The EMBO Journal vol.6 no.2 pp.369–373, 1987

Orientation and molecular map position of the complement genes in the mouse MHC

Urs Müller[1], Danielle Stephan[1], Peter Philippsen[2] and Michael Steinmetz[1,3]

[1]Basel Institute for Immunology, Grenzacherstrasse 487, CH-4058 Basel, Switzerland and [2]Institut für Mikrobiologie und Molekularbiologie, Justus-Liebig-Universität Giessen, Frankfurterstrasse 107, D-6300 Giessen, FRG

[3]Present address: Central Research Units, F.Hoffmann-La Roche & Co., Ltd, CH-4002 Basel, Switzerland

Communicated by M.Steinmetz

Over the past few years six gene clusters have been isolated from the major histocompatibility complex (MHC) of the BALB/c mouse encompassing a total of 1600 kb of DNA and 48 genes. The molecular distances between these gene clusters and the orientation of four of the six clusters on chromosome 17 is not known. Here we use pulse-field gradient gels and Southern blot hybridization to establish large-scale genomic restriction maps covering several hundreds of kb surrounding the three gene clusters located in the *K*, *I*, *S*, and *D* regions of the MHC. Comparison of the maps orients the complement gene clusters in the *S* region with the *21-OHB* gene pointing towards the *K* end and the *C2* gene pointing towards the *D* end of the MHC. The distances between the E_α and *21-OHB* genes is 430 kb and between the *C2* and *TNF-α* genes at least 420 kb.

Key words: gene map/major histocompatibility complex/pulse-field gradient gels

Introduction

Genes coding for class I and class II cell surface glycoproteins that play a role in the presentation of foreign antigens to T lymphocytes are located in the major histocompatibility complex (MHC) together with a number of other homologous and non-homologous genes. So far 48 genes have been cloned from the MHC of the BALB/c mouse and ordered into six gene clusters defined by overlapping cosmid clones (see Steinmetz, 1986; Steinmetz *et al.*, 1986; Stephan *et al.*, 1986; Müller *et al.*, 1987). A cluster of 600 kb containing two class I and seven class II genes defines the proximal end of the MHC and is separated by an unknown distance from a 500-kb cluster with 13 class I, the tumor necrosis factor (*TNF-α*) and lymphotoxin (*TNF-β*) genes (Steinmetz *et al.*, 1986; Stephan *et al.*, 1986; Müller *et al.*, 1987) (Figure 1a). These two gene clusters span the *K* and *I*, and the *D* and *Qa* regions, respectively. From the *S* region between these two clusters a 300-kb gene cluster has been cloned containing six genes coding for certain complement components and 21-hydroxylase, an enzyme involved in steroid biosynthesis (Chaplin *et al.*, 1983; Steinmetz *et al.*, 1984; White *et al.*, 1984 and this paper). The molecular map position and the orientation of this gene cluster in the MHC has been unknown so far because informative recombinant mouse strains do not exist. We used pulse-field gradient (PFG) gel electrophoresis and Southern blot hybridization to establish long-range genomic restriction maps for the four facing ends of the *I*, *S* and *D* region gene clusters. Comparison of the maps allowed us to orient the complement gene cluster and to place it 170 kb distal to the *I* region gene cluster. Its distance to the *D* region gene cluster, on the other hand, is at least 350 kb indicating that more genes might be located in this portion of the MHC still to be cloned.

Results and Discussion

Restriction enzymes useful for PFG gel analysis of the mouse MHC

PFG gels will separate DNA fragments up to at least 2000 kb in length (Carle and Olson, 1984; Schwartz and Cantor, 1984) and in conjunction with infrequently cutting restriction enzymes and Southern blot hybridization have been used to study gene organization in the MHC of man (Hardy *et al.*, 1986; Ragoussis *et al.*, 1986). Among the restriction enzymes useful for the construction of large-scale genomic restriction maps are those specific for 6- or 8-bp long recognition sites containing one or two CpG dinucleotides, since CpG is known to be under-represented in mammalian DNA. We have mapped the six MHC gene clusters, isolated from the MHC of the BALB/c mouse and encompassing a total of 1600 kb of DNA (Steinmetz *et al.*, 1986; Stephan *et al.*, 1986) for some of these restriction enzymes. In Table I the number of sites found for six restriction enzymes is compared to their recognition sequence. It is interesting that the combination of two CpG dinucleotides with A and T nucleotides in a hexanucleotide recognition sequence is particularly rare in the MHC of the BALB/c mouse. For instance, *Pvu*I recognizing the hexanucleotide CGATCG has less than half as many sites as *Not*I, recognizing the octanucleotide GCGGCCGC.

