Human HMG box transcription factor HBP1: a role in hCD2 LCR function

Talgat Zhuma, Richard Tyrrell1, Belaid Sekkali, George Skavdis, Alexander Saveliev2, Mauro Tolaini, Kathleen Roderick, Trisha Norton, Steve Smerdon1, Steve Sedgwick3, Richard Festenstein1 and Dimitris Kioussis4

Division of Molecular Immunology, 1Division of Protein Structure and 3Division of Yeast Genetics, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK
2Present address: Department of Medicine, Imperial College of Science, Technology and Medicine, MRC Clinical Sciences Centre, Hammersmith Hospital, Du Cane Road, London WC12 0NN, UK
4Corresponding author

The locus control region (LCR) of the human CD2 gene (hCD2) confers T cell-specific, copy-dependent and position-independent gene expression in transgenic mice. This LCR consists of a strong T cell-specific enhancer and an element without enhancer activity (designated HSS3), which is required for prevention of position effect variegation (PEV) in transgenic mice. Here, we identified the HMG box containing protein-1 (HBP1) as a factor binding to HSS3 of the hCD2 LCR. Within the LCR, HBP1 binds to a novel TTCATT-CATTCA sequence that is higher in affinity than other recently reported HBP1-binding sites. Mice transgenic for a hCD2 LCR construct carrying a deletion of the HBP1-binding sequences show a propensity for PEV if the transgene integrates in a heterochromatic region of the chromosome such as the centromere or telomere. We propose that HBP1 plays an important role in chromatin opening and remodelling activities by binding to and bending the DNA, thus allowing DNA–protein and/or protein–protein interactions, which increase the probability of establishing an active locus.

Keywords: chromatin/HBP1/HMG box/locus control region/position effect variegation

Introduction

Locus control regions (LCRs) direct expression of linked heterologous and homologous transgenes in a tissue-specific, position-independent, transgene copy number-dependent manner (Grosveld et al., 1987). Studies on the human CD2 (hCD2) and β-globin LCRs have demonstrated that the LCR achieves the position-independent expression in T cells and erythroid cells of transgenic mice, respectively, by overcoming heterochromatin-mediated position effect variegation (PEV; Festenstein et al., 1996; Milot et al., 1996a). The hCD2 LCR consists of at least two distinct elements: a transcriptional enhancer and a region without enhancer activity (designated HSS3) necessary for preventing PEV (Lake et al., 1990; Festenstein et al., 1996). The functional differences between the LCR and enhancer were evident from the early experiments on the human β-globin LCR using transgenic and cell transfection approaches. It has been shown that whereas the HSS3 and HSS4 regions of the β-globin LCR hardly increase the level of expression of the linked β-globin gene in cell transfection assays, they are able to render expression of the transgene copy number-dependent and site integration-independent (Talbot et al., 1990; Pruzina et al., 1991). Modular structures of LCRs have also been described for the human ADA (Aronow et al., 1995) and the T-cell receptor α (αTCR; Ortiz et al., 1997) gene loci. It appears, therefore, that LCRs consist of multiple elements such as enhancers and sequences necessary for establishing an open chromatin configuration, thereby preventing mosaic expression of transgenes when they are integrated into heterochromatic areas such as telomeres or the centromeres (Kioussis and Festenstein, 1997).

PEV, or the phenomenon of mosaic expression within a cell lineage of genes placed adjacent to a euchromatin–heterochromatin junction, has been well characterized in yeast and Drosophila (Allshire et al., 1994; Weiler and Wakimoto, 1995; Grunstein, 1997; Wallrath, 1998). Although PEV in mammals has not been studied in detail, certain characteristics of PEV described in mice bear a strong resemblance to those seen in other organisms. For instance, it appears that the decision whether to express the hCD2 transgene carrying the hCD2 LCR with HSS3 deleted is a stochastic one and, once made, it is maintained through subsequent cell divisions (Kioussis and Festenstein, 1997, 1998). Similar observations were made in experiments on human β-globin (Milot et al., 1996a) and β-lactoglobulin transgenic mice (Dobie et al., 1996). Mosaic expression patterns of transgenes have also been described in other systems (Pravtcheva et al., 1994; Robertson et al., 1995; Guy et al., 1996). Although most of the studies on PEV in mammals have been carried out using transgenic technology (Garrick et al., 1996; Dobie et al., 1997; Kioussis and Festenstein, 1997; Grosveld et al., 1998), a growing body of evidence has emerged implicating position effects in several human diseases caused by chromosomal translocations (Milot et al., 1996b; Kleinjan and van Heyningen, 1998). To date, experiments indicate that LCRs have the ability to overcome PEV effects. However, in spite of many exciting breakthroughs, the mechanism of action of the LCRs remains a subject of hot debate (Higgs, 1998) and very little is known about proteins that bind to LCRs and mediate the chromatin-opening and PEV-preventing function.

