Far upstream regions of class II MHC Ea are necessary for position-independent, copy-dependent expression of Ea transgene

Susan Carson and Michael V. Wiles
Molecular Immunology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK and Basel Institute for Immunology, Postfach, CH-4005 Basel, Switzerland

Received February 23, 1993; Revised and Accepted March 29, 1993

ABSTRACT

The chromatin upstream of the class II MHC Ea gene contains specific, DNase I hypersensitive (DH) sites (groups I–V), overlapping and extending the promoter proximal and distal control regions. To determine whether the Ea DH groups I–V define a functionally important chromatin domain or locus control region (LCR), we have used wild type Ea gene constructs to generate transgenic mouse lines from strains that do not express an endogenous Ea gene product. Constructs contained either DH groups I–V ‘Longs’ or DH groups I–II ‘Shorts’, of the hypersensitive sites defined within 20 kb 5’ of Ea. We show that position-independent, copy number-dependent expression of the Ea gene occurs only with the Long construct (8/8 transgenic mouse lines, over a range of copy numbers, 1–30 copies); in contrast, the Short constructs are subject to position-dependent effects. This suggests that the region delineated by Ea DH groups I–II is necessary but not sufficient as an LCR, which requires the presence of the upstream regions containing DH III–V for complete position-independent, copy number-dependent expression. These results introduce an Immunologically-important, putative LCR which can be used to target genes to cells of the B cell lineage, as well as to other class II MHC expressing cells, and highlight the importance of chromatin structure analysis as a means to locate DNA regions of regulatory interest which are dispersed over a large distance.

INTRODUCTION

The major histocompatibility complex (MHC) is a polymorphic multigene family encoding two classes of cell surface glycoproteins class I and class II, and has been extensively studied in mice and humans. The mouse class II MHC region contains six genes (Ob-Ab-Aa-Eb1-Eb2-Ea) encoding two cell surface heterodimers (AαAβ and EaEβ) which are expressed on key components of the immune system, including B cells, dendritic cells and can be induced by γ-interferon on other antigen presenting cells, such as macrophages. This pattern of expression reflects their role in foreign antigen processing and recognition by T cells. (For reviews and references therein: (1–6)). The mechanisms underlying the tissue-specific regulation of expression of class II MHC molecules are still not fully understood. Studies defining the immediate 5’ flanking sequences of class II MHC genes and their associated DNA-binding proteins (5–10), as well as transgenic mouse models containing DNA constructs covering the promoter proximal sequences have been examined in detail by other groups (5,11–13). However, the influence of more distant cis-acting sequences/regions on class II MHC expression has not yet been characterised.

In other systems, far upstream and/or downstream sequences of tissue-specific genes or gene families have been shown to be involved in position-independent, copy number-dependent expression of associated genes in transgenic mice. These locus control regions (LCR) were first defined for the β-globin system (14,15) and their ability to orchestrate lineage specific expression of sequences linked in cis has made them key elements in the development of gene therapy technology. As the class II molecules are pivotal to the immune response, the class II locus may be an important future target for gene therapy. Therefore, it was important to reexamine the regulation of class II genes in transgenic mouse models, in order to determine whether far upstream regions were required for position-independent, copy number-dependent expression as in documented LCRs.

Such regions of regulatory interest have often been shown to correlate with the presence of DNase I hypersensitivity (DH) sites, as in the case of the β-globin locus, and other developmentally regulated gene families (16,17). These DH sites represent nucleosome-free regions in chromatin which are accessible to trans-acting proteins. Such proteins can include those involved in DNA conformational changes (18, for review) or in structure recognition, such as the high mobility group proteins (HMGs) (19–21) as well as, sequence-specific transcriptional regulatory proteins. In fact, the establishment of a DH site most probably represents one of the earliest steps necessary to commit a gene to be functionally active in a particular cell. Also, specific DH sites have been used to pinpoint elements, not necessarily
sequenced or with a known function, that appear to control the accessibility of chromatin domains to trans-acting regulatory factors, again even before a gene is expressed (22–24). As the action of an LCR is postulated to be mediated by a variety of transacting factors interacting with promoter/enhancer type sequences, with subsequent effects on chromatin domains (25, and references therein); we would expect chromatin structure analysis (DH sites) to be of particular use in localising regions of regulatory interest.

