MOLECULAR AND CELLULAR BIOLOGY, May 2005, p. 3864–3874 Vol. 25, No. 10
0270-7306/05/$08.00+0 doi:10.1128/MCB.25.10.3864–3874.2005

The Human β-Globin Locus Control Region Can Silence as Well as Activate Gene Expression

Yong-Qing Feng,[1] Renaud Warin,[1] Taihao Li,[1] Emmanuel Olivier,[1] Arnaud Besse,[1] Amanda Lobell,[1] Haiqing Fu,[2] Chii Mei Lin,[2] Mirit I. Aladjem,[2] and Eric E. Bouhassira[1]*

Department of Medicine, Division of Hematology, and Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York,[1] and Laboratory of Molecular Pharmacology, National Cancer Institute, Bethesda, Maryland[2]

Received 17 September 2004/Returned for modification 25 October 2004/Accepted 1 March 2005

Using recombinase-mediated cassette exchange to test multiple transgenes at the same site of integration, we demonstrate a novel chromatin context-dependent silencer activity of the β-globin locus control region (LCR). This silencer activity requires DNase I hypersensitive sites HS2 and HS3 but not HS4. After silencing, the silenced cassettes adopt a typical closed chromatin conformation (histone H3 and H4 deacetylation, histone H3-K4 methylation, DNA methylation, and replication in late S phase). In the absence of the LCR at the same site of integration, the chromatin remains decondensed. We demonstrate that the LCR is necessary but not sufficient to trigger these chromatin changes. We also provide evidence that this novel silencing activity is caused by transcriptional interference triggered by activation of transcription in the flanking sequences by the LCR.

The β-globin gene loci have long been used to study the role of chromatin structure in the regulation of transcription during development and differentiation in mammalian cells, in part because the β-like globin genes are very tightly regulated. During ontogeny, the ε-globin gene is activated in the yolk sac and silenced by the tenth week of gestation. At that time point, the γ-globin genes become the predominantly expressed β-like globin genes. In turn, around birth, the γ-globin genes are silenced and replaced by the β- and δ-globin genes. The mechanism of silencing of the globin genes and of globin transgenes during development has been under intense investigation but remains controversial (5, 6, 8, 11, 17, 29).

Regulation of globin transgenes also has been studied extensively because silencing has been and remains an impediment to gene therapy of the hemoglobinopathies (32). The insertion of transgenes at random sites of integration results in position effects that are believed to be caused by the juxtaposition of regulatory elements in the transgene and at the site of integration. At some sites of integration, the expression of the transgene is relatively homogenous for all cells, but the level of expression varies depending on the chromosomal integration site (stable position effects). At other sites, a stochastic process progressively silences the transgene, resulting in mosaic expression patterns and, eventually, a complete shutdown of expression (silencing or variegating position effects). The rate of silencing is a characteristic of the site of integration (15, 43).

The mechanisms of position effects, and in particular of transgene silencing, are poorly understood. Whether transgene silencing and the programmed gene silencing that occurs during development and differentiation share the same basic mechanisms is also unknown. In the last few years, we have attempted to address some of these questions using the locus control region (LCR) and the β-globin promoter as a model.

The human β-globin LCR, a group of five DNase I-hypersensitive sites located 6 to 20 kb upstream of the ε-globin genes, has been shown to be the major *cis* regulator of all five β-like globin genes, since in its absence the expression of all globin genes is almost totally suppressed at all stages of development (4). It was initially reported that the LCR is dominant over all position effects (18), and this property became the defining characteristic of LCRs (27). However, we (1) and others (21, 31) have reported that the LCR is not completely dominant over position effects in either cell culture or transgenic mice. Indeed, even 150-kb yeast artificial chromosomes containing all of the known regulatory sequences of the β-globin gene cluster are subject to both stable and variegating position effects in transgenic mice.

Using recombinase mediated cassette exchange (RMCE), a method of performing site-specific chromosomal integration for mammalian cells (12), we previously observed that in mouse erythroleukemia (MEL) cells, transgenes containing sequences from the β-globin locus are subject to both stable and silencing position effects (16). In particular, cassette 234-β-EGFP, which contains the β-globin promoter, the enhanced green fluorescent protein (EGFP) coding sequence, and the mini-LCR (a collection of the central regions from hypersensitive sites 2, 3, and 4 of the LCR) (13) was shown to be expressed at a high level when integrated in one orientation (termed the permissive orientation) but silenced when it was integrated at the same site but in the opposite orientation (termed the nonpermissive orientation). Importantly, this orientation-dependent silencing was observed at two different sites of integration, RL4 and RL6, suggesting that the phe-

* Corresponding author. Mailing address: Division of Hematology/Department of Medicine, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461. Phone: (718) 430-2188. Fax: (718) 824-3153. E-mail: bouhassi@aecom.yu.edu.

nomenon was not specific for one site of integration. At RL4, the silencing took less than 8 weeks, while at RL6, more than 16 weeks were required. In the present report, we took advantage of the rapid rate of silencing at RL4 to further investigate the mechanisms of silencing.