In this study, we identified HBP1 as a protein that binds to HSS3 and contributes to the hCD2 LCR function of...
preventing PEV. HBP1 belongs to the HMG family of proteins, with LEF1 being the closest homologue (Travis et al., 1991). HBP1 was cloned originally as a protein able to complement a potassium channel defect in yeast (Lesage et al., 1994). The function of this protein remains the subject of current investigation, and recently several groups have reported the ability of HBP1 to interact with proteins of the retinoblastoma family, to induce morphological transformation of cells in culture and to act as a transcriptional repressor of the cyclin D1, p21 and N-myc genes (Lavender et al., 1997; Tevosian et al., 1997; Gartel et al., 1998; Shih et al., 1998; Yee et al., 1998). Here we show that HBP1 binds to the HSS3 sequences within the hCD2 LCR and appears to contribute to the prevention of PEV. Our data suggest that the role of HBP1 in the cell cycle and differentiation may be mediated via chromatin remodelling activity and the establishment of open chromatin structures.

Results

Yeast one-hybrid assay identifies human HBP1 as binding to HSS3 of the hCD2 LCR

To identify trans-acting factors that bind to the HSS3 region of the hCD2 LCR and, potentially, determine the LCR function, an in vitro DNase I footprint analysis of the HSS3 region was performed using nuclear protein extracts. The DNase I footprinting of the 116 bp Ndel–BsrGI fragment from HSS3 resulted in identification of the FT1 and FT2 regions protected with both thymus and liver extracts (Figure 1). Analysis of the footprinted sequences revealed that the FT1 region contains a TTCA motif, which is repeated five times within the region and is also present in the FT2 region (Figure 1). To identify proteins binding to the FT1 and FT2 sequences, the Matchmaker (Clontech) Jurkat T-cell cDNA library was screened using the one-hybrid assay in yeast. Using the FT1 sequence as a bait, we isolated a library plasmid carrying a 1.3 kb insert homologous to the rat HBP1 cDNA (Lesage et al., 1994). Sequence comparison has shown that the isolated 1.3 kb fragment consists of a 0.35 kb region encoding the C-terminal part of the protein including the HMG box DNA-binding domain plus 0.95 kb of the 3′-untranslated region (3′-UTR). A full-length open reading frame (ORF) cDNA of the human HBP1 (Figure 2A) was subsequently isolated from the Jurkat λ-ZAP cDNA library (Stratagene) screened with the 1.3 kb hHBP1 cDNA probe isolated in the yeast one-hybrid screen. Sequence comparison of the isolated ORF of the human HBP1 protein with that of the rat HBP1 revealed 91% homology in the region coding for the HMG box domain and 89% in the rest of the sequence. At the protein level, homology was 100% in the HMG domain and 91% in the N-terminal part of the protein.

hHBP1 HMG box binds to a novel TTCATTCA motif within the hCD2 LCR

To identify the precise HBP1-binding site within the LCR, a recombinant HMG box domain (residue 433–513) protein was overexpressed in Escherichia coli using the pET-22b expression system (Novagen). The recombinant protein was used to footprint a 116 bp LCR fragment containing the FT1 and FT2 protected regions identified with thymus and liver nuclear extracts. The FT1 region contains two HBP1-binding sites with apparently different affinities (Figure 2B). Thus, the ATCGTTCCATTCAAGG (FT1.S) region was visible using low concentrations of the recombinant box (25 ng), whereas protection of the other region, TACACCCTATTCAATCCTT (FT1.W), became apparent with higher concentrations (100 ng). To verify the boundaries of the FT1 footprint, we used oligonucleotides, 16 bp each, spanning a 48 bp region containing the FT1 footprint, were used in gel retardation assays (Figure 3). Such analysis showed that three oligonucleotides ATCGTTCCATTCA (FT1.J), CGTTCCATTCA (FT1.K), and CGTTCCATTCA (FT1.L) were able to complement a potassium channel defect in yeast. Here we show that HBP1 binds to the motif within the hCD2 LCR and appears to contribute to the prevention of PEV.
ATTCAAC (FT1.K) and TTCATTCAACAGA (FT1.L), sharing a common TTCATTCA motif, bind the recombinant HBP1 HMG box better than the other oligonucleotides. The TTCATTCA motif is located within the FT1.S region corresponding to the higher affinity footprint. Thus, it appears that the TTCATTCA site contains redundant binding elements, which were not digested by DNase I even in lane 1, without the recombinant protein.