Genomic restriction maps

We decided to use four of the infrequently cutting restriction enzymes, namely *Mlu*I, *Not*I, *Nru*I and *Pvu*I, for the construction of genomic restriction maps. Figure 1a shows three gene clusters of the MHC of the BALB/c mouse which we intended to link at the molecular level. Seven single-copy hybridization probes were isolated from the ends of the three gene clusters. Accord-

Table I. Frequency of sites in the MHC for some enzymes containing two CpGs in their recognition sequence[a]

Enzyme	Number of sites	Recognition sequence[b]
*Pvu*I	8	CGATCG
*Nru*I	15	TCGCGA
*Mlu*I	19	ACGCGT
*Not*I	19	GCGGCCGC
*Sac*II	>60	CCGCGG
*Bss*HII	>60	GCGCGC

[a]The six cloned BALB/c gene clusters described by Steinmetz *et al.* (1986) and Stephan *et al.* (1986) were mapped for the enzymes indicated.
[b]CpG dinucleotides are underlined.

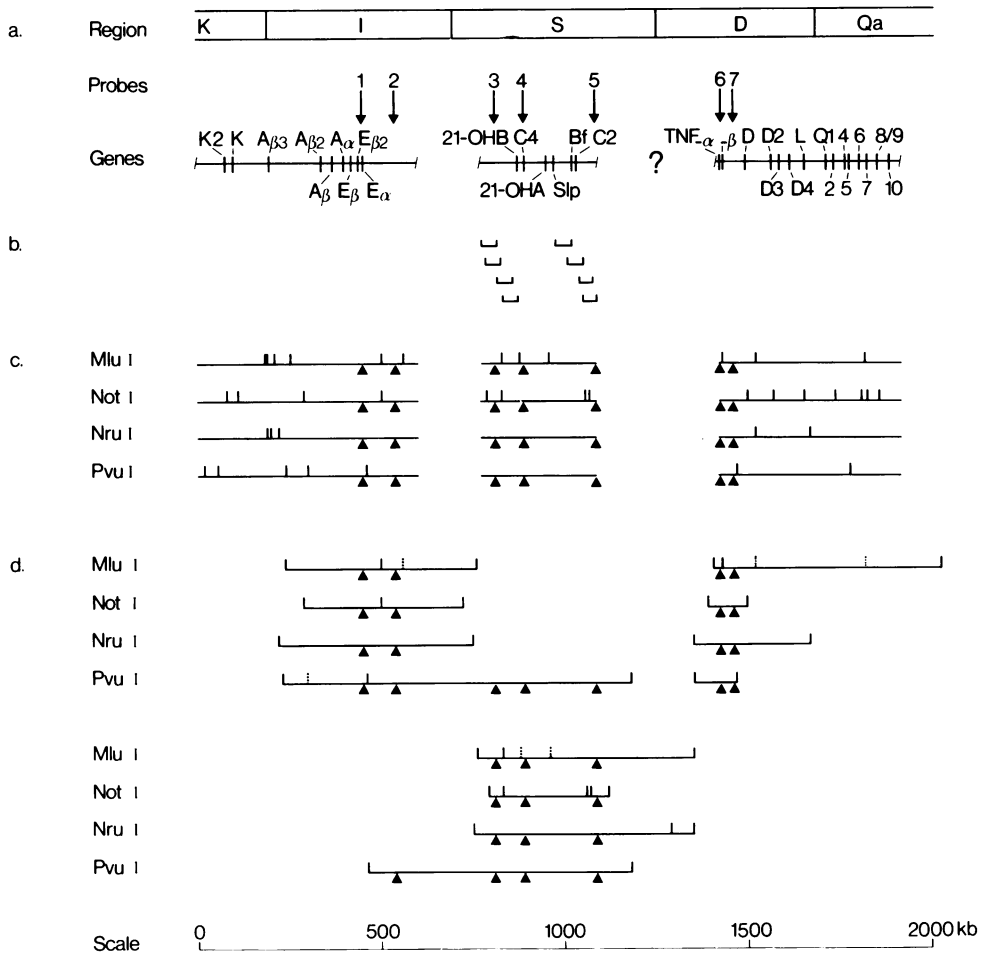

Fig. 1. Organization of MHC genes in the BALB/c mouse. **(a)** Three gene clusters, described previously (Steinmetz *et al.*, 1984,1986; Stephan *et al.*; Müller *et al.*, 1987) are spaced and oriented on the molecular map based on results described in this paper. Genetic regions of the mouse MHC, defined by recombination events, are shown on top. Probes used for Southern blot analysis are: 1, probe 7 (Steinmetz *et al.*, 1982); 2, probe 10 (Steinmetz *et al.*, 1982); 3, 0.4-kb *Mbo*I fragment isolated from II 9.19 (see below); 4, the *C4* cDNA clone pMC4/21 (Tosi *et al.*., 1984); 5 1.5-kb *Mbo*I fragment isolated from cosmid clone III 1.1 (see below); 6, 0.7-kb *Rsa*I fragment isolated from cosmid clone II 3.5 (Stephan *et al.*, 1986); 7, probe 3 (Stephan *et al.*, 1986). Probe locations are indicated by arrows throughout the figure. *K2*, *K*, *D*, *D2*, *D3*, *D4*, *L*, *Q1*, *Q2*, *Q4*, *Q5*, *Q6*, *Q7*, *Q8/9* and *Q10* are class I genes; $A_{\beta 3}$, $A_{\beta 2}$, A_{β}, A_{α}, E_{β}, $E_{\beta 2}$ and E_{α} are class II genes; *21-OHA* codes for 21-hydroxylase while *21-OHB* is non-functional; *C4*, *Slp*, *Bf*, *C2* code for complement and related components; *TNF-α* and *-β* code for tumor necrosis factor and lymphotoxin, respectively. The question mark indicates that the distance between the *S* and the *D/Qa* region gene clusters, shown to be at least 350 kb, could be larger. **(b)** Overlapping cosmid clones defining the ends of the *S* region gene cluster are shown. From left to right they are: III 5.1, III 12.1, II 9.19, II 1.1, II 2.1, II 2.19, II 6.19, III 1.1. Cosmid clones for the middle portion of the *S* gene cluster and for the *K*, *I*, *D* and *Qa* region gene clusters have been described previously (Steinmetz *et al.