3130–3137 *Nucleic Acids Research, 1999, Vol. 27, No. 15* *© 1999 Oxford University Press*

The polyoma virus enhancer cannot substitute for DNase I core hypersensitive sites 2–4 in the human β-globin LCR

Keiji Tanimoto, Qinghui Liu, Jörg Bungert and James Douglas Engel*

Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60208-3500, USA

Received April 1, 1999; Accepted May 27, 1999

ABSTRACT

The polyoma virus enhancer (PyE) is capable of conferring integration position-independent expression to linked genes in stably transfected erythroid cells after joining to DNase I hypersensitive site (HS) 5 of the human β-globin locus control region (LCR). In attempting to separate the chromatin opening activity of the LCR from its enhancer activity and to investigate contributions of the individual HS core elements to LCR function, the human β-globin LCR HS2, HS3 and HS4 core elements were replaced with the PyE within the context of a yeast artificial chromosome (YAC) bearing the whole locus. We show here that, in contrast to its function in cultured cells, the PyE is unable to replace HS core element function *in vivo*. We found that the PyE substitution mutant LCR is unable to provide either chromatin opening or transcriptional potentiating activity at any erythroid developmental stage in transgenic mice. These data provide direct evidence that the human β-globin LCR core elements specify unique functions that cannot be replaced by a ubiquitous enhancer activity.

INTRODUCTION

The human β-globin genes are sequentially organized on chromosome 11, with the ε-globin gene located 5'-most, followed by the two γ-globin genes (Gγ and Aγ), while the adult δ- and β-globin genes are at the 3'-end. This sequential order of the globin genes also reflects the developmental pattern in which these genes are expressed in erythroid cells. Although sequences proximal to the individual globin genes are sufficient to provide stage and tissue specificity to the genes, a locus control region (LCR), located from 6 to 22 kb 5' of the ε-globin gene, is important in initiating and maintaining high levels of globin gene transcription throughout erythroid development (details reviewed in 1). The LCR is composed of an array of DNase I hypersensitive sites (HS1–5), four of which (HS1–4) are erythroid cell specific and developmentally stable. A large number of experiments that examined the function of these HSs in cell transfection assays (transient and stable) and in transgenic mice have shown that HS2, HS3 and HS4 contain information sufficient for directing high level expression of linked reporter genes in erythroid cells. The HS regions have been mapped by functional tests to individual 'core' transcriptional activating activities (2–4) and the importance of these core sequences has been supported by subsequent phylogenetic homology comparisons (5), showing a high degree of conservation in the HS core sequences, but not outside the cores, between humans, mice, goats and rabbits.

Several observations have led to the view that the LCR acts as a 'master' regulator of the β-globin locus genes. Perhaps most compelling among these data was the finding that a naturally occurring deletion of most of the LCR, leaving the globin genes unaltered, results in severe γδβ-thalassemia (diminished expression of the γ-, δ- and β-globin genes) in the afflicted individual (the Hispanic deletion; 6). In this same locus, neither DNase I HSs within the LCR nor general DNase I sensitivity within the locus are detectable after transfer of the mutant chromosome into murine erythroleukemia (MEL) cells, demonstrating that the deleted sequences are important not only to amplify expression of the globin genes but also for maintenance of an open chromatin conformation in the globin locus. This pathophysiological observation was experimentally confirmed by Reik *et al*. (7), who showed that deletion of HS2–5 (sequences lying within the Hispanic deletion) from the human β-globin locus displayed no expression of the human β-globin genes when transferred into MEL cells. However, in contrast to the Hispanic deletion, the remaining DNase I HS and general DNase I sensitivity within the gene targeted locus were unaffected.

Many recent efforts have focused on the introduction of mutations into the LCR within the context of the entire β-globin locus. Several experiments were performed examining the endogenous mouse locus using embryonic stem (ES) cell gene targeting. Fiering *et al*. (8) deleted HS2 from the mouse β-globin locus and found a mild reduction in expression of the adult β^{maj}-globin gene; expression of the embryonic-specific (εy and βH1) genes were unaffected when compared with wild-type animals. Hug *et al*. (9) deleted mouse HS3 using the same strategy and found a similar, quite mild phenotype. Finally, Epner *et al.* (10) deleted the whole LCR (murine HS1–6) in ES cells and analyzed expression from the mutant chromosome after transfer into the human erythroleukemia cell line

*To whom correspondence should be addressed. Tel: +1 847 491 5139; Fax: +1 847 467 2152; Email: d-engel@nwu.edu
Present address:
Jörg Bungert, Department of Biochemistry and Molecular Biology, University of Florida, PO Box 100245, Gainesville, FL 32610-0245, USA

K562. Surprisingly, the mutant locus formed an open chromatin structure and the mouse βh1 globin gene was expressed at a low but significant level. One interpretation of these experiments is that the remaining HSs, or other elements located within and/or possibly even outside the known locus, can partially compensate for the activity of the LCR (10).

In analogous experiments analyzing human β-globin locus transgenes, the phenotype of deleting individual HS elements in the intact locus is more variable, probably because the mutant loci are more affected by neighboring repressive elements in the vicinity of the random transgene integration sites (11). Stamatoyannopoulos and colleagues analyzed deletions of both HS3 (HS3Δ; 12) and HS2 (HS2Δ; 13) in the context of a large human β-globin yeast artificial chromosome (YAC). In the HS3Δ transgenic mice, gene expression in the embryonic yolk sac was most severely affected. In contrast, HS2Δ transgenic mice displayed 30–50% reduction of all β-globin transcripts. Milot *et al*. (11) deleted HS2, HS3 and HS4 in the context of a linked cosmid construct harboring the β-globin locus. They detected variegated β-globin gene expression and concluded that an incomplete LCR (e.g. one bearing single HS deletions) is unable to uniquely render an open chromatin environment upon any insertion site and is therefore not always able to confer position-independent, high level expression to the *cis*-linked locus. We have examined transgenic mice harboring human β-globin YACs bearing deletions of either the HS2, HS3 or HS4 core elements and shown that expression of all of the human β-locus genes, at all developmental stages, was impaired (14,15). Most germane to the present study, substitution of the HS3 core element for that of HS2 within the context of a YAC severely affected transcription of the genes while leaving the hypersensitivity of the LCR generally unperturbed (15), suggesting that the LCR transcriptional and chromatin opening activities might be distinct and therefore separable.