ATTCAAC (FT1.K) and TTCATTCAACAGA (FT1.L), sharing a common TTCATTCA motif, bind the recombinant HBP1 HMG box better than the other oligonucleotides. The TTCATTCA motif is located within the FT1.S region corresponding to the higher affinity footprint. Thus, it appears that the TTCATTCA site contains redundant binding elements, which were not digested by DNase I even in lane 1, without the recombinant protein.

**The TTCATTCA motif is a single site.** The TTCATTCA motif identified is partly homologous to the (T/A)(T/A)CAA(A/T)G sequence which corresponds to the binding consensus for several of the HMG proteins including LEF1, TCF1, ROX1, SOX4, SRY and STE11 (Landsman and Bustin, 1993; Oosterwegel et al., 1993; Grosschedl et al., 1994). Because of this resemblance, we speculated that the identified TTCATTCA motif may, in fact, consist of two overlapping TTCATTCA binding sites.
HBP1 role in hCD2 LCR function

Fig. 3. Fine mapping of the HMG box-binding motif. (A) Gel retardation assay using the recombinant hHBP1 HMG domain and overlapping FT1.B–FT1.Q double-stranded oligos covering the FT1 region. The recombinant protein was incubated with the 5’ end-labelled oligonucleotides in the presence of excess poly(dI–dC). The intensity of the bands was assessed using a PhosphorImager. Retarded complexes are visible using the FT1.I–FT1.O oligos, with maximum intensity corresponding to the FT1.J–FT1.L probes. (B) Deletion summary diagram representing the alignment of the FT1.A sequence and the FT1.B–FT1.Q overlapping oligos covering the FT1.A area (the sense strand sequences are shown). The FT1.A sequence (underlined) comprises nucleotides from 1396 to 1446 of the hCD2 LCR containing the 32 bp FT1 footprint (Figure 1) identified with the thymus and liver nuclear extract. The TTCA TTCA TTCA motif common for the three oligos FT1.J–FT1.L, which bind the HMG box better than the others, is highlighted by a rectangle. The two arrows mark the 30 bp region deleted from the hCD2 LCR in subsequent experiments with transgenic mice. The FT1.Q oligo represents the site generated by the junction of the two 8 bp regions on either side of the deleted 30 bp sequence. In a gel retardation experiment, no binding by the HMG box was observed. The region within FT1.A used as a bait to clone hHBP1 in the one-hybrid assay is highlighted by larger letters.

hHBP1 box-binding site then the GGCATTCCATCA (FT1.J.mC) and TTCATTCCAGGCA (FT1.J.mE) mutants would still be able to bind to the HMG box through the remaining intact TTCATTCCA motif. If, however, TTCATTCCA was a single binding site, then these mutant oligonucleotides would not be able to bind the protein. Similar mutations had been reported to abolish the TCF1- and SOX4-binding sites (Oosterwegel et al., 1991; van de Wetering et al., 1993). The gel retardation experiment showed that the oligonucleotides corresponding to the above mutants indeed were unable to bind the hHBP1 HMG box. Thus, TTCATTCCA appears to represent a single binding site (Figure 4).