*, 1984,1986; Stephan *et al.*, 1986). **(c)** Restriction maps established by single and double digests of cloned DNAs. For other enzymes see Steinmetz *et al.* (1984,1986) and Stephan *et al.* (1986). **(d)** Restriction maps obtained by Southern blot hybridization of BALB/c DNA separated on PFG gels after single and double digestions. Probes 1 and 2 were used to map the distal end of the *I* region (**upper panel**, left), probes 3, 4 and 5 to map the *S* region **(lower panel)** and probes 6 and 7 to map the proximal end of the *D* region (**upper panel**, right). Sites shown by stippling, present in cloned DNA, were not found when genomic DNA was digested.

ing to their location, these probes should allow us to confirm and extend the restriction maps established using cloned DNA (Figure 1c). High mol. wt DNA, prepared from BALB/c spleen cells in agarose blocks, was digested with each of the four enzymes alone and in all pairwise combinations, separated on PFG gels, transferred to nylon membranes and hybridized sequentially with the seven probes. Figure 2 shows results obtained for a Southern blot with *Mlu*I, *Nru*I and *Pvu*I single and double digests sequentially hybridized with the probes. Sizes of restriction fragments were determined by comparison with co-migrating oligomers of phage λ DNA and are summarized in Table II.

Hybridization of *Mlu*I-digested DNA with probe 3 identified a 70-kb fragment together with a 580-kb partial fragment while probes 4 and 5 picked up a 530-kb fragment and the same 580-kb partial digestion product. Comparison of this result with the restriction map of the complement gene cluster indicates that two of the three *Mlu*I sites found in cloned DNA are not recognized when genomic DNA is digested. The most likely explanation for this finding is methylation of the corresponding sites in genomic DNA although sequence polymorphism in the BALB/c mice used cannot be excluded. Hybridization of *Not*I-digested DNA with probes 3 and 4 identifies the restriction fragments expected from the analysis of cloned DNA. Probe 5 identifies a 50-kb *Not*I fragment extending outside of the cloned region. No restriction sites for *Nru*I and *Pvu*I have been found in the cloned region. Hybridization of genomic DNA gives a 520-kb *Nru*I fragment (and a 600-kb partial) and a 700-kb *Pvu*I fragment with probes 3, 4 and 5. *Mlu*I + *Nru*I, *Mlu*I + *Pvu*I and *Nru*I + *Pvu*I double digests place the *Nru*I and *Pvu*I sites outside of the cloned region as shown in Figure 1d.