Following this line of reasoning, we have recently defined a series of DNase I hypersensitive sites flanking the class II MHC locus and have shown that these sites are specific to cell lines which are developmentally programmed to express the class II MHC gene products (i.e. B cell lineage cell lines), or to those actually expressing class II MHC at the cell surface (26). The DH sites mapping within 20 kb upstream of the Ea gene (Ea DH I–V) were particularly suited to further analysis, as one of the earliest transgenic mouse models of immunologically important genes made use of inbred mouse strains (H-2b or H-2d haplotype) which do not express surface I-E due to a specific 650 bp deletion covering part of the Ea promoter and first exon, while Eb expression remained normal (27–29). In what are now classical studies, when wild-type Ea gene constructs of either the H-2b or H-2d haplotype (30–32) were injected into oocytes of the H-2b/H-2d strains, appropriately regulated, cell surface expression of the E antigen was seen and E-restricted responses were restored.

We have made use of this model system and measured Ea mRNA abundance, in order to study the functional importance of the Ea upstream regions in transgenic mice. We discuss whether the regions containing these apparently developmentally programmed patterns of DNase I hypersensitive sites reflect the action of a locus control region (LCR) by conferring tissue-specific, position-independent, copy-dependent expression on a wild-type Ea gene.

**MATERIALS AND METHODS**

**Transgenic mice**

Mouse cosmid 32.1 (BALB/c genomic library, H-2d haplotype (33) was obtained from Dr M. Steinmetz. The Long construct, containing Ea DHS I–V were a 35 kb Mlu I/Xho I fragment isolated from cosmid 32.1 (Figure 1a–c). The Short construct, containing Ea DHS I–II was a 9kb fragment obtained from restriction enzyme digests containing Ea DHS I–II strains, appropriately regulated, cell surface expression of the E antigen was seen and E-restricted responses were restored.

Oocytes were obtained from C57Bl/6 x SJL F1; both these inbred mouse strains, have a deletion covering the promoter and part of the Ea exon 1 and therefore do not express their own endogenous Ea (27–29). The oocytes were microinjected with the DNA constructs, following established protocols (35). Offspring were tested for the transgene by PCR on tail-derived DNA. PCR primers covering the deletion present in promoter of the Ea upstream regions in transgenic mice. We discuss whether the regions containing these apparently developmentally programmed patterns of DNase I hypersensitive sites reflect the action of a locus control region (LCR) by conferring tissue-specific, position-independent, copy-dependent expression on a wild-type Ea gene.

DNA and RNA analysis

DNA was isolated from cell lines or tissues by established methods (34). DNA from transgenic mice was digested with restriction enzymes and electrophoresed through agarose gels to determine the integrity of the construct as well as copy number. Integrity of the construct was determined using Mlu I/Xho I restriction enzyme digests and Kpn I digests. The Mlu I/Xho I digests were electrophoresed using a Field Inversion Gel Electrophoresis apparatus in order to resolve large DNA fragments, following protocols previously described (37).

The Kpn I digests were electrophoresed on 0.8% agarose gels and Southern blotted to determine copy number (see Figure 2d and figure legend). Under these conditions we detect a single band (3.6 kb Kpn I fragment) which is equivalent to the two dysfunctional endogenous host Ea genes, or a band at 4.2 kb in BALB/c mice. In Ea transgenic mice, two bands are resolved, the endogenous copy plus the transgene, at 4.2 kb. As such, the endogenous gene was used as a DNA loading reference representing two copies. Thus by comparing the signal ratio of the Ea endogenous gene to transgene signal, we could directly relate signal to copy number. Copy number determination used a PhosphorImager (Molecular Dynamics, Sunnyvale, CA, USA.) permitting linear quantification of signal intensity (38).

DNA was isolated from different tissues following the AGPC technique (39). DNA was analysed by dot blots using Gene Screen Plus (Du Pont Company, NEN Products Boston, MA, USA) with conditions as suggested by the manufacturer. Probes for the hybridisation were 32P-labelled by random priming (40). The relative abundance of β-actin, Ea and Eb RNAs were determined in duplicate in at least three experiments. β-actin was used to control for RNA loading, and Eb was used as an internal control for the number of class II MHC expressing cells. Eb RNA expression was assumed to represent two copies of the Eb gene per genome, and was approximately constant as compared to β-actin for all transgenic mouse lines tested. All mice were normalised to BALB/c expression of Ea and Eb genes, representing two copies per genome, and Ea transgenes are expressed per gene copy per genome.