The TTCATTCCA motif has a higher affinity than other HBP1-binding sites reported. While this work was in progress, it was reported that the mouse HBP1 HMG domain is able to bind the AGAATGGG, TCAATGGG and AAAATGGG motifs located within the N-myc promoter. Three TCAATGGG motifs were also reported within the cytomegalovirus (CMV) promoter and thought to be implicated in promoter silencing induced by the mHBP1 protein (Tevosian et al., 1997; Yee et al., 1998). Because the TTCATTCCA motif identified differed from the sequences reported in the N-myc promoter, it was decided to compare the HBP1 HMG box-binding ability with the sequences found in the LCR and N-myc1
promoter. A band shift experiment (Figure 4) showed that the ATCGTTCATTCATTCA (FT1.J) oligo binds the recombinant HBP1 HMG box better than the TGCTGAGAATGGGAAG (MYC1), GACCTCAATGGGGG (MYC2) and AGTGCAAAATGGGAGG (MYC3) oligos corresponding to the 466–581, 492–507 and 317–332 sequences within the N-myC promoter (as in DDBJ/EMBL/GenBank accession No. X632811; the core motifs are highlighted by larger letters). Thus it appears that TTCATTCATTCA is the highest affinity HBP1-binding motif described to date.

The hHBP1 HMG box binding is sequence specific. To assess the DNA-binding specificity of hHBP1, we tested whether the recombinant HMG box would recognize the binding sites of other HMG box-containing proteins. The hHBP1 HMG domain does not bind to the oligonucleotides containing consensus binding sequences of the LEF1/TCF1, SOX4 and SRY proteins (Figure 4) (Giese et al., 1991; van de Wetering et al., 1991, 1993). Thus, it appears that the recombinant hHBP1 HMG domain binds to the novel TTCATTCATTCA motif of the LCR in a sequence-specific manner.

Recombinant HBP1 HMG domain has DNA-binding characteristics identical to the binding activity found within a T-cell nuclear extract.

To demonstrate that the protein present in the thymus nuclear extract which binds to the FT1 sequences is HBP1, a band shift experiment was performed using the thymus nuclear extract with the FT1.J oligonucleotide containing the hHBP1 HMG box-binding motif, TTCATTCATTCA. The experiment revealed a single retarded band that could not be competed by either LEF1/TCF1- or SOX4-binding sites or other non-specific oligonucleotides (Figure 5A), demonstrating that the observed binding is sequence-specific.
of the LEF1/TCF1 proteins to the HBP1-binding motif identified within FT1.

Deletion of the FT1 region from the LCR results in PEV

It has been shown previously in our laboratory that deleting the HSS3 region from the LCR results in PEV in transgenic mice carrying a ΔHSS3-LCR hCD2 minigene (Festenstein et al., 1996). To determine in vivo the contribution of the FT1 region and, potentially, the role of P1 in the regulation of LCR function, a 30 bp sequence including the FT1 footprint was deleted from the LCR. The pattern of expression of the hCD2 minigene with such an LCR (designated ΔFT1-LCR) was analysed in transgenic mice (Figure 7A). Seven transgenic mouse lines, further referred to as FT1.1–7, were generated and the expression of the transgenic hCD2 on thymocytes was analysed by flow cytometry, which allows measurement of the hCD2 protein in individual cells. The pattern of transgene expression in the FT1.1–7 transgenic lines was compared with that in the Mg4 line carrying a hCD2 minigene integrated in the centromere. The hCD2 minigene in the Mg4 line is integrated at the centromere but contains the full LCR and, therefore, expression of the transgene does not variegate (Festenstein et al., 1996). To quantitate the degree of variegation, we estimated the percentage of thymocytes falling outside the 1.3× SD (standard deviation) interval measured from the mean of hCD2-positive thymocytes. Such calculations for the Mg4 line resulted in an average $K_{PEV}$ of 9%, indicating that the transgenic hCD2 minigene is expressed unimodally on virtually all thymocytes, so that the peak’s shape is close to the normal distribution, which corresponds to $K_{PEV} = 10\%$. Analysis of the ΔFT1 transgene expression on thymocytes has shown that in the FT1.1 transgenic line, the $K_{PEV}$ value was 29%, indicating that 20% more cells were found outside the normal distribution peak than in the non-variegating Mg4 line. In addition, the FT1.2 line also had noticeably more hCD2-negative cells ($K_{PEV} = 16\%$) than the Mg4 background control. Fluorescence in situ hybridization (FISH) experiments showed that the ΔFT1-LCR hCD2 transgene in the FT1.1 and FT1.2 lines was integrated in the centromere and centromere border, respectively (Figure 7B).