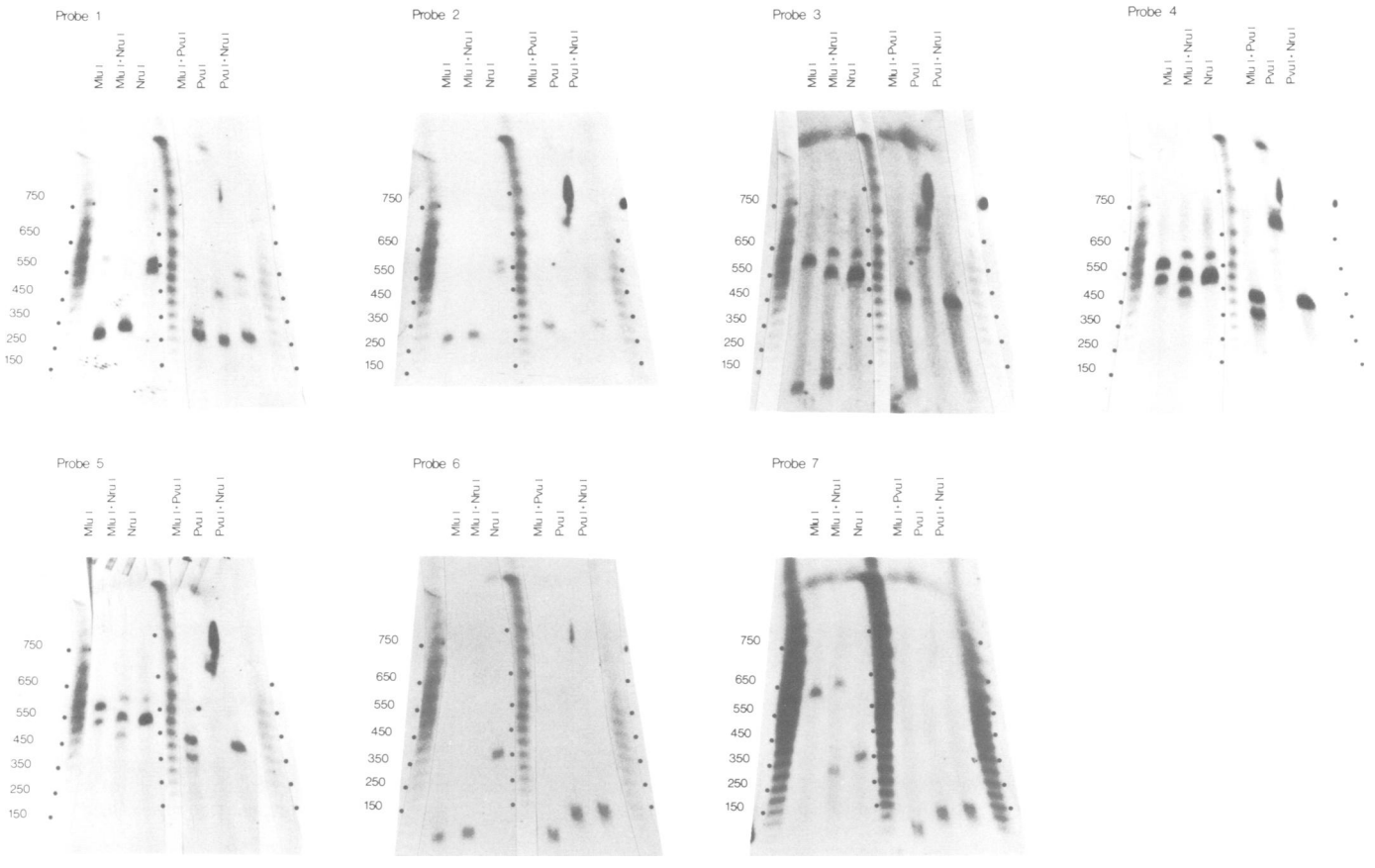

Fig. 2. PFG gel analysis of BALB/c DNA with *I*, *S* and *D* region probes. Hybridization probes are those shown in Figure 1. Enzymes used for digestion are indicated. Phage λ oligomers, identified by hybridization with ^{32}P-labeled λ DNA, were run in parallel and used as size markers. Sizes are given in kilobase pairs. All autoradiograms were obtained from a single filter hybridized sequentially with probes 2 (together with labeled λ DNA), 5, 3, 6, 4, 1, and 7 (together with labeled λ DNA).

Table II. Sizes (kb) of restriction fragments determined by Southern blot hybridization of PFG gels[a]

Restriction enzyme	Probe						
	1	2	3	4	5	6	7
*Mlu*I	260 (+530,600)	270	70 (+580)	530 (+580)	530 (+580)	<50	610
*Not*I	220	220 (+500,600)	<50 (+270)		50	110	
*Nru*I	540 (+570)	540 (+570)	520 (+600)	520 (+600)	520 (+600)	350	350
*Pvu*I	250 (+450)	700	700	700	700	120	120
*Mlu*I + *Not*I	220	220			50	<50	
*Mlu*I + *Nru*I	260	270	70 (+550,620)	480 (+550,620)	480 (+550,620)	<50	290 (+610)
*Mlu*I + *Pvu*I	250	270	70 (+440)	360 (+440)	360 (+440)	<50	50
*Not*I + *Nru*I	220	220			50	110	
*Not*I + *Pvu*I	190	220	<50 (+270)		50	100	
*Nru*I + *Pvu*I	250 (+450,520)	300 (+550)	410	410	410	120	120

[a]BALB/c spleen DNA was digested and analysed as described in Materials and methods. Partial digestion products are given in parentheses.