The probes used for Southern and Northern Blot hybridisation were Ea cDNA (41), Eb cDNA (42), Ea 695 bp (Pst I/Kpn I fragment covering the second exon of Ea) and 5'Ea 1.1 kb (Sac I/Sal I 1.1 kb fragment isolated from cosmid 32.1) (26) and β-actin cDNA (43).

**DNase I hypersensitivity assay**

The conditions used for the DNase I hypersensitivity assay were as described (26) using the technique of indirect end labelling (44). Increasing concentrations of DNase I (Boehringer-Mannheim) were expressed as units per 10⁶ nuclei. Incubations were for 10 minutes at 37°C.
Nuclei were isolated from cell lines or spleen from transgenic animals or inbred mouse strains. Wehi 231 is a mouse (BALB/c) B cell lymphoma line (45). The probes used for indirect end labelling were: Ea 695 bp (Figure 4b), and 5'Ea 1.1 kb (Figure 4A) (See Figure 4D for exact location).

RESULTS
Transgenic mice containing Long and Short Ea constructs
The rationale behind these experiments was that, if sequences previously 'localised' by the DNase I hypersensitivity (DH) sites were of functional importance, then their presence or absence would have an effect on the regulated expression of the Ea transgene, as shown by position-independent and copy-dependent expression. Classic work (11-13) had shown that the promoter proximal and B cell control regions (delineated in our work by Ea DH I-II) were essential to tissue specific expression, and it was therefore necessary to include these sequences in both Short and Long constructs (Figure 1) used in the present work. The Short constructs contain the Ea\textsuperscript{d} gene in addition to 2 kb of 5' Ea sequence and 1.4 kb of 3' Ea sequence isolated as a single Bgl I fragment (Figure 1d). The Long constructs contain the Ea\textsuperscript{d} gene, in addition to 23 kb of 5' Ea sequence and 4.5 kb of 3' Ea sequence (Figure 1c) and was designed to show whether the presence of the upstream regions containing DH groups III-V were of functional importance to position-independent, copy-dependent expression of the Ea transgene. No additional DNase I hypersensitive sites other than DH sites I-V (Figure 1b) were found in the upstream or downstream regions present in the Long construct (data not shown). The DNA region upstream of DH site V was present in order to permit the isolation of a single contiguous fragment of DNA by the restriction enzymes Mlu I/Xho I (Figure 1c). Eight independent Ea\textsuperscript{d} Long (Figure 1c) and two Ea\textsuperscript{d} Short (Figure 1d) transgenic mice lines were generated (see Materials and methods). All Ea\textsuperscript{d} transgenic mice lines were healthy and transmitted the Ea\textsuperscript{d} transgene to their offspring.

Both Long and Short constructs are expressed in transgenic mice
We analysed ten, newly generated transgenic mouse lines, plus two pre-existing transgenic lines Ea-16 and We32-25 (30, 36) for Ea transgene copy number and abundance of Ea RNA in spleen (Figure 2 and Table 1). Ea transgene copy number (Figure 2D) was determined by Southern blots of DNA derived from transgenic mice spleens, digested with the restriction enzyme Kpn I, and then hybridised using a DNA probe specific for the second exon of Ea. The signal was quantified by PhosphorImager analysis and copy number normalised to the two endogenous Ea copies/genome present. The abundance of Ea RNA was determined by dot blot analysis, the signals obtained were quantified by PhosphorImager and normalised for loading using a \beta-actin cDNA probe.

By comparison to the endogenous Ea gene (C57Bl/6×SJL F1), the transgenic animals have Ea\textsuperscript{d} copy numbers ranging from 1-200 (Figure 2D). BALB/c spleen RNA was used as an Ea positive control, its expression level being equivalent to two intact copies of the Ea gene. The ratio of Ea to \beta-actin was determined by dot blot analysis, the signals obtained were quantified by PhosphorImager and normalised for loading using a \beta-actin cDNA probe.