Interestingly, the FT1.7 line gave rise to litters that contained littersmates with two different transgene expression pattern phenotypes within the same litter. One group of littersmates, named FT1.7B, had very few hCD2-negative thymocytes ($K_{PEV} = 28\%$), whereas the other group, FT1.7A ($K_{PEV} = 13\%$), had very few hCD2-negative thymocytes. The frequency of the two phenotypes is roughly equivalent, with half the transgenic littersmates showing the variegating pattern. That littersmates in these two groups do belong to the same line was confirmed by Southern hybridization (not shown). FISH analysis revealed that the transgene in the FT1.7 line is located close to the telomere (Figure 7B). As mammalian telomere-induced PEV has not been described in detail, it is not clear whether the observed variable phenotype is due to the telomeric location of transgene.

The rest of the transgenic lines, FT1.3–6, expressed the transgenic hCD2 on thymocytes in a fairly unimodal manner, resulting in an average $K_{PEV}$ within the 8–13%
range. FISH analysis has shown that the ΔFT1-LCR hCD2 transgene in these lines was integrated in the long arm of the chromosomes (Figure 7B; data not shown). A summary of the degree of variegation of transgene expression for all seven transgenic ΔFT1-LCR hCD2 lines is presented in Figure 8A. We found that the extent of variegation varied particularly within FT1.1, FT1.2 and FT1.7 variegating transgenic lines, resulting in higher standard deviations. For example, among the 18 mice analysed for FT1.1, line variegation ranged from a $K_{\text{PEV}}$ of 13 to 38%, with an average $K_{\text{PEV}}$ of 29%. Such a dispersion within the same line could be due to effects of genetic background on PEV, so that the same transgene variegates differently on the CBA and C57Bl/10 backgrounds (R.Festenstein and D.Kioussis, in preparation).

The bar chart in Figure 8A demonstrates that the FT1.1 line carrying a transgene integrated in the centromere has the highest $K_{\text{PEV}}$. The FT1.2 and FT1.7 mice carried the transgene integrated at the centromere and telomere borders, respectively, and had a relatively high number of...
HBP1 role in hCD2 LCR function

Fig. 8. Summary of the degree of variegation. (A) The histogram summarizes average K_{PEV} values for the FT1.1–7 and Mg4 transgenic lines. Each column represents the average K_{PEV} for a single line calculated as an arithmetic mean of the K_{PEV} of n mice analysed from the same line. The number of mice analysed is indicated under the column for each line (n). Columns 1–6 correspond to the FT1.1–6 lines, whereas columns 7 and 8 correspond to the FT1.7A and FT1.7B mice, respectively. The histogram demonstrates that the FT1.1 and FT1.7B mice have a significantly higher K_{PEV} than the others. The SD is represented by the bars above the respective columns; the representatives of the FT1.3 and FT1.5 lines had virtually identical K_{PEV}s and their SDs are not shown. Each of the FT1.1–7 lines carried 20, 22, 23, 8, 25, 6 and 5 copies of the transgene, respectively, with Mg4 carrying 7 copies (SDs were within a 20% range). No correlation between transgene copy number and degree of variegation was observed. (B) K_{PEV} depends on the distance to the heterochromatin. The graph shows an inverse correlation between K_{PEV} and the distance from the ΔFT1-LCR hCD2 transgene to a heterochromatic region such as the centromere or telomere. Each of the asterisks represents one of the FT1.1–7 transgenic lines. For each line, two mice were analysed and the arithmetic mean of distances measured from the transgene to the centromere or telomere on two daughter chromosomes was calculated. (FT1.7A and FT1.7B mice were analysed as distinct lines and are represented by separate asterisks.) The distance was measured using the NIH Image Software (NIH) and is presented on the graph in arbitrary units. In all lines, the SD was within a 15% range. The broken line correspond to K_{PEV} = 9% of the non-variegating control Mg4 line.

Discussion

To understand the mechanism underlying LCR function, we aimed to identify proteins that interact with the hCD2 LCR. Using DNase I footprint analysis, the binding sites of proteins interacting with the LCR and, potentially, regulating its function were mapped. These footprints were used as baits to screen a T cell-specific library in a one-hybrid assay that resulted in the isolation of the human HBP1. Screening of the one-hybrid and λ-ZAP T cell-specific cDNA libraries did not identify any HBP1 homologues, suggesting that HBP1 is a single gene. In addition, Northern blot analysis has identified only one mRNA band, supporting the hypothesis that the HBP1 protein probably exists in a single form.