To investigate whether the HS2, HS3 and HS4 core elements could be physically dissected from the LCR without affecting its chromatin structure, we substituted all of the cores with a heterologous enhancer (PyE), cloned from a polyoma virus (16,17) host range mutant that is active in hematopoietic cells (18). The PyE was substituted for the HS2, HS3 and HS4 core elements within a human β-globin YAC and then used to generate transgenic mice. In contrast to all previous studies, we observed a complete loss of LCR HS formation as well as the total loss of expression of all the globin genes, at every developmental stage, from the mutated transgene locus. Thus the present data clearly demonstrate that the core enhancer elements are indispensable for human globin gene expression as well as for LCR chromatin structure formation in transgenic mice and that a ubiquitous enhancer cannot substitute for the unique properties of the core element functions within the LCR.

MATERIALS AND METHODS

Modification of the right YAC vector arm by homologous recombination in yeast

The original 150 kb YAC construct harboring the human β-globin gene locus (which we called A201F4.1; 19,20) was previously modified by disrupting the *URA3* gene and simultaneously inserting a *LYS2* gene by homologous recombination using pRV1 (21), which we referred to as A201F4.2 (14). This modification allowed the use of *URA3* as a marker for subsequent mutagenesis steps in yeast. At the same time, the neomycin-resistance (neo^r) gene under control of the methallothionein-I (MT-I) promoter carried on pRV1 was also incorporated into the right vector arm. In order to use the YAC transgene in ES cells, we replaced the MT-I promoter by the phosphoglycerate kinase (PGK) promoter for elevated neo^r gene expression, since the PGK promoter is very active in ES cells (22). At the same time, we introduced a restriction enzyme site for *Sfi*I into the right YAC vector arm to generate A201F4.3 in order to facilitate structural analysis of the locus after integration into transgenic animals. It has also been reported that some distributed plasmids with the neomycin-resistance (Neo^r) gene have mutations in the neo coding region, thereby diminishing enzyme activity (23). Therefore, we used two different neo^r-containing plasmids, pPGK-neo (mutant Neo^r) and pMC1-neo-poly(A) (wild-type Neo^r, i.e. Glu182; Stratagene), in order to construct the right arm targeting vector, pRV2.

An *Eco*RI [in the multicloning site (MCS) 5' to the PGK promoter]–*Bal*I (within the coding region of the neo^r gene) fragment was excised from plasmid PGK-neo. A *Bal*I (in the coding region of the neo^r gene)–*Hin*dIII (within the MCS 3' to the gene) fragment was excised from pMC1-neo-poly(A). These two fragments were inserted into the *Eco*RI and *Hin*dIII sites of pRS306. *Eco*RI and *Bam*HI (at the 3'-end of the gene) sites were then sequentially removed from the resultant plasmid by digestion with the appropriate enzymes, and the plasmid blunt-ended and self-ligated (pRS306-Neo-ΔEB). The MT-I promoter–URA gene chimeric fragment was prepared by PCR using the following primers and pRV-1 as template: *URA* primer, 5'-ATG**TCTAGA**TCCTGTAGAGACCACATCATC-3'; MT-I primer, 5'-GAT**GGATCC***GGCCTTCGAGGCC*AGCCTGGTCAACAAAGTG-3'. Homologous sequences to the targeted sequences are underlined. We also included *Xba*I and *Bam*HI sites (bold) within the *URA* and MT-I primers, respectively, to facilitate subsequent cloning steps. An *Sfi*I site (italic) was also included within the MT-I primer. The PCR product was digested with *Xba*I and *Bam*HI and then cloned into the *Xba*I and *Bam*HI sites of pRS306-Neo-ΔEB to generate pRV-2. A unique *Eco*RI site between the *URA* and the MT-I promoter sequences was used to linearize the targeting vector for homologous recombination in yeast.

YAC mutagenesis and DNA isolation

The construction of substitution targeting vectors for HS core sequences (HS2, HS3 and HS4) has been described previously (14,15). Basically, these pRS306-based vectors contain the flanking sequences of individual HSs and substituted core elements inserted as *Xba*I–*Xho*I fragments. The PyE fragment was prepared by PCR using the following oligonucleotides to amplify the enhancer from plasmid pMC1-neo-poly(A) (Stratagene): PyE-S, 5'-CAGGGCCAGTGAATTCTCG(ag)-3'; PyE-A, 5'-GCGGGGTTTTCTAGACATTGGG-3'. The positions of endogenous *Xho*I (nucleotides in parentheses are not included within the synthetic oligonucleotide but are present on the template) and artificial *Xba*I restriction enzyme sites are underlined in the sequences. The PyE fragment was digested with *Xho*I and *Xba*I and ligated to the *Xho*I+*Xba*I-digested targeting vectors in order to clone this fragment in place of individual HS core elements. The core PyE fragments were sequenced in

all three targeting constructs (for HS2, HS3 and HS4) to verify that no mutations were generated during PCR. The resultant target vectors were linearized using unique restriction enzyme recognition sites in the HS flanking regions and used for homologous recombination in yeast. Most of the yeast methods used here were derived from a single source (24). Yeast integrative plasmid pRS306 contains the *URA3* selectable marker used to monitor homologous integration into and excision from the YAC construct, as previously detailed (14). The three independent mutations (for individual HS core elements) were consecutively introduced into the YAC A201F4.3. YAC DNA for microinjection was isolated from pulsed field gels according to standard protocols (25) with minor modifications (14).

Transgenic mice

Purified YAC DNA was injected into fertilized mouse oocytes (CD1; Charles River Breeding Laboratory) before transfer to foster mothers (CD1) as described (26). Tail DNA from offspring was initially analyzed for the presence of the left and right YAC vector arms by PCR and then for copy number and integrity of the transgenic β-globin locus by Southern blots of DNA prepared from transgenic tails and thymuses.

Preparation of agarose plugs for pulsed field gel analysis

Samples of 2×10^7 cells from the thymus (from <2-month-old animals) were suspended in 0.8 ml of cell suspension buffer (CSB) (10 mM Tris–HCl, pH 7.4, 20 mM NaCl, 50 mM EDTA, pH 8.0). This solution was then mixed with 2% low melting point agarose solution in CSB and poured into clamped homogeneous electric field (CHEF) plug molds. Solidified plugs were treated at 50°C in proteinase K buffer (100 mM EDTA, pH 8.0, 0.2% sodium deoxycholate, 1% Sarcosyl, 10 mg/ml proteinase K) for 48 h, followed by four washes with SE (75 mM NaCl, 25 mM EDTA) buffer for 1 h and two washes with TE (10 mM Tris–HCl, pH 7.4, 1 mM EDTA) buffer for 1 h.