Figure 1. Map of the class II Ea gene and transgene constructs. The map (a) shows the cosmid 32.1 (33) containing the Ea\textsuperscript{d} gene used in this study. Filled boxes indicate exons. The letters X, Y and W refer to control elements present in the upstream region of Ea and previously described as proximal promoter (through to -200 bp from the cap site) region and B cell control region (-1kb through to -2kb) respectively (5). The arrowed solid bar refers to a deletion of 650 bp covering the promoter region and the signal sequence present in the b and s haplotype (28). Sizes are in kilobase (kb). The arrows underneath the line drawing (b) refer to the locations of the DNase I hypersensitivity sites present in the upstream region of Ea: DH I (-100 bp), DH II (-1.2kb), DH III (-3.4 kb), DH IV (-5.4 kb) and DH V (-8.4 kb) (26). The large arrows refer to the central location of DH sites and the smaller arrows refer to minor DH sites. No additional B cell hypersensitive sites were found in the region upstream of Ea DH V. 'Long'(c) and 'Short'(d) refers to the constructs isolated from cosmid 32.1 using the Bgl I fragment (Figure 1d). The Long constructs contain the Ea\textsuperscript{d} gene in addition to 2 kb of 5' Ea sequence and 4.5 kb of 3' Ea sequence (Figure 1c) and was designed to show whether the presence of the upstream regions containing DH groups III-V were of functional importance to position-independent, copy-dependent expression of the Ea transgene. No additional DNase I hypersensitive sites other than DH sites I-V (Figure 1b) were found in the upstream or downstream regions present in the Long construct (data not shown). The DNA region upstream of DH site V was present in order to permit the isolation of a single contiguous fragment of DNA by the restriction enzymes Mlu I/Xho I (Figure 1c). Eight independent Ea\textsuperscript{d} Long (Figure 1c) and two Ea\textsuperscript{d} Short (Figure 1d) transgenic mice lines were generated (see Materials and methods). All Ea\textsuperscript{d} transgenic mice lines were healthy and transmitted the Ea\textsuperscript{d} transgene to their offspring.

Both Long and Short constructs are expressed in transgenic mice
We analysed ten, newly generated transgenic mouse lines, plus two pre-existing transgenic lines Ea-16 and We32-25 (30, 36) for Ea transgene copy number and abundance of Ea RNA in spleen (Figure 2 and Table 1). Ea transgene copy number (Figure 2D) was determined by Southern blots of DNA derived from transgenic mice spleens, digested with the restriction enzyme Kpn I, and then hybridised using a DNA probe specific for the second exon of Ea. The signal was quantified by PhosphorImager analysis and copy number normalised to the two endogenous Ea copies/genome present. The abundance of Ea RNA was determined by dot blot analysis, the signals obtained were quantified by PhosphorImager and normalised for loading using a \beta-actin cDNA probe.

By comparison to the endogenous Ea gene (C57Bl/6×SJL F1), the transgenic animals have Ea\textsuperscript{d} copy numbers ranging from 1-200 (Figure 2D). BALB/c spleen RNA was used as an Ea positive control, its expression level being equivalent to two intact copies of the Ea gene. The ratio of Ea to \beta-actin was determined by dot blot analysis, the signals obtained were quantified by PhosphorImager and normalised for loading using a \beta-actin cDNA probe.
in the spleen of all transgenic mice was equivalent to that in BALB/c and C57Bl/6 controls (data not shown, Table 1). The cell surface expression of the Ea\(^d\)/Eb\(^h\) heterodimer, remained constant over the range of Ea copy-numbers, being equivalent to that seen in the control BALB/c mouse spleen (data not shown). All of the Ea transgenic mouse strains contained intact copies of the transgene construct, as shown by Southern blot analysis (data not shown, see Materials and methods), and were probably integrated in tandem.

Long but not Short constructs give rise to position-independent, copy number-dependent expression of the Ea transgene

Figure 2A shows both the expression level and the copy number of the Ea transgene on the same bar graph. Transgenic mice containing the Short construct show position-integration effects at higher copy numbers and are therefore not under complete or appropriate control. Where the Ea copy number and the abundance level match, position-independent, copy number-dependent expression has been achieved, with the level of expression per gene copy being directly equivalent to that in BALB/c mice. This is clearly the case for all transgenic mice containing the Long construct, showing a copy number-dependent expression over the range of 1 to 30 copies.