Using yeast one-hybrid and various in vitro DNA-binding assays, we have shown that HBP1 interacts with the FT1 region of the LCR in a sequence-specific manner. The identified HBP1-binding motif TTCAATTCAATTCA is different from the canonical (T/A)(T/A)CAA(A/T)G consensus sequence of other HMG box proteins (Landsman and Bustin, 1993; Oosterwegel et al., 1993; Grosschedl et al., 1994) and represents the highest affinity among the HBP1-binding sites identified to date (Tevosian et al., 1997; Yee et al., 1998). Although the sites within the N-myc promoter were mapped originally using the mouse HBP1 HMG domain, it is reasonable to expect that both mouse and human HBP1 HMG domains recognize identical or highly similar sequences, since the human and mouse HMG boxes are identical at the amino acid level. Interestingly, the CCCATTTCT (MYC1), CCCATTTGA (MYC2) and CCCATTTT (MYC3) motifs within the N-myc promoter resemble the CCTATTCA motif within the weak hHBP1-binding site, FT1.W (Figure 2B). The presence of two HBP1 sites (FT1.S and
FT1.W) may account for the larger footprint caused by the nuclear extract as compared with the footprint mapped with the purified HBP1 HMG box. Alternatively, it is possible that the larger region of protection observed with nuclear extracts is due to a larger protein complex comprising HBP1 as the DNA-binding core.

The FT1 footprint is located within the HSS3 region, which does not have any enhancer activity, but is necessary for the prevention of PEV in transgenic mice (Festenstein et al., 1996). Here, we show that the deletion of FT1 results in the variegation of transgene expression when the transgene is integrated close to heterochromatic areas, such as centromeres or telomeres. Our results show that the FT1 region is required for the LCR function and suggest that HBP1 plays a role in the prevention of PEV. Therefore, our data implicate a chromatin-opening activity and suggest the possibility that HBP1 interacts with the chromatin remodelling machinery, such as SWI–SNF and/or deacetylase–acetyltransferase complexes. Involvement of DNA-binding proteins in chromatin remodelling processes has been demonstrated previously. For instance, it has been shown that the DNA-binding GAGA factor is able to act as a PEV modifier in the context of the NURF nucleosome remodelling complex, a member of the SWI–SNF family (Farkas et al., 1994; Tsukiyama et al., 1994, 1995a,b; Granok et al., 1995).

How HBP1 binding to the hCD2 LCR can mediate the prevention of PEV remains to be investigated. However, our results are consistent with other work on HBP1. Thus, it has been shown that HBP1 binds to RB family proteins (Lavender et al., 1997; Tevossian et al., 1997), and the latter are capable of interaction with the BRG1 and BRM components of the SWI–SNF complex as well as with the HDAC1 histone deacetylase (Dunaief et al., 1994; Strober et al., 1996; Brehm et al., 1998; Maghaghi-Jaulin et al., 1998). These data support our hypothesis that HBP1 could interact with the chromatin remodelling complexes by recruiting mediator molecules such as the RB family proteins. It is tempting to speculate that HBP1, which contains the same RB-binding LXCXE motif as HDAC1, is capable of competitively dissociating HDAC1 from chromatin complexes, thus preventing deacetylation of local histones and impeding the establishment of a repressive chromatin state.

It is also possible that HBP1 plays an architectural role in the LCR function, causing DNA bending and/or other structural changes and thus facilitating interaction between proteins that establish and maintain an active chromatin configuration. Such a role was suggested for the LEF1 HMG protein in the regulation of the ADA LCR (Haynes et al., 1996). The mutation of the LEF1-binding site within the ADA enhancer/LCR resulted in loss of site-independent transgene expression. The ability of HMG proteins to recognize altered DNA structures and bend DNA supports the notion that they play a central role in formation and functioning of enhanceosomes. Thus, it has been shown that the assembly and function of the αTCR enhancer complex is dependent on LEF1-induced bending (Giese et al., 1995; Love et al., 1995). Furthermore, it was found that the protein–protein interaction between components of the β-interferon enhanceosome are facilitated by HMG-I(Y) protein altering the enhancer DNA structure (Thanos and Maniatis, 1992; Du et al., 1993; Thanos et al., 1993; Du and Maniatis, 1994; Grosschedl, 1995).