Reverse transcriptase–polymerase chain reaction (RT–PCR) analysis

Total RNA from at least two animals from each of the transgenic lines was extracted from yolk sac and anemic adult spleen using ISOGEN (Nippon Gene) based on the AGPC method (27). First strand cDNA was synthesized with Moloney murine leukemia virus (M-MuLV) reverse transcriptase (Gibco BRL) using 2.5 μg of total RNA in a 20 μl reaction volume. The PCR was performed with 0.4 μl of cDNA in a 10 μl reaction volume (200 μM each dNTP, 40 ng of each primer set, 0.2 U of AmpliTaq, 0.5 μCi of [α-^{32}P]dCTP, 1× Perkin-Elmer PCR buffer (2.5 mM $MgCl_2$) under the following conditions: 94°C for 30 s, 58°C for 1 min and 72°C for 1 min. An aliquot of each PCR reaction was electrophoresed on 8% polyacrylamide gels, dried and subjected to autoradiography. The PCR primers used were: hε, 5'-CAATCACTAGCAAGCTCTCAGG-3' (sense) and 5'-GGGCTTGAGGTTGTCCATGTTT-3' (antisense); hγ, 5'-GATGCCATAAAGCACCTGGATG-3' (sense) and 5'-TTGCAGAATAAAGCCTATCCTTGA-3' (antisense); hβ, 5'-AACTGTGTTCACTAGCAACCTCAA-3' (sense) and 5'-GAGTGGACAGATCCCCAAAGGA-3' (antisense); mα, 5'-CTGATTCTGACAGACTCAGGAAGAA-3' (sense) and 5'-CCTTTCCAGGGCTTCAGCTCCATAT-3' (antisense).

All the primer sets flank introns so that the amplified products corresponding to cDNA could be easily distinguished from those amplified from contaminating genomic DNA.

DNase I HS mapping

DNase I HS mapping was carried out essentially according to the procedure of Forrester *et al*. (28). Adult transgenic mice were made acutely anemic by phenylhydrazine injection as described (14). Single cell suspensions were then rinsed once with PBS(–) and resuspended in 3 ml of RSB buffer (10 mM Tris–HCl, pH 7.4, 10 mM NaCl, 5 mM $MgCl_2$) at 4°C. The cells were lysed by adding 3 ml of RSB plus 0.5% NP-40 and douncing 15 times with a type B pestle homogenizer. Nuclei were harvested by centrifugation at low speed in a clinical centrifuge at 4°C and then rinsed once with 10 ml of RSB at 4°C. Nuclei were then aliquoted into several fractions (300 μl each), containing approximately 10^6 nuclei, and then treated with DNase I (either 0, 1, 5 or 20 mg/ml final concentration, respectively) at 37°C for 15 min.

After quenching the reactions by the addition of an equal volume (300 μl) of a solution containing 1% SDS, 20 mM Tris–HCl, pH 7.4, 0.6 M NaCl, 10 mM EDTA and 100 μg/ml proteinase K, the nuclei were digested with proteinase K at 55°C overnight, after which the DNA was purified by extraction with phenol/$CHCl_3$ and ethanol precipitated. Following resuspension, the DNA was digested with *Eco*RI or *Pvu*II, electrophoresed on agarose gels and then transferred to nylon membranes followed by hybridization to radiolabeled DNA probes. For mapping DNase I HS within the LCR, a probe corresponding to the 3'-flanking region of HS4 [*Xba*I (nt 2352)–*Hin*dIII (nt 3266); www.psuglobin.edu) was used. For mapping of HS1 and HS5, the *Hin*dIII (nt 13769)–*Xba*I (nt 14235) fragment and *Eco*RI–*Bgl*II (~0.5 kb) fragment 5' to HS5 were used, respectively. The probe used to map the mouse β^{maj}-globin gene promoter HS was a 498 bp *Pst*I–*Sac*I fragment from the 5'-region of the gene (a generous gift of Dr Mark Groudine).

RESULTS

Mutagenesis of the human β-globin LCR by homologous recombination in yeast

A mutant polyoma virus enhancer (which we refer to hereafter as PyE) was identified from one of the host range polyoma virus mutants (PyF101) that is able to grow not only in differentiated mouse cells but also in embryonic cells (29). This PyE is a ubiquitous enhancer that is active in a broad range of cell types and is a stronger transcriptional activator than the wild-type enhancer. We replaced three of the LCR HS core sequences (HS2, HS3 and HS4, ranging from 200 to 400 bp in size) within human β-globin YAC A201F4.3 with the PyE enhancer (~150 bp) using homologous recombination in yeast (Fig. 1A). To verify that the correct mutations were introduced, the wild-type and mutant YAC DNAs were purified from yeast cells, digested with *Bam*HI and *Hin*dIII (as indicated in Fig. 1B), electrophoresed, blotted to nylon membranes and hybridized with probes specific for the HS core sequences or PyE sequences. None of the probes for individual HSs hybridized to fragments from the mutant YAC, although the same probes hybridized to the anticipated bands from wild-type YAC

