bp downstream of the recombinational point in CAS3(R0) (Figure 4). The possible origin of the T nucleotide at position 1959 in the CAS3(R18) sequence, which does not match the expected sequence of the *cas3* haplotype is discussed in the legend to Figure 3.

Table I. MHC alleles of inbred and recombinant mice used[a]

Strain	Genetic locus		
	K	I-A	I-E
B10.BR	*k*	*k*	*k*
C57BL/6	*b*	*b*	*b*
CAS3	*cas3*	*cas3*	*cas3*
CAS3(R0)	*cas3* |	*b*	*b*
CAS3(R11)	*b* |	*cas3*	*cas3*
CAS3(R18)	*cas3* |	*b*	*b*
CAS3(R23)	*k* |	*cas3*	*cas3*
CAS3(R25)	*cas3* |	*b*	*b*

[a]Alleles are listed according to Klein *et al.* (1983) and Steinmetz *et al.* (1986). Locations of recombination events are indicated by vertical lines.

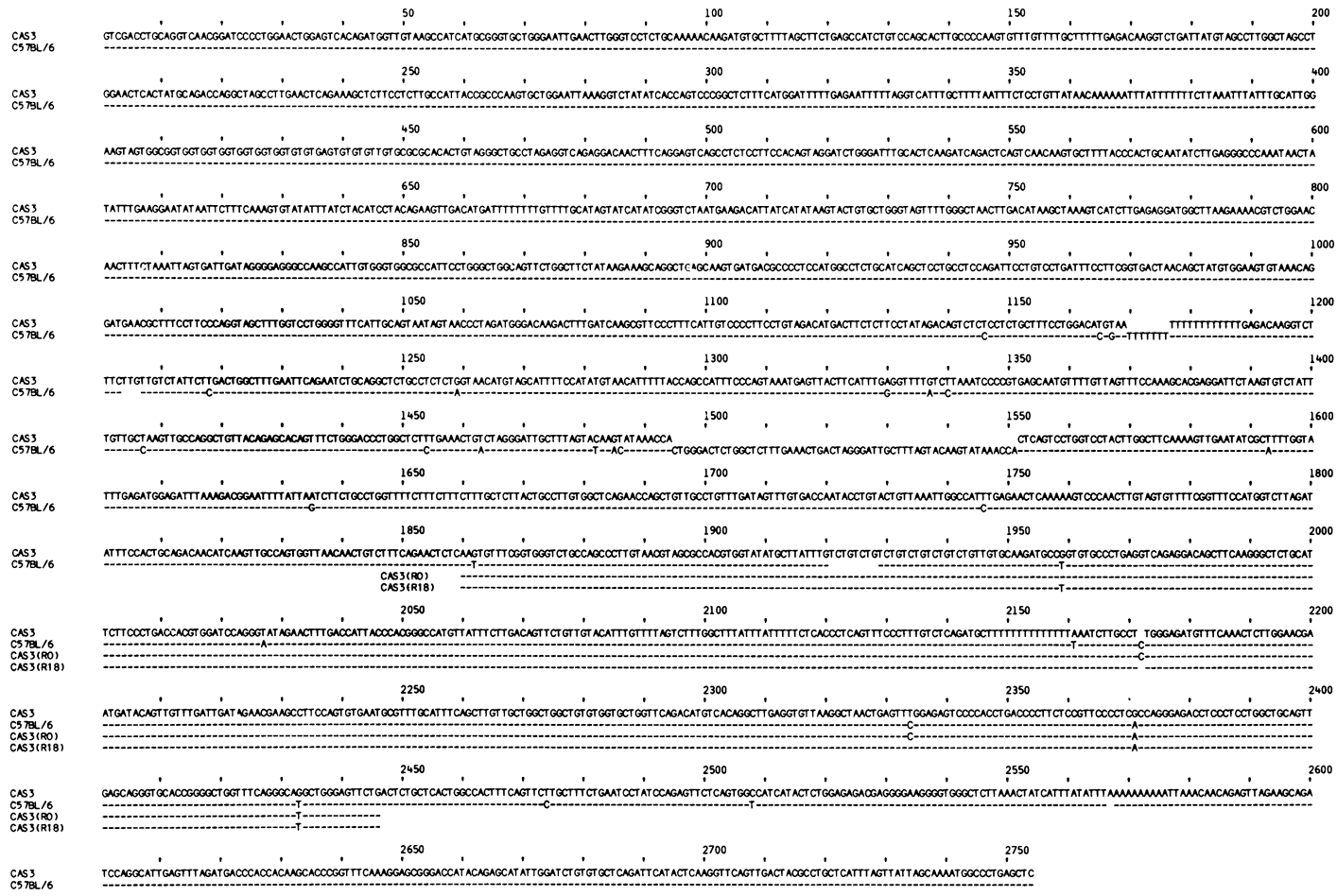

Fig. 4. Nucleotide sequence of a stretch of 2.8 kb of DNA spanning the recombination breakpoints in the CAS3(R0), CAS3(R11), CAS3(R18) and CAS3(R25) recombinant mouse strains. The nucleotide sequences were determined as described in Materials and methods for the two parental *cas3* and *b* haplotypes (in mouse strains CAS3 and C57BL/6, respectively) and for the two recombinant mouse strains CAS3(R0) and CAS3(R18) (only from positions 1860–2446). Nucleotides identical to the CAS3 sequence, shown at the top, are indicated by horizontal lines; deletions are shown as gaps; substitutions and insertions are listed. The polymorphic nucleotides of the restriction sites for *Taq*I, *Msp*I and *Sph*I, which confine the breakpoints in CAS3(R11) and CAS3(R25), are located at positions 1218, 1959 and 2433, respectively. For the inconsistent T nucleotide at position 1959 in CAS3(R18) see legend to Figure 3.

Thus, the two recombinational events have occurred at distinct but spatially close positions. Furthermore, both are homologous equal recombinational events; that is, there are no additions or removals of nucleotides associated with the meiotic recombination.

Discussion

Characteristics of meiotic recombinational hotspots in the MHC

On the basis of the present findings and those made earlier (Shiroishi *et al.*, 1982; Steinmetz *et al.*, 1982,1986; Begovich and Jones, 1985; Kobori *et al.*, 1984; Lafuse *et al.*, 1986) meiotic recombinational hotspots in the MHC are characterized by several features.