While designing the experiment for assessing the effect of the FT1 deletion on hCD2 expression in transgenic mice, we considered a theoretical possibility of perturbations of enhanceosomes and LCR-binding proteins by the deletion per se. To minimize such a possibility, we deleted 30 bp of nucleotides exactly, corresponding to a whole number (three) of full turns of the DNA double helix, and thus minimizing local perturbations. We have found that in thymocytes of the FT1.1 and FT1.2 variegating mice, the HSS3 site remains hypersensitive (data not shown), indicating that the FT1 deletion per se causes minimal disturbances to the surrounding chromatin and does not prevent the other proteins from binding to the LCR and establishing an open DNase I-hypersensitive chromatin configuration.

Furthermore, the effect of the deletion is not as dramatic as that caused by the deletion of the whole HSS3 region, suggesting that, in addition to HBP1, other HSS3-binding proteins (such as the FT2-binding factors and others) also contribute to the prevention of PEV. Potential involvement of other proteins in the regulation of the LCR function was demonstrated in the experiment with a partial deletion of the HSS3 region which leaves FT1 and FT2 intact, but removes other footprints. Mice carrying such transgenes were subject to PEV (Festenstein et al., 1996), indicating that additional HSS3-binding proteins are also involved in prevention of PEV. The idea that several LCR-binding proteins are responsible for prevention of PEV is consistent with the mass action model of PEV (Locke et al., 1988), which implicates many elements in the establishment of euchromatin or heterochromatin structures (Kioussis and Festenstein, 1997).

The notion that heterochromatin spreads along the chromosomes from the centromere or telomere to the long arms is in line with our results demonstrating the inverse correlation between the degree of PEV and the distance of the transgene from the heterochromatin centres: the closer the transgene is placed to the centromere or telomere, the higher the degree of variegation (Renauld et al., 1997). Factors that determine how far the heterochromatin spreads in each cell are unknown and, according to the mass action model (Locke et al., 1988), this is a stochastic process depending on the local concentration of various chromatin components. This model is consistent with recent findings (Festenstein et al., 1999) demonstrating that the overexpression of centromere–heterochromatin-associated M31 protein in variegating ΔHSS3-LCR hCD2 mice results in an enhancement of variegation. It is worth noting that a similar overexpression of M31 in the ΔFT1-LCR hCD2 variegating mice did not result in enhancement of variegation (data not shown), perhaps indicating that the ΔFT1-LCR transgene is less sensitive to the repressive influence of heterochromatin than ΔHSS3-LCR hCD2. Another factor that may affect transgene silencing in mammals is the number of copies of the transgene (Garrick et al., 1998). However, in our study, we did not see any correlation between transgene copy number and degree of variegation (Figure 8A).

LCRs are powerful elements that are able to regulate expression of genes in a developmental- and tissue-specific manner by opening chromatin, thus rendering the gene
locus transcriptionally active (Kioussis and Festenstein, 1998). We show here that HBP1 contributes to the regulation of LCR function. This, taken together with the known role of HBP1 as a regulator of differentiation (Lesage et al., 1994; Lavender et al., 1997; Tesevian et al., 1997; Gartel et al., 1998; Shih et al., 1998; Yee et al., 1998), suggests a connection between chromatin opening and lineage commitment, indicating that the role of HBP1 in the cell cycle and differentiation may be mediated via chromatin remodelling.

**Materials and methods**

**Isolation of human HBP1 cDNA**
Screening of the Matchmaker one-hybrid and λ-ZAP Jurkat T cell cDNA libraries was performed according to the manufacturer’s instructions (Clontech and Stratagene, respectively). DNA was sequenced using an ABI PRISM Dye-Terminator Cycle Kit (Perkin Elmer) and Automated Sequencer 377 (Applied Biosystems) as per the manufacturer’s instructions. The data were analysed using ABI Sequence Analysis, Factura, AutoAssembler (ABI–Perkin Elmer) and LaserGene (DNA*) softwares.