Genomic restriction maps for the distal end of the *I* and the proximal end of the *D* region gene clusters were constructed in similar ways and are shown in Figure 1d. It is striking that probe 2 from the *I* region and probes 3, 4 and 5 from the *S* region pick up the same sized *Pvu*I fragment of ~700 kb. In fact, the two genomic maps constructed for the *I* and *S* regions can be overlapped as shown in Figure 1d if indeed the 700-kb *Pvu*I fragment identified with *I* and *S* region probes is identical.

Independent evidence that the 700-kb *Pvu*I fragment seen with *I* region probes is the same as the one seen with *S* region probes was obtained through the analysis of mouse strains B10.WR7 and B10.BUA1. These strains carry at least three *Slp* genes (Levi-Strauss *et al.*, 1985; Rosa *et al.*, 1985; Robins *et al.*, 1986). When analysed on PFG gels probe 4 from the *S* region identified 950- and 850-kb *Pvu*I fragments in B10.WR7 and B10.BUA1 DNA, respectively, which were larger than the 700-kb *Pvu*I fragment identified in BALB/c DNA (Figure 3). This finding is compatible with the gene amplification in the *S* region. Interestingly, probe 2 from the *I* region hybridized to the same *Pvu*I fragments in B10.WR7 and B10.BUA1 DNA as