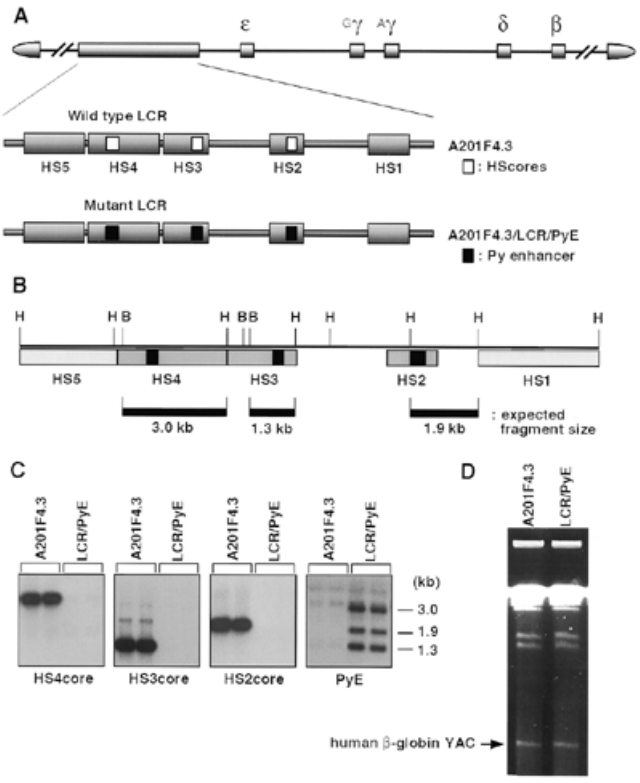

Figure 1. Mutagenesis of the human β-globin YAC by homologous recombination in yeast. (**A**) Schematic representation of the human β-globin YAC and LCR. The top line shows the diagrammatic representation of the human β-globin YAC [called A201F4.3, originally derived from A201F4 (19), 150 kb]. The position of the β-like globin genes are shown relative to the LCR. A detailed structure of the LCR from wild-type and mutant YACs is shown with the location of individual HSs (HS1–5) indicated in the bottom two lines. HS core sequences in HS2, HS3 and HS4, delineated by open boxes, were replaced with PyE as described in the text (solid boxes). (**B**) Restriction enzyme map and probes for the wild-type and mutant LCRs. The position of *Hin*dIII (H) and *Bam*HI (B) sites are shown as vertical lines. Fragments of the expected size can be detected with individual HS core probes (for the wild-type LCR) or PyE probe (mutant LCR) as indicated. (**C**) Southern blot analysis of YAC DNA from yeast. DNA from yeast clones bearing the wild-type (A201F4.3) or LCR mutant human β-globin YACs (LCR/PyE) was digested with *Bam*HI and *Hin*dIII and separated on regular agarose gels. After transfer to nylon membrane, the DNA was hybridized to HS core probes (HS4, HS3 and HS2 cores) or PyE probe. (**D**) PFGE of yeast DNA. Human β-globin YAC DNA was separated from yeast chromosomes by PFGE and visualized by ethidium bromide staining of the gel.

DNA, indicating that the HS core sequences were indeed deleted from the mutant. When the same blot was rehybridized with a PyE fragment, the expected fragments were detected only in the mutant YAC DNA, confirming that the mutations were successfully incorporated. Ethidium bromide staining of the pulsed field gel (Fig. 1C) showed no difference in size between the wild-type and mutant YACs, demonstrating that no gross rearrangement had occurred during the successive mutagenesis steps in yeast.

Structural analysis of wild-type and mutant human β-globin YAC transgenic mice

Wild-type and mutant YAC DNAs were purified from pulsed field gels and injected into fertilized oocytes of CD1 mice to generate transgenic animals. Tail DNA was prepared from the offspring and screened by PCR using left and right YAC vector arm-specific primer sets and positives were then further analyzed by Southern blot hybridization.

In order to determine the transgene copy numbers in each of the lines, we first probed *Pst*I-digested mouse tail DNA with fragments corresponding to the left (L-end) and right (R-end) vector arms of the YAC. These fragments are located outside the *Pst*I restriction enzyme sites within the YAC vector arms (as shown in Fig. 2A), thus enabling analysis of the ends of the transgene at the integration sites. Using this strategy, we can also determine the arrangement of multiply integrated copies of the transgene (e.g. whether they integrate in a head-to-head, head-to-tail or tail-to-tail configuration) (30).

We generated four independent transgenic lines with the core enhancers of HS2, HS3 and HS4 substituted by the PyE (nos 304, 342, 348 and 355). Figure 2B shows that all four lines appear to bear a single copy of the mutant human β-globin YAC, as revealed by the appearance of only single bands using both the L-end and R-end probes (Fig. 2B). Further internal fragment analysis confirmed this assumption (data not shown).

One potential problem with YAC transgenes is the possibility that a broken piece(s) of DNA might integrate within or near an intact transgene, which could therefore alter expression from the locus. It is therefore crucial to carefully examine the integrity of these transgenes. To facilitate the analysis of YAC integrity in transgenic mice, we introduced a restriction enzyme site for *Sfi*I into the right vector arm of the 155 kb human β-globin YAC (A201F4.3) by homologous recombination in yeast (Fig. 2A). In this modified YAC, the whole globin gene locus is included within two *Sfi*I restriction enzyme fragments, one of which (10 kb) contains HS5 and HS4 and the other (100 kb) spans a region from within the LCR (5' to HS3), including all the β-like globin genes (ε, γ and β), into the right vector arm. By detecting these two fragments on pulsed field gels with several probes spanning the locus, we concluded that all four lines carry intact, single copy globin transgene loci and that no other globin locus fragments were integrated. High molecular weight DNA was prepared from the thymus of transgenic animals, embedded in agarose plugs and digested with *Sfi*I. After gel electrophoresis the DNA was transferred to nylon membranes and hybridized to various probes from the human β-globin locus and YAC vector arm: HS5, HS4-3', ε globin gene, β globin gene and the right arm (inside the *Sfi*I site; Fig. 2A and C). All of the probes detected single fragments of the expected size in all four transgenic lines, indicating that these lines carry complete, single copy transgenes.

Expression in PyE/LCR-substituted human β-globin YAC transgene mice

In order to examine the effect of the LCR mutations on expression of the human globin genes during murine erythroid development, we collected yolk sacs from 9.5 day embryos and anemic spleens (from 1-month-old animals) and analyzed β-like globin gene expression by semi-quantitative RT–PCR (14). Figure 4 shows that none of the human β-globin genes are expressed in the mutant lines. We also analyzed two multicopy lines bearing the mutant YAC and again failed to detect expression of the human β-globin genes (data not shown). In contrast, similar samples from transgenic mice carrying the