**Overexpression of recombinant HMG box in E.coli and generation of rabbit antisera**
To overexpress a recombinant HBP1 HMG box, the DNA encoding the HMG box was amplified from isolated hHBP1 cDNA using the primers 5'-GGGGAATTCCATGGAAGTGCAGAAAGCATTGAT-3' and 5'-GGGAGGCTTTTGATTGGAATTGCTGATGTTTCCTTTT-3'. The PCR product was digested with Ndel and HindIII, cloned into a pET-22b plasmid and transformed into E.coli BL21 (DE3) (Novagen). Induction and purification of recombinant were performed following the manufacturer’s recommendations and the protein was stored at −20°C.

The recombinant protein was injected into rabbits and immune antisera were raised. The LEF1 antisera was a generous gift from Dr. R. Grosschedl.

**Preparation of nuclear protein extract. DNAse I footprint and band shift assays**
Nuclear extracts from fresh mouse tissue (thymus and liver) were prepared as described (Dignam et al., 1983). Probes were labelled either at the sense or the antisense strands by T4 kinase with [γ-32P]ATP. For a typical binding reaction, nuclear protein extract (30 μg) or recombinant HMG (50 ng), probe (3000 c.p.m.) and poly(dI–dC) (1 μg) were incubated in 20 μl containing 20 mM HEPES pH 7.9, 1 mM MgCl2, 60 mM KCl, 8% glycerol. After treatment with DNase I (Boehringer Mannheim), the probe was phenol/chloroform purified and analysed on an 8% polyacrylamide–8 M urea sequencing gel. Maxam–Gilbert A and B sequencing reactions were performed using a kit (Biotechnology System NEN RP) according to the manufacturer’s instructions.

For band shift assays, double-stranded oligonucleotides were purified on a non-denaturing polyacrylamide gel and labelled by T4 kinase with [γ-32P]ATP. For a typical binding reaction, nuclear extract (0.2 μg) or recombinant HMG (20 ng) was incubated with probe (20 000 c.p.m., equaling 0.1 ng) in the presence of poly(dI–dC) (20 ng). The mixture was loaded onto a 5% acrylamide gel and electrophoresed in 0.5× TBE at 200 V up to 4 h. In competition experiments, non-labelled competitor DNA was added with the poly(dI–dC).

**Mice**
Mice [CBA/Ca and C57Bl/10 and (CBA/Ca×C57Bl/10)F1] were bred at the National Institute for Medical Research.

**Evaluation of hCD2 expression and calculation of K_{PEV}**
A total of 10⁵ thymocytes were incubated for 30 min at 4°C with CD4 RED (Boehringer Mannheim), CD8 PE (Catlog Laboratories) and fluorescein isothiocyanate (FITC)-conjugated or biotinylated anti-hCD2 (OKT11) (Festenstein et al., 1996) antibodies and subsequently analysed using a Beckton Dickinson FACS sorter and CellQuest programme.

K_{PEV} was calculated as the percentage of thymocytes falling outside of the 1.3× SD interval measured from the mean of hCD2-positive thymocytes. The K_{PEV} for normal distribution is 10%. K_{PEV} for the variegating lines was calculated as the percentage of thymocytes falling outside of the 1.3× SD interval measured from the mean of peak of hCD2-positive thymocytes using the SD of the non-variegating control Mg4 line (SD_{Mg4}). For the lines with a transgene expression level significantly different from that of Mg4, the SD was adjusted according to the formula SD = SD_{Mg4}×P/PMg4 where P and PMg4 are the means of the peaks of hCD2-positive thymocytes in the considered line and Mg4 control, respectively.

**Fluorescence in situ hybridization (FISH)**
Metaphase spreads were obtained from transgenic mouse spleens cultured for 2 days after lipopolysaccharide stimulation (Sigma, final concentration 20 ng/ml). The hCD2 DNA probe was labelled and hybridized with the metaphase spreads following procedures previously described (Festenstein et al., 1996). Slides were then mounted in antifade (Vector) and counterstained in 4’,6-diamidino-2-phenindole (DAPI). They were examined using a Zeiss Axioshot fluorescence microscope and the images collected by cooled CCD camera (Photometrics) using capture software (Digital Scientific).

**Acknowledgements**
T.Z. was supported by the Association for International Cancer Research. B.S. was supported by grant No. PI970203 from the European Community and G.S. was supported by a grant from the Leukaemia Research Fund.

**References**
Gartel,A.L., Gouzman,E., Tsevian,S.G., Shih,H., Yee,A.S. and

4605
T.Zhuma et al.


Received July 9, 1999; revised and accepted September 29, 1999.