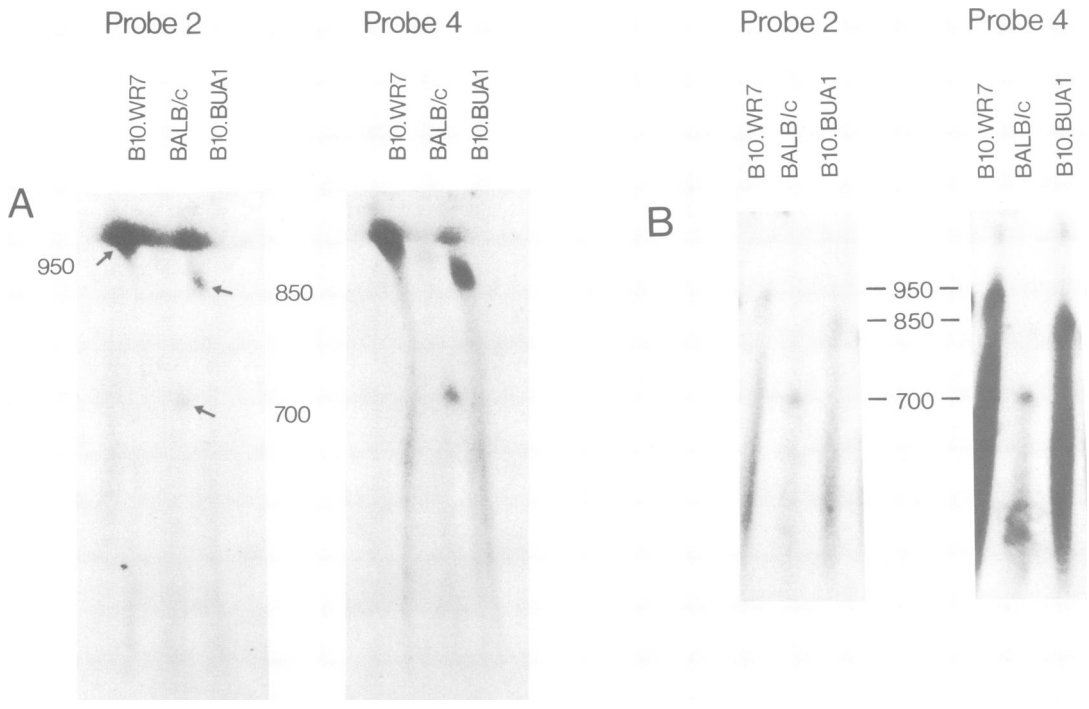

Fig. 3. PFG gel analysis of BALB/c, B10.WR7 and B10.BUA1 DNA with *I* and *S* region probes. Mouse spleen cell DNAs were prepared, digested with *Pvu*I and analysed by sequential hybridization with probes 2 and 4 (see Figure 1) as described in Materials and methods. The PFG gel shown in **(a)** was run with a switching interval of 60 s, while for the one shown in **(b)** an interval of 100 s was used. Sizes are given in kilobase pairs and were determined by comparison to λ oligomers run in parallel.

seen with the *S* region probe (Figure 3). Since the same size shifts are identified with *I* and *S* region probes we conclude that these probes indeed hybridize to the same *Pvu*I fragment.

Orientation and molecular map position of the complement genes

The identification of the overlapping *Pvu*I fragment and the location of the flanking *Mlu*I, *Not*I and *Nru*I sites determines the orientation of the complement gene cluster as shown in Figure 1a. Previously, chromosome transfer experiments have suggested an inverted orientation of the complement gene cluster within the mouse MHC (Weis *et al.*, 1986). This prediction, however, was based on the analysis of one transferent and the occurrence of gene rearrangements during transfer and establishment of the transformed cells could not be excluded. For the human MHC conflicting proposals for the orientation of this cluster have been made, based on the analysis of present day haplotypes and attempts to connect these haplotypes on an evolutionary scale with as few recombination events as possible (Marshall *et al.*, 1984; Wilton and Charlton, 1986). Since the gene order within the class II and the complement region is the same in mouse and man (White *et al.*, 1984; Hardy *et al.*, 1986), we predict that in the human MHC the *DR*, *21-OHB*, *C4B*, *21-OHA*, *C4A*, *Bf*, *C2*, *TNF-α*, *TNF-β*, and *HLA-B* genes will be ordered in the same way as their murine counterparts.

Physical map of the mouse MHC

The size of the gap between the class II and the complement genes not yet cloned from the BALB/c MHC is ~170 kb. On the other hand, the distance between the complement genes and the *TNF-α* and *-β* genes is at least 350 kb but could be larger (Figure 1). It is interesting to note that the genes coding for 21-hydroxylase, the complement components and TNF-α and *-β* are organized in pairs of related genes and have the same 5′ to 3′ orientation, opposite to the orientation of all the class I genes in the *D* and *Qa* regions (Chaplin *et al.*, 1984; White *et al.*, 1984; Stephan *et al.*, 1986; Müller *et al.*, 1987; this paper). Based on the density of genes already identified in the cloned portions of the mouse MHC it is very likely that more genes, e.g. neuraminidase (Figueroa *et al.*, 1982; Womack and David; 1982), will be found between *C2* and *TNF-α*.