MATERIALS AND METHODS

Plasmids. All plasmids were constructed using standard methods. Cassette EGFP contains the EGFP coding sequence (Clontech, Palo Alto, CA). Cassette β-EGFP contains a −350 to +50 (relative to the cap site) fragment of the β-globin promoter fused at the ATG of the EGFP cassette. Cassette 234-β-EGFP has the mini-LCR, as defined in reference 16, fused to the β-EGFP cassette. Cassette 4-β-EGFP has HS4, as defined in reference 16, fused to the β-EGFP cassette. In cassette 234inv-β-EGFP, the mini-LCR of 234-β-EGFP was inverted by digestion with appropriate restriction enzymes and religation. In cassette 234-β-EGFPinv, the β-EGFP was inverted by digestion with appropriate restriction enzymes and religation. In cassette phosphoglycerokinase (PGK)-EGFP, a −500 to +1 (relative to the cap site) fragment of the mouse PGK1 promoter was fused to the EGFP coding sequence. Cassette 234-PGK-EGFP was constructed by inserting the mini-LCR of cassette 234-β-EGFP just upstream of the PGK promoter in cassette PGK-EGFP. Sequences of all plasmids are available on request.

RMCE. RMCE is a method to perform site-specific chromosomal integration of cassettes in mammalian cells using the Cre recombinase. RMCE was performed as previously described (16). Briefly, previously constructed target cells containing the L1-CMV-HYTK-1L cassette (a cytomegalovirus [CMV]–hygromycin–herpes simplex virus–thymidine-kinase fusion selectable cassette [HYTK] flanked by inverted L1 Lox sites) inserted at a site of integration termed RL4 were coelectroporated with 50 μg of a Cre expression plasmid and 200 μg of an exchange plasmid containing the cassette of interest. Clones having lost the HYTK cassette and integrated with the exchange cassette were selected for by using 10 nM ganciclovir. Replacement of the HYTK cassette by the cassette present in the exchange plasmid was then verified by Southern blotting for at least 12 clones. The frequency of exchanges varied between 50 and 100% of the ganciclovir-resistant clones. Each series of transfections performed included a 234-β-EGFP control that was used for normalization of the florescence-activated cell sorter (FACS) results. At least three, but typically five to eight clones with each cassette in each orientation were then monitored by FACS for at least 3 months.

FACS analysis. Dead cells were gated out on the basis of morphological parameters (forward scatter and side scatter) and propidium iodide exclusion. The percentage of cells expressing the EGFP transgene was estimated using untransfected MEL cells as the negative control. The mean expression of the EGFP cassette was determined by multiplying by 100 the ratio of the mean fluorescence in the green channel of the tested cassette to the mean fluorescence in the green channel of the control cassette (234-β-EGFP) measured on the same day. Data presented in Fig. 1 to 4 were collected 3 months after integration of the cassettes. This 3-month time point was chosen because at that time, except for the 234-β-EGFPinv cassette, all cassettes were either completely silenced or were expressed in close to 100% of the cells. Before that time, the silencing process was not always completed, complicating the analysis. Mean expression data for the 234-β-EGFPinv were calculated 2 months after transfection.

Methylation analysis. Southern blots and bisulfite sequencing were performed as described previously (13). Primers for the promoter analyses were +bisProm-1, 5′ TATGAAAATAGGAAAAGAAAATAAATTTTG; −bisProm-1, 5′ CCCCCCGGATCCCTAACTTTTATACCCAACCCTAA; +bisProm-2, 5′CCC CCCGGATCCGTATTTTTGGATAGTTAGGTGGT; and −bisProm-2, 5′CC CCCCGGATCCAATTAACCAACCCTAAAATATAA.

Cloning of RL4. Genomic DNA extracted from cells with cassette 234-β-EGFP in the permissive orientation was digested with EcoRV, SspI, ScaI, or StuI, and ligated to partially double-stranded, NH2-blocked, phosphorylated adaptors (upper strand, 5′ GTA ATA CGA CTC ACT ATA GGG CAC GCG TGG TCG ACG GCC CGG GCT GGT 3′; lower strand, 5′ phosphate ACC AGC CC-NH2). Nested PCR with adaptor-specific primers 5′ GTA ATA CGA CTC ACT ATA GGG C (first PCR), and 5′ ACT ATA GGG CAC GCG TGG T (second PCR) and gene-specific primers (HS4 region) 5′ CAC TAA GAA AAA GAA TGA GAT GTC TAC (first PCR) and 5′ TAT ACC CTG CGT CCC CTC TTG TGT ACT (second PCR) were then performed. These secondary PCRs yielded unique bands for the EcoRV and PvuII digests. The subcloning and sequencing of these bands yielded overlapping sequences containing the end of cassette 234-β-EGFP and about 500 bp of sequence corresponding to the flanking region on one side of the integration site.