The data shown in Figure 3 indicates the linear nature of this position-independent, copy number-dependent expression in Long transgenics. The Ea expression level obtained with the Long construct mice approaches the actual ratio of that seen in BALB/c mice of one times the expression level per genomic copy of the Ea\(^d\) gene. In contrast, the expression level observed with the Short construct mice deviates markedly from the expected ratio and position effects are present above five copies/genome. This is not due to a simple depletion of transcription factors as twelve copies of the transgene in Long 7, and 30 copies in Long 12 gave the expected expression ratio. The Eb/\beta-actin expression remains relatively constant for all of the transgenics (see Table 1 (Long 11 is the exception, see below) and data not shown). This indicates that the number of B cells/spleen, as well as B cell viability is unaffected by these high levels of Ea expression. When the Ea transgene expression was normalised to Eb expression instead of to \beta-actin alone, the Ea transgene expression was again shown to be position-independent and copy number-dependent for the Long constructs (final column in Table 1, including Long 11). However, the surface expression of the class II Ea/\beta heterodimer remains constant irrespective of Ea RNA abundance levels, this is probably due to limiting amounts of Eb RNA (data not shown).

Copy number-dependent DNase I hypersensitive sites are maintained in transgenic mice

If the DH sites of the integrated transgenes were of functional use they should be maintained in an appropriate copy-number dependent fashion in transgenic mice. Using high copy number Long transgenics, one would expect a copy-dependent signal in the DNase I hypersensitivity assay. Figure 4A and 4B show DNase I hypersensitive assays performed on BALB/c and transgenic mice spleens as well as a BALB/c B cell line, Wehi 231, using approximately equal number of nuclei (Figure 4C).
PhosphorImager analysis of Southern blots showed that all of the Ea DH sites had been reestablished in the Long 12 mouse (see figure legend to Figure 4C). DH numbers used refer to the previously described hypersensitivity sites which are present during all stages of B cell development (26). The map of the region is indicated below the digests (Figure 4D) and shows the DNA probes used as well as the position of the DH sites. That the upstream hypersensitive sites are reformed appropriately, lends weight to the hypothesis for a functional role of these sites in the expression of the Ea gene. In addition, the intensity of the sub-bands shown in Figure 4 (A – C) shows that the factors which establish these sites are not limiting. Further, the presence of these hypersensitive sites is appropriately regulated and these sites are found only in those tissues which normally express class II MHC, as DH sites corresponding to the Ea transgene were not detected in brain tissue isolated from Long 12 transgenics (data not shown).

**DISCUSSION**

We have shown in this paper that position-independent, copy number-dependent expression of the Ea gene in transgenic mice, requires far upstream regions containing five DNase I Hypersensitive sites (DH 1–V). These DH sites are present upstream of the Ea gene in a variety of transformed cells at different stages of B cell development, but not in non-B cell lines that do not express Class II MHC (26). These areas of hypersensitivity are now shown to be functionally important in vivo, as they mark regions which comprise a putative locus control region (LCR). This is the first time that strict position-independent, copy number-dependent expression has been shown for a mouse gene in a transgenic mouse model, and these mice now provide a source of a stably-integrated transgene of immunological importance. In addition, the use of these hypersensitive sites as markers for regions which are of regulatory importance as well as for use in gene targeting experiments is now possible.

**An LCR among the Ea upstream regions?**

By definition, any DNA sequence/region that does not permit position-independent, copy number-dependent expression of a linked gene in all transgenic mice, is not an LCR. We therefore, compared 12 independently derived mouse lines; including 8 Longs, 2 Shorts (Figure 1), as well as the previously described Ea-16 and We32-25 (30,36), analysing their DNA and RNA quantitatively using a PhosphorImager. In our analysis, copy-dependent expression (Figures 2 and 3) is lost above five copies of Ea in the Short transgenics indicating that the region containing DH I and II (2.0 kb 5' of Ea) does not contain the appropriate elements to isolate the transgene from adverse chromatin effects, or does not contain all of the enhancer elements necessary for position-independent, copy number-dependent expression. Therefore, the DNA upstream of the Dh site II is critical, as only the Long transgenics show a strict site-independent integration and copy number-dependent expression of the Ea construct (Figure 3). The position-dependent effects seen in Short transgenic mice were mentioned previously by other groups who reported that only 3/8 mice containing 2.0 kb 5' of Ea expressed the transgene at all (11). Such integration effects were also seen using the Short construct plus or minus the X and Y box, as the phenotype of interest (thymic expression) was not seen in all cases (46). In addition, the exact Ea copy number of the transgenic mice from other groups is not known, and/or the number of mice tested was not large. However, based on the results reported in this paper, we can say that the position effect seen in the Short transgenic mice, is probably due to the absence of the upstream regions containing DH 111–V. In these cases, this is not due to a limitation of transcriptional factors, since the Long transgenics have a range of copy number and equivalent expression per gene from 1 to 30 copies (Figures 2 and 3).