By definition, recombination occurs at hotspots with an unusually high frequency. The recombination frequency within the 9.5 kb of DNA around the E_β gene, to which 12 recombinational events have been mapped so far, is ~0.1% which contrasts with an expected frequency of 0.005% for a 10-kb stretch of DNA, assuming that recombination in the mouse genome would be completely random [1600 cM (genetic length mouse genome) × 10 kb/3 × 10^6 kb (physical length mouse genome)]. Thus, the hotspot in the E_β gene enhances recombination at least 20-fold. An even stronger enhancement of recombination is found in the $A_{\beta 3}/A_{\beta 2}$ recombinational hotspot. As shown here, the five recombinational events, isolated at a frequency of 0.6%, could be confined to 3.5 kb of DNA. Thus, at least a 300-fold enhancement is found (3 × 10^6 kb × 0.6 cM/3.5 kb × 1600 cM).

Even though the recombinational events are highly clustered, they do not occur at the same position. As shown here at least three of the five events at the $A_{\beta 3}/A_{\beta 2}$ hotspot have occurred at different positions, separated by at least 200 and 160 bp, respectively.

Recombination at hotspots is homologous and equal down to the nucleotide level. No evidence for either deletion or insertion of nucleotides was found at the two recombinational breakpoints in CAS3(R0) and CAS3(R18), the only ones that have so far been characterized by DNA sequencing.

The presence or absence of certain recombinational hotspots depends on the haplotype. Thus, the $A_{\beta 3}/A_{\beta 2}$ hotspot is observed in crosses between *cas3* and either *b* or *k* MHC haplotypes but not in crosses between *b* and *k* haplotypes. For instance, the recombination of the *b* and *k* haplotypes that gave rise to the B10.MBR mouse chromosome is outside of the $A_{\beta 3}/A_{\beta 2}$ recombinational hotspot (Steinmetz *et al.*, 1986). Also, the E_β recombinational hotspot appears to be present in several MHC haplotypes but not in all (Lafuse *et al.*, 1986).