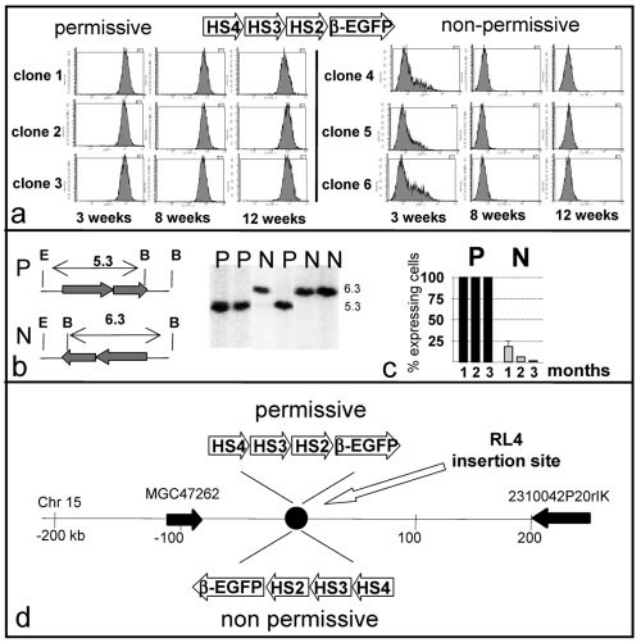

FIG. 1. Silencing of cassette 234-β-EGFP at RL4. (a) FACS analysis of six MEL cell clones with cassette 234-β-EGFP inserted at RL4. The *x* axis shows EGFP fluorescence (FL-1); the *y* axis shows cell numbers. Dead cells were gated out on the basis of forward scatter and side scatter and on the basis of propidium iodide exclusion. The cassette is stably expressed when it is inserted in the permissive orientation, but is rapidly silenced in the nonpermissive orientation. (b) Insertion of cassette 234-β-EGFP at RL4 was demonstrated by Southern blotting using the EGFP coding sequence as a probe. P, permissive orientation; N, nonpermissive orientation; E, EcoRV; B, BglII. The histogram on the right depicts the percentage of cells expressing the EGFP transgene as a function of time. (c) Map of site RL4. After integration by RMCE, the cassettes studied are integrated at position 97,816,844 of mouse chromosome 15 (February 2003 assembly in the genome browser). The closest gene to the site of integration (MGC47262) is about 87 kb away.

Replication timing analyses were performed as described previously (22). Briefly, asynchronously growing cells were labeled with the thymidine analog BrdU for 90 min and fractionated based on DNA content after Hoechst 33342 labeling using a Vantage FACS. Newly replicated, BrdU-substituted DNA was isolated by immunoprecipitation. After two rounds of immunoprecipitation and DNA isolation, mitochondrial DNA sequences (mMT primers) were used to verify that each fraction contained similar quantities of amplifiable DNA strands. Samples containing 8 nanograms of DNA were analyzed by real-time PCR on an ABI 7900 system. The abundance of an amplified sequence was calculated using genomic DNA standards with a standard curve ranging from 0.005 ng to 50 ng per reaction. Each measurement was performed in triplicate. The relative abundance for each probe/primer combination was calculated as the ratio of the number of molecules amplified from a specific cell cycle fraction to the number of molecules amplified from the cell cycle fraction where amplification was maximal. The analyses were performed at least 3 months after integrations. The clones in the nonpermissive orientation were therefore completely silenced.

Chromatin immunoprecipitations (ChIP) were carried out as described previously (22). Briefly, MEL cells were treated with 1% formaldehyde in the growth medium, washed, and subject to sonication. Chromatin was isolated, precleared by incubation with protein A, and incubated with antibodies (anti-acetyl-histone H3 [06-599], anti-acetyl-histone H4 [06-866], and anti-dimethyl-histone H3 [Lys 4; 07-030], all from Upstate Biotechnology, Lake Placid, NY). After precipitation with protein A beads and a series of washes, DNA was isolated and subject to real-time PCR analysis using the primers and probes listed below. The enrichment of specific sequences in precipitated DNA relative to genomic "input"

TABLE 1. Primers used in this study[a]