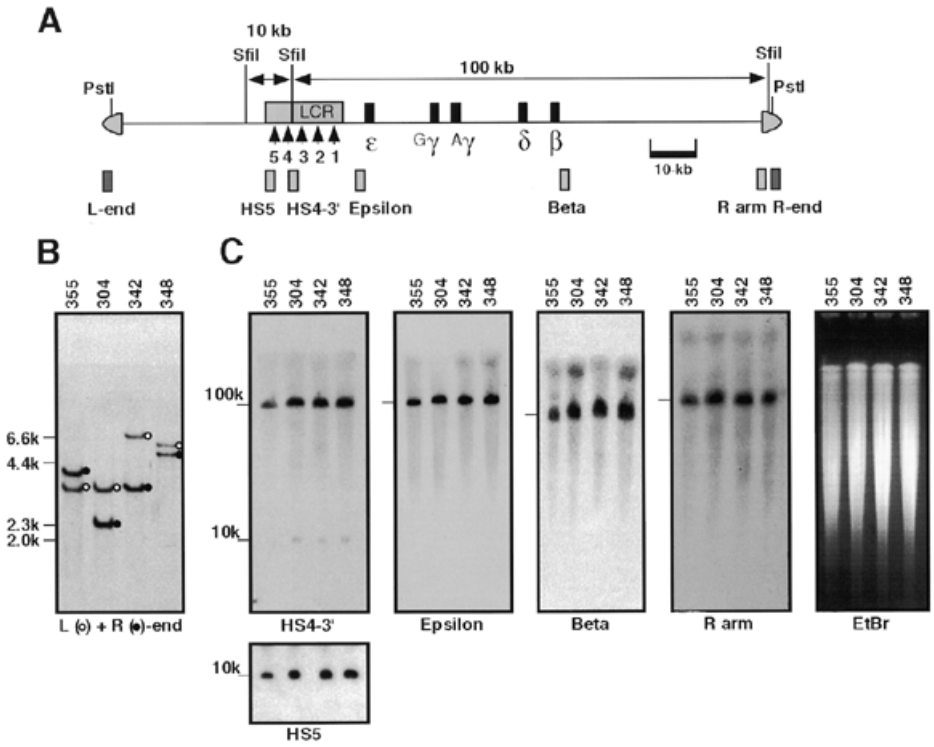

Figure 2. Structural analysis of the β-globin YAC in transgenic mice. (**A**) Schematic representation of the human β-globin YAC A201F4.3. *Sfi*I restriction enzyme sites are located 5' to HS5, between HS4 and HS3 and in the right arm of the YAC. The position of *Pst*I sites, present in the left and right vector arms, is also shown. Probes used for short- and long-range fragment analysis are indicated below the map. (**B**) Short-range analysis of integrated transgenes. Tail DNA from F_1 transgenic progeny was digested with *Pst*I, separated on agarose gels and transferred to nylon membranes. Hybridization was first performed with the L-end (L, open circle) and R-end (R, filled circle) probes, separately (data not shown) and then together. The bands detected with the probes for the left and right arms are indicated. (**C**) Long-range analysis of the integrated transgenes. DNA was isolated from transgenic thymus cells and digested with *Sfi*I in agarose plugs, separated on pulse field gels and transferred to nylon membranes. Hybridization was carried out using several probes spread throughout the β-globin locus.

wild-type human β-globin locus showed, as expected, high levels of ε- and γ-globin transcript accumulation in the embryonic yolk sac and abundant expression of the adult β-globin gene in definitive erythroid cells isolated from the spleen. This result was remarkable in light of the fact that even in the absence of a linked LCR, low level transcription from individual genes could be detected in numerous previous transgenic experiments (31–34). However, it should also be noted that many of the previous studies were performed using plasmid globin gene constructs, which often integrate in significantly higher copy numbers.

DNase I HS mapping of mutant and wild-type β-globin LCRs

One of the established functions of the human β-globin LCR is to provide an open chromatin environment throughout the locus (35). In order to examine the effect of substituting the LCR core enhancer with the PyE on HS formation as one reflection of the state of chromatin in this locus, we performed DNase I HS mapping experiments on erythroid cells from one wild-type (no. 31) and three mutant (nos 304, 342 and 348) human β-globin YAC transgenic lines. We first tested whether or not the substitution of HS with the PyE disrupted formation of HS in the erythroid chromatin of the YAC transgenic mice. Figure 4A depicts the structure of the wild-type and mutant LCRs and the predicted results if various HSs are detected within it. Figure 4B shows that none of the transgenic lines harboring the mutant LCR displayed HS formation at the anticipated positions. In contrast, strong HSs were easily detected in the LCR of wild-type YAC transgenic mice. As an internal control, the same samples were probed with a fragment derived from the (endogenous) mouse β^{maj}-globin gene promoter. Figure 4B shows that the endogenous HS is efficiently formed in erythroid chromatin in all of the transgenic animals.

Reik *et al.* (7) deleted an LCR fragment from HS2 to HS4 from human chromosome 11 and then analyzed chromatin structure and expression of the globin genes after transfer of the mutated chromosome into MEL cells. These authors found that there was no expression of the human β-like globin genes from the mutated globin locus but that the mutant locus retained efficient formation of the remaining HSs (HS5 and HS1). From these data, together with mapping the general sensitivity within the locus, the authors concluded that the LCR is required for activating the β-like globin genes, but was not required for the maintenance of an open chromatin configuration. We therefore analyzed the formation of HS5 and HS1 in the PyE-substituted LCR mutant transgenic mice. Figure 5A shows the experimental strategy to detect the formation of HS5 and HS1 using the probes EB and HS1-3', respectively. DNase I-treated DNA was digested with *Pvu*II, blotted to nylon membrane and hybridized with probes specific for human HS. As shown in Figure 5B, HS1 and HS5, as well as

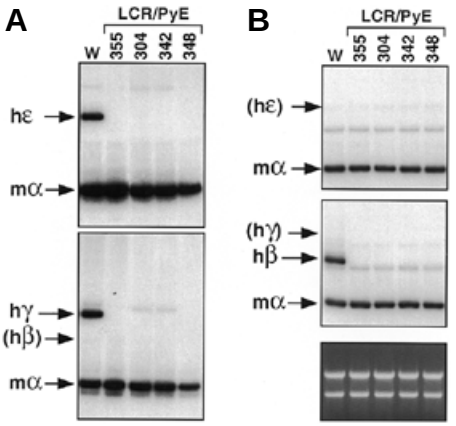

Figure 3. RT–PCR analysis of β-globin gene expression in transgenic mice. (**A**) Embryonic yolk sac expression of the transgenes. Total RNA was prepared from yolk sacs of 9.5 d.p.c. embryos. cDNA was synthesized beginning with 2.5 μg of total RNA and subjected to RT–PCR analysis. (Upper) An autoradiogram of 18 cycles of a radioactive PCR reaction with mixed primer sets (E mix) specific for mouse α- and human ε-globin genes. (Lower) Twelve PCR cycle reactions with mixed primer sets (GB mix) for mouse α-, human γ- and human β-globin genes. (**B**) Adult spleen expression of transgenes. One- to two-month-old transgenic mice were made anemic by injecting phenylhydrazine 6 days before surgery. Total RNA was prepared from anemic spleens and 2.5 μg was used for synthesizing cDNA. (Upper) The PCR result with E mix at 12 cycles. (Lower) The result with GB mix at 12 cycles. Ethidium bromide staining of the total RNA samples separated on agarose gel is shown at the bottom to illustrate the quality of the RNA.