The enhancement of meiotic recombination is not dependent

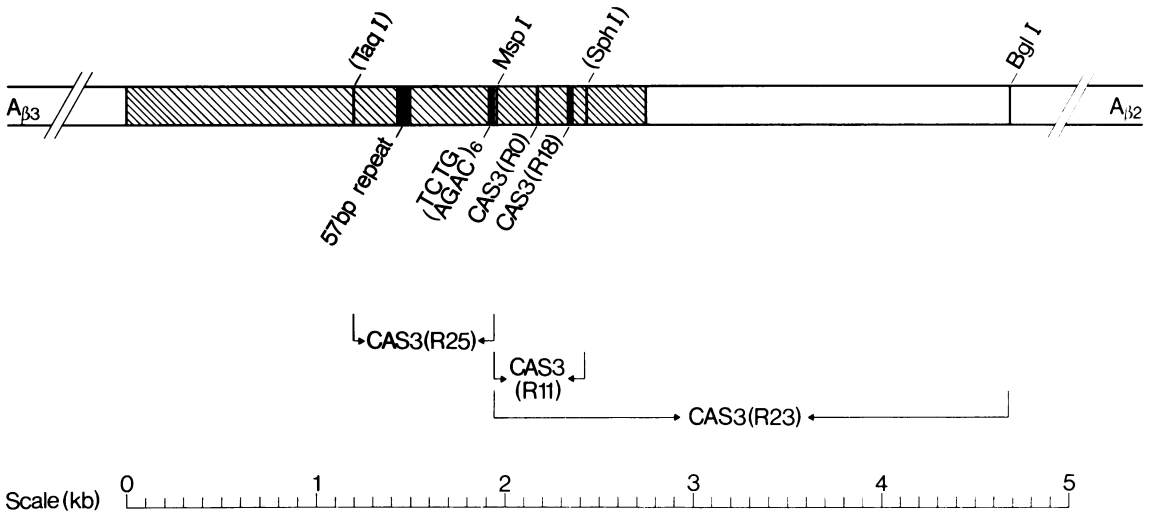

Fig. 5. Location of three specific sequence elements with respect to the breakpoints in the five recombinant strains. The three specific elements, the 57-bp repeat, the $(CAGA)_6$ repeat and the LTR core sequence are discussed in the text. The *Msp*I and *Bgl*I restriction sites are present in CAS3 but not in C57BL/6 DNA, while the *Taq*I and *Sph*I sites are specific for C57BL/6.

on the pairing of wild-derived MHC haplotypes (e.g. *cas3*, *cas4*) with those of the laboratory inbred strains (e.g. *b*, *k*, *s*). For instance, high frequency recombination in the $E\beta$ gene is observed when different laboratory strains are crossed (Steinmetz *et al.*, 1982) and recombination between *K* and *I* region marker loci also occurs at an unusually high frequency when *cas3* and *cas4* MHC haplotypes are crossed (K.Fischer-Lindahl, unpublished results).

Molecular basis of meiotic recombinational hotspots

High frequency recombination at hotspots might be due to certain nucleotide sequences being preferentially recognized by the recombination machinery or might result from an open and therefore easily accessible chromatin configuration during meiosis. At present we cannot rule out either possibility, although we favor the specific-sequence model.

In the region sequenced we have identified three elements which we would like to discuss with respect to the molecular basis of the hotspot. First, a 57-bp long sequence present in CAS3 DNA at position 1438 is found to be duplicated in the C57BL/6 genome (Figures 4 and 5). We think it is very unlikely that this structural non-homology between CAS3 and C57BL/6 DNA explains the presence of the recombinational hotspot. Based on our present understanding of the mechanism of recombination a mismatch should decrease rather than increase recombination frequencies.