This paper demonstrates the usefulness of PFG gel analysis in establishing gene maps extending over millions of basepairs. Summing up the sizes of the cloned regions and distances determined so far by PFG gel analysis gives a minimum estimate of 2300 kb for the MHC of the BALB/c mouse. The determination of gap sizes betwen cloned regions is important if chromosomal walking experiments are planned to link these regions. Finally, the results presented make a comparison between molecular and genetic maps possible for the central portion of the mouse MHC. A distance of 450 kb between E_α and *C4* and at least 600 kb between *C4* and *D* correlates with a recombination frequency (Snell *et al.*, 1976) of 0.11 and 0.26%, respectively. Since the average recombination frequency in the mouse genome is 1%/2000 kb, there is no need at the moment to postulate the presence of recombinational hotspots (Steinmetz *et al.*, 1982,1986; Uematsu *et al.*, 1986) in this portion of the mouse MHC.

Materials and methods

Cosmid clones and hybridization probes

Cosmid clones II 1.1, II 2.1, II 2.19, II 6.19 and II 9.19 were isolated from a BALB/c cosmid library (Steinmetz *et al.*, 1984) constructed using vector pNNL, clones III 1.1, III 5.1 and III 12.1 from a BALB/c cosmid library constructed using vector Lorist B. Lorist B (a cosmid vector based on the replication origin of phage λ and conferring kanamycin resistance) was obtained from P.Little, Chester Beatty Laboratories, London, UK. Cosmid clones were isolated by conventional chromosomal walking procedures (Steinmetz *et al.*, 1985). Hybridization probes (see legend to Figure 1), except probe 4 which is a cDNA clone,

were isolated from the cosmids by restriction enzyme digestion and preparative agarose gel electrophoresis and subcloned into pUC8.

PFG gel analysis of BALB/c DNA

Preparation of high mol. wt DNA in agarose blocks from BALB/c spleen cells, restriction enzyme digestion, PFG gel electrophoresis, transfer to Zeta probe membranes (BioRad), and hybridization with oligolabeled probes was as described (Carle and Olson, 1984; Schwartz and Cantor, 1984; Steinmetz *et al.*, 1986; Van Ommen and Verkerk, 1986) and according to protocols obtained from D.Barlow and H.Lehrach (EMBL, Heidelberg, FRG). In brief, BALB/c spleens were disrupted to obtain single cells, 7×10^5 cells/block ($6 \times 2 \times 5$ mm) were suspended in agarose, and DNA was prepared by incubation in 0.5 M EDTA (pH 8.0), 1% sodium lauroyl sarcosine, 2% proteinase K at 50°C for 48 h with constant rotation. After several washes at room temperature in 10 mM Tris–HCl, 1 mM EDTA (pH 8.0) (TE) and two washes at 50°C in TE plus 0.04 mg/ml PMSF, digestions were performed with 40–80 U of enzyme for 5 h at 37°C. Blocks were then washed in TE and transferred to pre-formed slots of a 20 × 20 cm PFG gel. The PFG electrophoresis unit was purchased from EMBL, Heidelberg, FRG. Gels were run in 0.25 × TBE buffer (Van Ommen and Verkerk, 1986) for 40 h at 14°C and 450 V with a switching interval of 60 s. After staining with ethidium bromide gels were photographed and prepared for transfer by sequential incubation in 0.25 M HCl for 2 × 10 min and for 60 min in 0.5 M NaOH, 1.5 M NaCl. Transfer was with 0.4 M NaOH overnight. Filters were then dried by baking (2 h, 80°C, vacuum oven) and hybridized as described (Steinmetz *et al.*, 1986). For hybridization with a new probe, filters were stripped by two incubations in 0.1 × SSC, 0.1% SDS at 95°C. Autoradiography was with two intensifying screens at −70°C for 1–3 days.

Acknowledgements

We thank Denise Barlow and Hans Lehrach for detailed protocols concerning the preparation and digestion of high mol. wt DNA in agarose blocks and the handling of PFG gels, Peter Little for the cosmid vector Loris B, Felipe Figueroa and Jan Klein for B10.WR7 and B10.BUA1 mice, Susan Carson for comments and Catherine Plattner and Anne Iff for typing the manuscript. The Basel Institute for Immunology was founded and is supported by F.Hoffmann-La Roche & Co., Ltd, Basel, Switzerland.