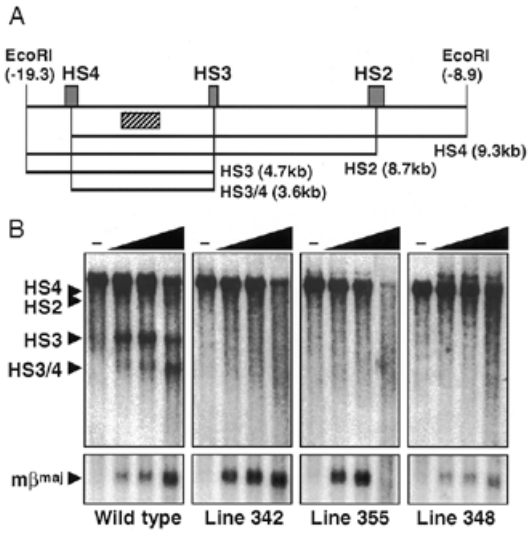

Figure 4. DNase I HS mapping of the human LCR chromatin in wild-type and PyE mutant YAC transgenic mice. (**A**) Schematic representation of the 10.4 kb *Eco*RI fragment of the LCR. The locations of the HS core elements HS2, HS3 and HS4 (located between –19.3 and –8.9 kb relative to the ε-globin gene promoter) and the probe used for detecting the fragments (HS4-3') are indicated as shaded or hatched boxes, respectively. Fragments generated by digestion with *Eco*RI and DNase I are shown with their predicted sizes. (**B**) Nuclei were isolated from anemic spleen, aliquoted and treated with different concentrations of DNase I [shown at the top of each panel as either no treatment (–) or increasing amounts of DNase I]; the reactions were stopped, DNA was isolated and then digested with *Eco*RI (Materials and Methods). After agarose gel electrophoresis, the DNA was transferred to nylon membranes and hybridized to either HS4-3' (Upper) or mouse β^{maj} probes (Lower). HSs formed in the transgenic wild-type human β-globin LCR or in the endogenous mouse β^{maj}-globin gene promoter region are indicated by arrowheads. Samples from left to right: transgenic lines carrying the wild-type human β-globin locus YAC (wild-type) and three of the LCR mutant lines (342, 355 and 348).

HS4 (HS3 is not separated from the parental band under these conditions), were clearly formed in wild-type, but not in mutant, YAC transgenic mice. Although HS1 might be formed weakly in one of the mutant transgenic lines (no. 348), the intensity of the band is significantly lower than that of the corresponding fragment found in the wild-type locus. The appearance of a weak HS1 band in only one of the mutant lines might thus be due to a position of integration effect. The results taken together clearly show that substitution of the core enhancer of HS2, HS3 and HS4 by the PyE results in loss of HS formation in LCR.

DISCUSSION

It is well documented that undifferentiated mouse embryonal carcinoma (EC) cells are resistant to productive infection by wild-type polyoma virus (36). Several groups have characterized polyoma virus mutants that were capable of productively infecting EC cells and in most cases these mutations mapped to the polyoma virus enhancer (reviewed in 37). These host range mutant enhancers have been used for expressing transgenes, especially in undifferentiated cells, since the mutants are expressed in a wide range of tissues, including adult spleen and bone marrow (38). Fujimura *et al*. (29) identified three mutants (PyF101, PyF111 and PyF441) which have single nucleotide changes within the enhancer. One of these enhancers, PyF101, has been shown to be active in the early embryo, starting from the blastocyst stage, and has therefore been used for expressing selectable markers in ES cells (39). Further analysis showed that the PyF101 enhancer, when inserted into the long terminal repeat (LTR) of M-MuLV, is functional in hemopoietic progenitor cells as well as in EC cells (18). Based on these observations, we chose to use the mutated (F101) enhancer (PyE) in the present experiments, because the mutation generated an enhancer which is active during hematopoiesis and very early embryogenesis.

In order to analyze the overall effect of substituting the human β-globin LCR HS core elements with heterologous enhancers on the expression of the β-like globin genes, we mutagenized the human β-globin YAC by homologous recombination in yeast (Fig. 1). Using this strategy, we replaced the core enhancer elements of HS2, HS3 and HS4 with the PyE and generated four lines of mutant human β-globin YAC transgenic mice. One of the most important but technically challenging aspects of using large DNAs for generating transgenic animals is to prove the integrity of the transgene after it has integrated into mouse chromatin. In order to facilitate this process, we decided to introduce a new *Sfi*I restriction enzyme site into the right arm of the human β-globin YAC (A201F4.3; Fig. 2A). Using the modified YAC to generate wild-type and mutant human β-globin transgenic mice, we could clearly show that all of the lines analyzed in this study contain at least one intact copy of the YAC transgene (Fig. 2C). By combining the pulsed field gel electrophoresis (PFGE) data with the

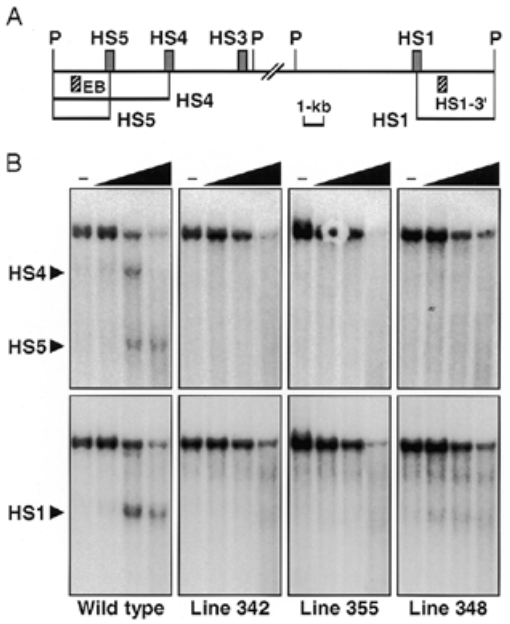

Figure 5. DNase I HS mapping of HS5 and HS1. (**A**) Schematic representation of the 5' and 3' portions of the LCR. The location of the HS elements HS5, HS4, HS3 and HS1 and the probes used for detecting the fragments (EB and HS1-3') are indicated as shaded or hatched boxes, respectively. Fragments generated by digestion with *Pvu*II (P) and DNase I are shown as horizontal thick lines. (**B**) The same DNase I-treated samples as used in Figure 4 were digested with *Pvu*II. After agarose gel electrophoresis, the DNA was transferred to nylon membranes and hybridized to either EB (Upper) or HS1-3' (Lower) probes. HS formed in the transgenic human β-globin LCR are indicated by arrowheads. Samples from left to right: transgenic lines carrying the wild-type human β-globin locus YAC (wild-type) and three mutant lines (342, 355 and 348).

results of end fragment analysis (qualitative copy number analysis; Fig. 2B) and internal fragment assay (comparing the intensity of internal restriction enzyme fragments from the YAC with fragments from the endogenous genome in a quantitative copy number analysis; not shown), we concluded that all four transgenic lines analyzed here carry single copies of the globin gene locus.