Further support for the functional importance of the regions containing these DH sites in establishing position-independent integration of the transgene, is reflected in the copy number-dependent reestablishment of the DH sites in the Long 12 transgenics (Figure 4). This shows that the presence of regulatory factors necessary for the establishment of a stable, open chromatin domain is not limiting. This also implies that the Eb message or protein is the limiting factor in the expression of the Eu/β heterodimer on the cell surface of these mice. This was confirmed by the unchanged level of cell surface expression of the E complex in transgenics as compared to BALB/c controls (data not shown, (47)).

Table 1. Summary of transgenic expression data

<table>
<thead>
<tr>
<th>Mouse</th>
<th>E alpha copy number</th>
<th>*Expression level E alpha</th>
<th>E beta (2 copies)</th>
<th>Ea/Eb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balb/C</td>
<td>2</td>
<td>2.00</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>C57BL6</td>
<td>α2</td>
<td>0.0</td>
<td>1.6</td>
<td>–</td>
</tr>
<tr>
<td>L5</td>
<td>1</td>
<td>0.8</td>
<td>1.8</td>
<td>0.9</td>
</tr>
<tr>
<td>L8</td>
<td>1</td>
<td>0.9</td>
<td>2.0</td>
<td>0.9</td>
</tr>
<tr>
<td>L6</td>
<td>3</td>
<td>1.6</td>
<td>1.8</td>
<td>1.8</td>
</tr>
<tr>
<td>L10</td>
<td>4</td>
<td>3.7</td>
<td>1.7</td>
<td>4.4</td>
</tr>
<tr>
<td>L11</td>
<td>6</td>
<td>3.5</td>
<td>0.8</td>
<td>8.8</td>
</tr>
<tr>
<td>L12</td>
<td>12</td>
<td>10.2</td>
<td>1.6</td>
<td>12.8</td>
</tr>
<tr>
<td>L9</td>
<td>30</td>
<td>28.8</td>
<td>2.3</td>
<td>25.0</td>
</tr>
<tr>
<td>S3</td>
<td>3</td>
<td>3.5</td>
<td>2.0</td>
<td>3.5</td>
</tr>
<tr>
<td>We32-25</td>
<td>2.7</td>
<td>n/d</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ea alpha 16</td>
<td>12</td>
<td>1.3</td>
<td>n/d</td>
<td>–</td>
</tr>
<tr>
<td>S1</td>
<td>200</td>
<td>10.1</td>
<td>1.6</td>
<td>12.6</td>
</tr>
</tbody>
</table>

* RNA abundance level normalised to β-actin, and then expressed in comparison to Balb/C (i.e. two copies equals an abundance of 2).
∞ indicates that the gene is present (two copies) but in a non-functional state.
Evidence for tissue specific, copy-dependent, position-independent expression has been shown mainly in cross-species transgenics; such as, human β-globin (14, 15, 48), CD2 (49), α-globin (50), and class I HLA-B7 (51, 52) genes as well as for chicken lysozyme (53), although absolute correlation between expression, tissue distribution and copy number has not been seen in all cases. This latter point, could be due to species differences involving transacting factors regulating appropriate expression, or variations in quantitation techniques and the number of transgenic lines analysed. The evidence in this paper for the class II MHC Ea gene supports the view that, in principle, all tissue specific genes/multigene families would have a 'locus control region' as presently defined. Whether this includes a combination of enhancer-type elements which overcomes any 'silencing' effects of the surrounding chromatin on the integrated transgene, and/or elements which interact with the nuclear matrix directly to isolate the transgene from inactive chromatin (54–56), needs further analysis of potential locus control regions.

The presence of these regulatory regions, stably integrated and of a known copy number among the Ea upstream regions, permits the transgenic mice described here to be used in the following studies. It is now possible to investigate whether it is 'just' extra DNA sequence, with no apparently defined hypersensitive sites, which leads to position-independent expression of the transgene by designing a minimal 'Long' construct containing only the defined DH sites (experiments in progress). It is also possible that the upstream DH sites mark the presence of other genes which either, have a direct effect on the expression of Class II MHC or more generally, on commitment to the B cell lineage. For these studies on early B cell ontogeny, the Long transgenic mice, particularly those with high copy numbers, could be used in recently described stromal/factor-dependent systems which give rise to non-transformed pre B/progenitor B cell lines (57–60). As the factors associated with the presence of these hypersensitive sites are not limiting, recent advances in in vitro footprinting (61) should permit the dissection of which particular proteins are binding to the DH site sequences indicated in this work. Additionally, the use of easily visualised reporter genes could be combined with these DH sites in order to visualise earlier stages of commitment, as the pattern of DH sites (Figure 4) precedes the expression of class II MHC in cell lines of the B cell lineage (26). The reporter gene would then serve as a marker of when the chromatin became 'open' to regulatory and transcription factors. Recent work using lac Z gene constructs
has shown that as few as 1/1000 cells (62) can be visualised, which would be a potentially powerful tool in visualising early stages of B cell commitment and differentiation in vivo.