Second, the stretch of 36 bp to which we could confine the breakpoint in CAS3(R18) (Figures 4 and 5) has 72% homology (26 out of 36 bp) to a conserved portion of certain solitary LTR sequences, of which there are about 500 copies dispersed throughout the mouse genome. These solitary LTR sequences are ~500 bp in length, resemble insertion elements and show homology to LTRs of xenotropic but not ecotropic murine leukemia viruses (Wirth *et al.*, 1983). The 36-bp long homologous sequence is present in both CAS3 and C57BL/6 DNA. What, if any, significance it has with respect to the recombinational hotspot is unclear.

Third, a TCTG tetramer repeated six times in CAS3 but only four times in C57BL/6 DNA, is found at position 1921 (Figures 4 and 5). The complementary strand, $(CAGA)_{4-6}$, of this repeat sequence is reminiscent of the CAGG-repeat sequence in the $E\beta$ recombinational hotspot. If the CAGA repeat is part of the structural basis of the $A_{\beta3}/A_{\beta2}$ hotspot, the $(CAGA)_6$ repeat in CAS3 must be a much better recombination site than the $(CAGA)_4$ repeat in C57BL/6 DNA. It is possible, however, that the presence of a second sequence element is required for the repeat sequence to be recognized by the putative recombinase. Such a second element might be located some distance away from the repeat sequence and could be missing in C57BL/6 DNA which would explain the absence of the $A_{\beta3}/A_{\beta2}$ recombinational hotspot in C57BL/6. In the region sequenced we have not been able to identify a candidate sequence element for such a second site.

We note that several interesting similarities appear to exist between the action of the $A_{\beta3}/A_{\beta2}$ meiotic recombinational hotspot and Chi hotspots for recombination in phage λ (for a review on Chi hotspots see Smith, 1983). Like Chi, the $A_{\beta3}/A_{\beta2}$ hotspot acts locally, appears to stimulate recombination maximally near itself and preferentially on one side, acts in a dominant fashion and enhances only homologous exchange. Chi is known to be active only in the presence of a second site (cos in phage λ) located in *cis* and in the proper orientation.

It is obvious that DNA sequence comparisons between recombinational hotspots can only suggest possibilities. This emphasizes the importance of developing functional assays for recombinational hotspots. There now exist various *in vivo* and *in vitro* recombination systems that could be used to test DNA fragments for the presence of sequences which would enhance homologous recombination (Liskay and Stachelek, 1983; Smith and Berg, 1984; Jasin *et al.*, 1985; Kucherlapati *et al.*, 1985; Lin *et al.*, 1985; Shaul *et al.*, 1985; Smithies *et al.*, 1985; Treco *et al.*, 1985; Thomas *et al.*, 1986). To find out whether meiotic recombinational hotspots are a result of chromosome folding, DNase I could be used to search for a correlation between DNase I-hypersensitive sites and hotspots. Identification of the structural basis of recombinational hotspots is not only interesting in itself, but they might also be useful elements for targeting transfected genes to specific sites in mammalian DNA.

Materials and methods

Construction and screening of λEMBL3 genomic libraries

Mouse liver DNAs were partially digested with *Mbo*I and size-selected for 19–23 kb long DNA fragments by sucrose gradient fractionation. λEMBL3 DNA (Frischauf *et al.*, 1983) was digested with *Bam*HI and a small aliquot was end-

labelled with [α-^{32}P]dGTP using the large fragment of *Escherichia coli* DNA polymerase I. The labelled EMBL3 DNA fragments were mixed with the unlabelled ones and digested with *Eco*RI. Completion of digestion was checked by agarose-gel electrophoresis and autoradiography. The small linker fragments were removed by fractionation on a Sepharose 4B column. Vector arms were then ligated to size-selected mouse DNA fragments and the mixture was packaged *in vitro*.

About 1 × 10^6 plaques were plated on five square dishes (23 × 23 cm, NUNC) for each library. Screening of the libraries was carried out according to Maniatis *et al.* (1984) except that membranes used were Gene Screen Plus (New England Nuclear).

The filters were hybridized with oligolabelled probes (Feinberg and Vogelstein, 1984) using conditions as described by Steinmetz *et al.* (1985). After hybridization, filters were washed in 3 × SSC, 0.1% at 65°C twice for 20 min each, followed by two washes in 0.1 × SSC, 0.1% SDS at 65°C for 20 min each.