We analyzed the developmental expression pattern of the 150 kb human β-globin locus in four independent transgenic lines after substitution of all three HS core sequences with the 145 bp PyE. Surprisingly, none of the transgenic lines expressed β-like globin genes at any developmental stage: in the embryonic yolk sac (Fig. 3A), fetal liver (14.5 d.p.c.; data not shown) or adult spleen (Fig. 3B). We also analyzed two transgenic lines carrying two copies of the mutant globin locus and found that expression of the human globin genes was still undetectable. Even after vast over-amplification of the RT–PCR reactions (30 cycles, whereas we normally quantify adult β-globin expression after 12 cycles), we failed to detect expression of any of the human β-globin genes (data not shown).

It is well known that transgenes can suffer from position effects at the site of integration in higher eukaryotic cells, leading to variegated gene expression. Milot *et al*. found that, although the wild-type LCR has the ability to overcome position effects (referred to as chromatin opening activity; 35), β-globin transgenes suffer from silencing effects from surrounding chromatin once the LCR is mutated (11). Even under these circumstances, however, low but significant levels of β-globin gene expression in all of the mutant transgenic lines could be detected. The phenotype of the mutants analyzed in the experiment described here is apparently more severe than what we (14,15) and Milot *et al*. reported for single LCR HS mutants.

In order to examine whether the substitution of the core enhancer elements by PyE would lead to a change in hypersensitivity, we analyzed HS formation within the LCR in mutant as well as wild-type transgenic mice (Fig. 4). DNase I HS mapping clearly showed that the PyE does not form HSs when situated within the LCR. It has been shown that the polyoma virus enhancer region is hypersensitive to DNase I in cultured cell systems. Thompson *et al*. (40) mapped the DNase I HSs within the LTR of integrated M-MuLV proviral DNA and found two strong HSs in the 5'-LTR; one of these sites is localized in the enhancer region. Subsequently, the authors inserted an additional PyE (F101) or replaced the enhancer region of M-MuLV with PyE and analyzed the chromatin configuration. In both constructs, the authors reported strong HS formation within the PyE fragment and the hypersensitivity of the inserted PyE was considerably stronger than that of the M-MuLV HSs. It is clear from the present results that the PyE fragments integrated into the human β-globin LCR architecture do not have the ability to form HSs in erythroid cells of transgenic mice, suggesting that the context of the substitutions might actually repress the ability of the PyE to alter surrounding chromatin.

We next asked whether or not the remaining HSs (HS5 and HS1) are formed when HS2, HS3 and HS4 are substituted by PyE in transgenic mice. Reik *et al*. (7) showed that HS5 and HS1 are formed efficiently in the absence of LCR elements HS2–4 in MEL cells. We show here, however, that HS5 and HS1 are not formed in the PyE substitution mutant YAC transgenic mice (Fig. 5B). This result does not necessarily mean that the whole globin locus in the mutant transgenes is within an inactive chromatin configuration, but it does indicate that the mutated LCR is not actively involved in the chromatin opening activity. One of the possible explanations for the discrepancy between our results and those of Reik *et al*. (7) could be attributable to the fact that deletion of a large portion of the LCR in tissue cultured erythroid cells cannot address the role of the LCR in initiating chromatin opening in earlier stages of hematopoiesis or erythropoiesis. In other words, the LCR may be indispensable for initiating chromatin opening, but may not be required for maintaining an open chromatin structure. Another possibility is that the PyE is actively repressing the chromatin opening activity of the LCR, although such an activity has not been reported for the PyE.

In our recently described model for LCR function (15), we suggested that individual HS core elements as well as their flanking sequences interact with each other to generate a holocomplex, which is required for providing both position-independent and high level expression to the globin genes in transgenic mice. The absence of HS formation in the PyE-substituted mutant LCR could indicate a failure to reconstitute the holocomplex structure. Since we speculated that both chromatin opening and transcriptional enhancement are functions of the LCR holocomplex, complete loss of globin gene

expression by the PyE substitution lines is probably explained by the lack of both of these activities in mutant LCR.

Several studies (41,42) have indicated that LCR element HS5 in the human, or HS4 in the chicken, β-globin locus harbors an insulator function and suggest that these elements might be able to shelter the globin locus from neighboring regulatory elements in transgenic animals. Yu *et al*. (43) sequenced the HS5 region from the human locus and identified a potential matrix attachment site (MAR). In the same report, the authors showed that the globin MAR by itself has little effect on transcription of a linked reporter gene in stably transfected cells. However, the combination of the globin MAR sequence with the PyE in the same experiment leads to copy number-dependent and integration site-independent expression of the reporter gene (in non-erythroid HT1080 cells). These authors concluded that the globin MAR and PyE might work synergistically to counteract chromosomal position effects and to ensure copy number-dependent expression of the transgene. Although there are three PyEs and an intact HS5 region in the mutant YAC construct analyzed in this study, we detected no globin gene expression. The discrepancy between these two studies might be due to differences in the plasmid constructs versus YACs containing the whole globin gene locus or it might reflect differences in the activity of regulatory elements in tissue culture versus transgenic animals.

In summary, we disrupted the regulation of β-globin gene expression in transgenic mice by substituting the HS core sequences of the LCR with PyE enhancer fragments. It appears that substitution of the core enhancer elements by PyE creates a non-functional LCR, which is unable to confer HS formation and expression to the globin genes in transgenic mice. Our results clearly demonstrate that the LCR HS core elements embody unique activities required for the formation of a functional LCR which cannot be replaced by ubiquitous enhancer activities *in vivo*.

ACKNOWLEDGEMENTS

We are very grateful to Jie Fan, Weimin Song and Yoko Tanimoto for outstanding technical assistance and to all the members of our laboratory for insightful discussions. Fellowship support was provided by the Japan Society for the Promotion of Science (K.T.), the Naito Foundation (K.T.), the Cooley's Anemia Foundation (K.T. and Q.L.) and the American Heart Association (J.B.). Research support for this work was provided by the NIH (DK 52356 to J.B. and HL 24415 to J.D.E.).