References

Carle,G.F. and Olson,M.V. (1984) *Nucleic Acids Res.*, **12**, 5647–5664.

Chaplin,D.D., Woods,D.E., Whitehead,A.S., Goldberger,G., Colton,H.R. and Seidman,J.G. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6947–6951.

Figueroa,F., Klein,D., Tewarson,S. and Klein,J. (1982) *J. Immunol.*, **129**, 2089–2093.

Hardy,D.A., Bell,J.I., Long,E.O., Lindstein,T. and McDevitt,H.O. (1986) *Nature*, **323**, 453–455.

Lévi-Strauss,M., Tosi,M., Steinmetz,M., Klein,J. and Meo,T. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 1746–1750.

Marshall,W.H., Neugebauer,M., Baur,M.P. and Albert,E.D. (1984) In Albert,E.D., Baur,M.P. and Mayr,W.R. (eds), *Histocompatibility Testing 1984*. Springer Verlag, Berlin, pp. 313–316.

Müller,U., Jongeneel,C.V., Nedospasov,S.A., Fischer Lindahl,K. and Steinmetz,M. (1987) *Nature*, **325**, 265–267.

Ragoussis,J., van der Bliek,A., Trowsdale,J. and Ziegler,A. (1986) *FEBS Lett.*, **204**, 1–4.

Robins,D.M., Malissen,M., Hood,L., Ferreira,A., Walthall,D. and Mitchell,M. (1986) *Mol. Cell. Biol.*, **6**, 134–141.

Rosa,P.A., Sepich,D.S., Shreffler,D.C. and Ogata,R.T. (1985) *J. Immunol.*, **135**, 627–631.

Schwartz,D.C. and Cantor,C.R. (1984) *Cell*, **37**, 67–75.

Snell,G.D., Dausset,J. and Nathenson,S. (1976) *Histocompatibility*. Academic Press, New York.

Steinmetz,M. (1986) In Rigby,P.W.J. (ed.), *Genetic Engineering*. Academic Press, London, Vol. 5, 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., Malissen,M., Hood,L., Oern,A., Maki,R.A., Dastoornikoo,G.R., Stephan,D., Gibb,E. and Romaniuk,R. (1984) *EMBO J.*, **3**, 2995–3003.

Steinmetz,M., Stephan,D., Dastoornikoo,G.R., Gibb,E. and Romaniuk,R. (1985) In Lefkovits,I. and Pernis,B. (eds), *Immunological Methods*. Academic Press, New York, Vol. III, pp. 1–19.

Steinmetz,M., Stephan,D. and Fischer Lindahl,K. (1986) *Cell*, **44**, 895–904.

Stephan,D., Sun,H., Fischer Lindahl,K., Meyers,E.,Hämmerling,G., Hood,L. and Steinmetz,M. (1986) *J. Exp. Med.*, **163**, 1227–1244.

Tosi,M., Lévi-Strauss,M., Duponchel,C. and Meo,T. (1984) *Phil. Trans. R. Soc. London Ser. B*, **306**, 389–394.

Uematsu,Y., Kiefer,H., Schulze,R., Fischer Lindahl,K. and Steinmetz,M. (1986) *EMBO J.*, **5**, 2123–2129.

Van Ommen,G.J.B. and Verkerk,J.M.H. (1986) In Davies,K.E. (ed.), *Human Genetic Diseases*. IRL Press, Oxford, pp. 113–133.

Weis,J.H., Seidman,J.G., Housman,D.-E. and Nelson,D.L. (1986) *Mol. Cell. Biol.*, 441–451.

White,P.C., Chaplin,D.D., Weis,J.H., Dupont,B., New,M.I. and Seidman,J.G. (1984) *Nature*, **312**, 465–467.

Wilton,A.N. and Charlton,B. (1986) *Immunogenetics*, **24**, 79–83.

Womack,J.E. and David,C.S. (1982) *Immunogenetics*, **16**, 177–180.

Received on 8 December 1986

Note added in proof

We have isolated a new cosmid clone (II 3.31) which extends the *S* region gene cluster by 30 kb upstream of *21-OHB*. A probe isolated from the left end of II 3.31 hybridizes to the same 270 kb *Mlu*I fragment picked up by probe 2. Thus the orientation and map position shown in Figure 1a for the *S* region gene cluster is supported by an overlapping *Mlu*I fragment as well.