Restriction mapping

Recombinant λ DNAs were isolated according to Maniatis *et al.* (1982) and digested with restriction endonucleases in the buffers specified by the commercial suppliers of the enzymes. Digested DNA fragments were separated by electrophoresis on 0.3% agarose gels. Restriction enzyme sites were mapped by double-digestion (Maniatis *et al.*, 1982) and by partial digestion following the protocol described by Rackwitz *et al.* (1984).

Primer-directed DNA sequencing

For DNA sequencing a 2.8-kb *Sal*I – *Sac*I restriction fragment was isolated from λ clones B6.11, CAS3.31 and R0.11 and a 2.1-kb *Sal*I – *Sac*I fragment from R18.21 by agarose gel electrophoresis and electro-elution using an electrophoretic concentrator (ISCO). These fragments were then subcloned into M13mp18 and M13mp19 vector DNAs. M13 subclones were grown in JM109 (Yanish-Perron *et al.*, 1985) and single-stranded template DNAs were isolated and DNA sequencing reactions were carried out using a commercial kit (Amersham). For sequencing we used a 17-mer universal M13 primer (Amersham) and 11 17-mer primers synthesized to elongate previously determined sequencing runs (Strauss *et al.*, 1986). Primers started at positions 206, 451, 702, 931, 1148 and 2601 for the strand shown in Figure 5 and at positions 222, 1766, 1960, 2219 and 2617 for the complementary strand. This allowed us to determine the sequence of the strand shown in Figure 4 from positions 1 – 1565 and 2577 – 2754. The complementary strand was sequenced from positions 1 – 108 and 1417 – 2754.

Southern blot analysis

After restriction enzyme digestion, 10 μg of DNA fragments were separated by electrophoresis on agarose gels and transferred to Zeta probe membranes (BioRad) using 0.4 N NaOH as transfer buffer (Reed and Mann, 1985) or 10 × SSC (Maniatis *et al.*, 1982). After transfer, filters were rinsed twice in 2 × SSC, dried at 80°C under vacuum and hybridized as described in Steinmetz *et al.* (1986) with oligolabelled (Feinberg and Vogelstein, 1984) probes shown in Figure 3.

Acknowledgements

We thank G.R.Dastoornikoo for expert technical assistance, J.Kaufman and C.Steinberg for constructive criticism on the manuscript, and J.Hossmann for preparation of the typescript. Yasushi Uematsu was on leave of absence from the Department of Immunology, School of Medicine, Chiba University, Chiba, Japan. The Basel Institute for Immunology was founded and is supported by Hoffmann-La Roche & Co. Ltd, Basel, Switzerland.