REFERENCES

1. Stamatoyannopoulos,G. and Neinhuis,A.W. (1994) In Stamatoyannopoulos,G., Nienhuis,A.W., Majerus,P. and Varmus,H. (eds), *The Molecular Basis of Blood Disease*, 2nd Ed. W.B.Saunders Co., New York, pp. 107–155.
2. Pruzina,S., Hanscombe,O., Whyatt,D., Grosveld,F. and Philipsen,S. (1991) *Nucleic Acids Res.*, **19**, 1413–1419.
3. Philipsen,S., Talbot,D., Fraser,P. and Grosveld,F. (1990) *EMBO J.*, **9**, 2159–2167.
4. Talbot,D., Philipsen,S., Fraser,P. and Grosveld,F. (1990) *EMBO J.*, **9**, 2169–2178.
5. Hardison,R., Slightom,J.L., Gumucio,D.L., Goodman,M., Stojanovic,N. and Miller,W. (1997) *Gene*, **205**, 73–94.
6. Driscoll,M.C., Dobkin,C.S. and Alter,B.P. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 7470–7474.
7. Reik,A., Telling,A., Zitnik,G., Cimbora,D., Epner,E. and Groudine,M. (1998) *Mol. Cell. Biol.*, **18**, 5992–6000.
8. Fiering,S.,E., Epner,E., Robinson,K., Zhuang,Y., Telling,A., Hu,M., Martin,D.I.K., Enver,T., Ley,T.J. and Groudine,M. (1995) *Genes Dev.*, **9**, 2203–2213.
9. Hug,B.A., Wesselschmidt,R.L., Fiering,S., Bender,M.A., Epner,E., Groudine,M. and Ley,T.J. (1996) *Mol. Cell. Biol.*, **16**, 2906–2912.
10. Epner,E., Reik,A., Cimbora,D., Telling,A., Bender,M.A., Fiering,S., Enver,T., Martin,D.I., Kennedy,M., Keller,G. and Groudine,M. (1998) *Mol. Cell*, **2**, 447–455.
11. Milot,E., Strouboulis,J., Trimborn,T., Wijgerde,M., deBoer,E., An Langeveld,A., Tan-Un,K., Vergeer,W., Yannoutsos,N., Grosveld,F. and Fraser,P. (1996) *Cell*, **87**, 105–114.
12. Navas,P.A., Peterson,K.R., Li,Q., Skarpidi,E., Rohde,A., Shaw,S.E., Clegg,C.H., Asano,H. and Stamatoyannopoulos,G. (1998) *Mol. Cell. Biol.*, **18**, 4188–4196.
13. Peterson,K.R., Clegg,C.H., Navas,P.A., Norton,E.J., Kimbrough,T.G. and Stamatoyannopoulos,G. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 6605–6609.
14. Bungert,J., Dave,U., Lim,K.C., Lieuw,K.H., Shavit,J.A., Liu,Q. and Engel,J.D. (1995) *Genes Dev.*, **9**, 3083–3096.
15. Bungert,J., Tanimoto,K., Patel,S., Liu,Q., Fear,M. and Engel,J.D. (1999) *Mol. Cell. Biol.*, **19**, 3062–3072.
16. Tyndall,C., La Mantia,G., Thacker,C.M., Favaloro,J. and Kamen,R. (1981) *Nucleic Acids Res.*, **9**, 6231–6250.
17. de Villiers,J. and Schaffner,W. (1981) *Nucleic Acids Res.*, **9**, 6251–6264.
18. Valerio,D., Einerhand,M.P., Wamsley,P.M., Bakx,T.A., Li,C.L. and Verma,I.M. (1989) *Gene*, **84**, 419–427.
19. Gaensler,K.M.L., Burmeister,M., Brownstein,B.H., Taillon-Miller,P. and Myers,R.M. (1991) *Genomics*, **10**, 976–984.
20. Gaensler,K.M., Kitamura,M. and Kan,Y.W. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 11381–11385.
21. Srivastava,A.K. and Schlessinger,D. (1991) *Gene*, **103**, 53–59.
22. Soriano,P., Montgomery,C., Geske,R. and Bradley,A. (1991) *Cell*, **64**, 693–702.
23. Yenofsky,R.L., Fine,M. and Pellow,J.W. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 3435–3439.
24. Guthrie,C. and Fink,G.R. (1991) *Methods in Enzymology*, vol. 194.
25. Schedl,A., Larin,Z., Montoliu,L., Thies,E., Kelsey,G., Lehrach,H. and Schutz,G. (1993) *Nucleic Acids Res.*, **21**, 4783–4787.
26. Hogan,B., Costantini,F. and Lacy,E. (1986) *Manipulating the Mouse Embryo*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
27. Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
28. Forrester,W.C., Epner,E., Driscoll,M.C., Enver,T., Brice,M., Papayannopoulou,T. and Groudine,M. (1990) *Genes Dev.*, **4**, 1637–1649.
29. Fujimura,F.K., Deininger,P.L., Friedmann,T. and Linney,E. (1981) *Cell*, **23**, 809–814.
30. Liu,Q., Bungert,J. and Engel,J.D. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 169–174.
31. Magram,J., Chada,K. and Costantini,F. (1985) *Nature*, **315**, 338–340.
32. Townes,T.M., Lingrel,J.B., Chen,H.Y., Brinster,R.L. and Palmiter,R.D. (1985) *EMBO J.*, **4**, 1715–1723.
33. Chada,K., Magram,J. and Costantini,F. (1986) *Nature*, **319**, 685–686.
34. Kollias,G., Wrighton,N., Hurst,J. and Grosveld,F. (1986) *Cell*, **46**, 89–94.
35. Higgs,D.R. (1998) *Cell*, **95**, 299–302.
36. Swartzendruber,D.E. and Lehman,J.M. (1975) *J. Cell. Physiol.*, **85**, 179–187.
37. Amati,P. (1985) *Cell*, **43**, 561–562.
38. Krippl,B., Griep,A.E., Mahon,K.A., Bohnlein,E., Gruss,P. and Westphal,H. (1988) *Nucleic Acids Res.*, **16**, 8963–8976.
39. Thomas,K.R. and Capecchi,M.R. (1987) *Cell*, **51**, 503–512.
40. Thompson,T. and Fan,H. (1985) *Mol. Cell. Biol.*, **5**, 601–609.
41. Chung,J.H., Whiteley,M. and Felsenfeld,G. (1993) *Cell*, **74**, 505–514.
42. Li,Q. and Stamatoyannopoulos,G. (1994) *Blood*, **84**, 1399–1401.
43. Yu,J., Bock,J.H., Slightom,J.L. and Villeponteau,B. (1994) *Gene*, **139**, 139–145.