These transgenic mice are also of use in more classical immunological studies. In previous work, using Ea transgenic mice (Figure 1c, (47,63) or recombinant mice strains (64), the presence of AβEa mixed isotype molecules at the cell surface was shown. One would assume that such AβEα interchain hybrids would also be present in the higher copy number transgensics reported here. In addition, the Ea transgene concentration at either message or cytoplasmic protein level, is apparently not toxic to the cell. This may be due to the complete absence of functional Ea and/or non-polymorphism of the Ea gene product. In contrast, it has been reported that the products of Aα and Ab transgenes compete with the products of endogenous genes in a way that is detrimental to the immune system (65,66).

This is in contrast to other groups who showed more limited effects on the immune system of the transgenic animals (67,68). However, an upregulation of E molecules at the cell surface might be expected in mice with high levels of both Eb and Ea transgenes, as had been seen for Ab and Aa transgenes (68). This may be of use in looking at early events, such as occur during the establishment of tolerance; thereby facilitating the task of looking at low numbers of cells.

DH sites as markers for interesting regions

In the constructs described in this paper, only 2 kb of 23 kb of DNA from the Ea upstream region have been sequenced. The presence of DH sites provides a convenient ‘marker’ for regions which should be isolated and sequenced for further study. The fine analysis of the DH sites upstream of the Ea gene has been hampered by the presence of repetitive/unstable sequences making them difficult to subclone, and thereby reduce the upstream regions to an absolute minimum as in the β-Globin LCR studies (14). Alternatively, it should be possible to disrupt specific DH sites by homologous recombination, as has been recently described for the β-globin LCR (69). This could be used to define which regions are important for class II MHC Ea/locus expression, and to attempt to generate regulatory mutants that do not express class II MHC genes at all, but leave the locus intact. Further, the homologies between mouse and human MHC complexes, suggest that equivalent regions may exist far upstream of the human equivalent to Ea (DRA). Hypersensitive regions have been known to exist immediately upstream of the DRα gene (70), and recently the absence of DH sites in class II deficient, Epstein–Barr virus-transformed cell lines established from patients with Combined Immunodeficiency syndrome have been reported (71).

In summary, in addition to previously described erythroid specific, T cell specific, macrophage specific LCRs, we have provided evidence in this paper for a putative Locus Control Region for the class II MHC, which would allow specific expression of genes to cells of the B cell lineage, as well as to cell types where class II can be expressed. The transgenic mouse lines also provide a source for future studies on early developmental stages of the immune system. Finally, we have shown here that DNase I hypersensitivity sites can be used to pinpoint interesting regions of specific importance to the expression of the class II MHC Ea gene, both as composing part of a locus control region, and as possible markers for B cell commitment. Whether these upstream regions contain other genes, or cis-acting, regulatory elements specific for class II Ea MHC and/or the entire class II MHC locus, they point the way for the analysis of other ‘interesting’ sites in both the mouse and human systems.

ACKNOWLEDGEMENTS

The authors would especially like to thank Dr Jim Kaufman for his thorough and lucid critique of this paper, and Drs Frank Grosveld and Harold van Boehmer for critical and helpful comments. We would also like to thank Urs Millet of the Basel Institute for Immunology for the generation of the transgenic mice used in this study, and Britt Johannson and Kathleen Roderick for excellent technical assistance. Many thanks to Marina Kuhn and Alexandra Livingstone, Marie Kosco, Doris Scheidegger, and Barbara Hausmann for much appreciated help and advice during the course of this work. Additional thanks are due to the Animal House Staff in Basel and in London (in particular C. Hetherington, K. Sneglar and T. Norton) for excellent care and breeding of the transgenic mouse lines. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche Ltd. This work was supported by the Medical Research Council, UK.

REFERENCES