References

Begovich,A.B. and Jones,P.P. (1985) *Immunogenetics*, **22**, 523 – 532.
Feinberg,A.P. and Vogelstein,B. (1984) *Anal. Biochem.*, **137**, 266 – 267.
Flavell,R.A., Allen,H., Huber,B., Wake,C. and Widera,G. (1985) *Immunol. Rev.*, **64**, 29 – 50.
Frischauf,A.-M., Lehrach,H., Poustka,A. and Murray,N. (1983) *J. Mol. Biol.*, **170**, 827 – 842.
Jasin,M., de Villiers,J., Weber,F. and Schaffner,W. (1985) *Cell*, **43**, 695 – 703.
Jeffreys,A.J., Wilson,V. and Thein,S.L. (1985) *Nature*, **314**, 67 – 73.
Kmiec,E.B., Angelides,K.J. and Holloman,W.K. (1985) *Cell*, **40**, 139 – 145.
Kobori,J.A., Winoto,A., McNicholas,J. and Hood,L. (1984) *J. Mol. Cell. Immunol.*, **1**, 125 – 131.
Kobori,J.A., Strauss,E., Minard,K. and Hood,L. (1986) *Science*, in press.
Krayev,A.S., Kramerov,D.A., Skryabin,K.G., Ryskov,A.P., Bayev,A.A. and Georgiev,G.P. (1980) *Nucleic Acids Res.*, **8**, 1201 – 1215.
Krayev,A.S., Markusheva,T.V., Kramerov,D.A., Ryskov,A.P., Skryabin,K.G., Bayev,A.A. and Georgiev,G.P. (1982) *Nucleic Acids Res.*, **10**, 7461 – 7475.
Kucherlapati,R.S., Spencer,J. and Moore,P.D. (1985) *Mol. Cell. Biol.*, **5**, 714 – 720.
Lafuse,W., Berg,N., Savarirayan,S. and David,C.S. (1986) *J. Exp. Med.*, **163**, 1518 – 1528.
Lebo,R.V., Chakravarti,A., Buetow,K.H., Cheung,M.C., Cann,H., Cordell,B. and Goodman,H. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4808 – 4812.
Lin,F.-L., Sperle,K. and Sternberg,N. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1391 – 1395.
Liskay,R.M. and Stachelek,J.L. (1983) *Cell*, **35**, 157 – 165.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.
Orkin,S.H. and Kazazian,H.H.,Jr. (1984) *Annu. Rev. Genet.*, **18**, 131 – 171.
Rackwitz,H.-R., Zehetner,G., Frischauf,A.-M. and Lehrach,H. (1984) *Gene*, **30**, 195 – 200.
Reed,K.C. and Mann,D.A. (1985) *Nucleic Acids Res.*, **13**, 7201 – 7221.
Saha,B.K. and Cullen,S.E. (1986) *J. Immunol.*, **136**, 1112 – 1116.
Saito,H., Maki,R.A., Clayton,L.K. and Tonegawa,S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5520 – 5524.
Shaul,Y., Laub,O., Walker,M.D. and Rutter,W.J. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3781 – 3784.
Shiroishi,T., Sagai,T. and Moriwaki,K. (1982) *Nature*, **300**, 370 – 372.
Smith,G.R. (1983) *Cell*, **34**, 709 – 710.
Smith,A.J.H. and Berg,P. (1984) *Cold Spring Harbor Symp. Quant. Biol.*, **49**, 171 – 181.
Smith,G.R., Kunes,S.M., Schultz,D.W., Taylor,A. and Trinan,K.L. (1981) *Cell*, **24**, 429 – 436.
Smithies,O., Gregg,R.G., Boggs,S.S., Koralewski,M.A. and Kucherlapati,R.S. (1985) *Nature*, **317**, 230 – 234.
Steinmetz,M. (1986) In Rigby,P.W.J. (ed.), *Genetic Engineering 5*. Academic Press, London, pp. 117 – 158.
Steinmetz,M., Minard,K., Horvath,S., McNicholas,J., Frelinger,J., Wake,C., Long,E., Mach,B. and Hood,L. (1982) *Nature*, **300**, 35 – 42.
Steinmetz,M., Stephan,D., Dastoornikoo,G.R., Gibb,E. and Romanuik,R. (1985) In Lefkovits,I. and Pernis,B. (eds), *Methods in Molecular Immunology: Immunological Methods Vol. III*. Academic Press, New York, pp. 1 – 19.
Steinmetz,M., Stephan,D. and Fischer-Lindahl,D. (1986) *Cell*, **44**, 895 – 904.
Stephan,D., Sun,H., Fischer Lindahl,K., Meyer,E., Hämmerling,G., Hood,L. and Steinmetz,M. (1986) *J. Exp. Med.*, **163**, 1222 – 1244.
Straus,E.C., Kobori,J.A., Siu,G. and Hood,L.E. (1986) *Anal. Biochem.*, **154**, 353 – 360.
Thomas,K.R., Folger,K.R. and Capecchi,M.R. (1986) *Cell*, **44**, 419 – 428.
Treco,D., Thomas,B. and Arnheim,N. (1985) *Mol. Cell. Biol.*, **5**, 2029 – 2038.
Widera,G. and Flavell,R.A. (1984) *EMBO J.*, **3**, 1221 – 1225.
Wirth,T., Glögger,K., Baumruker,T., Schmidt,M. and Horak,I. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3327 – 3330.
Yanish-Perron,C., Vieira,J. and Messing,J. (1985) *Gene*, **33**, 103 – 119.

Received on 25 June 1986