Primer name and sequence	Position
EB 797; 5′CCTGTAGTAGCGAAGTATGG	−19991–−19845
EB 798; 5′CTCCTGCTCATGCTGTC	
EB 799; 5′AACCAAACCGAAACAGAG	−3677–−3541
EB 800; 5′TGGAACTGGCAGTGAG	
EB 431; 5′GGCTTGGATCCCGCGCTTCTCTGA	−161–0
EB 727; 5′GCCTTGCAGCACATCC	
EB 785; 5′GGAGCGCAATAGCCAG	123–242
EB 786; 5′CCTTACATTCCGTCCCT	
EB 801; 5′AGATCCACCGTCACAG	3315–3426
EB 802; 5′GCTAAGTCCCTCTCCTTC	
EB 805; 5′GGTTAATACAAAGGGTAGGGA	14432–14536
EB 806; 5′CCCATACAGTTACTCTGTG	
EB 807; 5′AAAGCGTTAGGAGCCA	20410–20546
EB 808; 5′ATATGAAACGTCTTGCCAC	
EB 740; 5′AACTACAAGACCCGCG	316–449
EB 741; 5′CGGCCATGATATAGACGT	

[a] Position numbering is relative to the RL4 integration site (except for the last pair of primers, for which the numbering is relative to the ATG of the EGFP coding sequence).

DNA was calculated as $2^{(\text{"input"Ct-IPCt})}$. (Ct indicates the cycle threshold number calculated as the average of results of three real-time PCRs.). The primers and probes used are as described in reference 22. The analyses were performed at least 3 months after integrations. The clones in the nonpermissive orientation were therefore completely silenced.

Nuclear RNA extraction. One million to 10 million MEL cells were lysed by resuspension in 10 mM Tris (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, and 0.5% NP-40 for 5 min on ice. Nuclei were recovered by low-speed centrifugation (150 × *g*; 5 min at 4°C) and rinsed twice with ice-cold phosphate-buffered saline, and RNAs were extracted with TRIzol (Life Technologies) following the manufacturer's instructions. DNA was then eliminated from the RNA preparation by digestion with DNase I using a DNA removal kit (Turbo-DNA Free; Ambion).

Quantitative RT-PCR reactions were performed with a Roche Light Cycler using a QIAGEN Quantitech SYBR green RT-PCR kit. The primers used are listed in Table 1. PCR products from each primer pair in Table 1 were gel purified, and DNA concentrations were determined using a PicoGreen double-stranded DNA quantitation kit (Molecular Probes) according to the manufacturer's instructions. These purified PCR fragments were then used to create standard curves which were used to determine the number of β-2-microglobulin (B2M) and flanking sequence transcripts in nuclear RNA samples extracted from at least two independently derived clones for each cassette in each orientation. Control experiments with no template were performed on every run with each set of primers to detect any contamination by carryover PCR product. Control experiments with no reverse transcriptase were performed on each RNA sample to ascertain that there was no DNA contamination. Amounts of RNA corresponding to about 500,000 molecules of B2M were used in each reaction. Each PCR was performed in duplicate on at least two independently derived clones.

RESULTS AND DISCUSSION

The RL4 site resides in murine chromosome 15. The RL4 integration site was created by random integration of the plasmid L1HYTK1L (16). To better understand the silencing process, it was important to determine the exact site of integration and to identify potential regulatory sequences nearby.

To confirm our previous results at RL4 and obtain fresh genomic DNA necessary to clone the RL4 site, we integrated cassette 234-β-EGFP at RL4 by RMCE, identified by Southern blot clones with the cassette integrated in each orientation, and measured the expression of EGFP using a FACS (Fig. 1a and b). As expected, all clones exhibited a high level of fluorescence that was stable over time when the expression cassette was integrated in the permissive orientation, while rapid silencing occurred in all clones with the cassette in the nonpermissive orientation (Fig. 1a and c).

PCR-mediated genome walking (37) and sequencing yielded a 500-base fragment on one side of the cassette that mapped to the tip of chromosome 15, according to the mouse genome sequence (February 2003 assembly). This result was extended by PCR and sequencing analysis with primers designed to amplify the junction fragment on the other side of the integrated cassette, and by fluorescent in situ hybridization analysis with bacterial artificial chromosome rp24-497n14, which mapped to the putative integration site (data not shown).

Bioinformatics analysis revealed that the closest gene is 87 kb away on one side and more than 200 kb away on the other side (Fig. 1d). An analysis of the region of human chromosome 4 syntenic with the RL4 site in the mouse supported the conclusion that the integration site does not contain any nearby gene. A close examination of the repetitive elements present in the vicinity of the site of integration did not reveal any notable characteristics both in terms of amount and type of repeats present. We conclude from this analysis that the orientation-dependent silencing is not caused by integration inside or near another gene.

Silencing at RL4 depends on the orientation of the promoter and EGFP coding sequence relative to the flanking sequence. To identify the *cis*-acting regulatory sequences in the transgene that are involved in the silencing, we inserted at site RL4 new cassettes in which either the gene (cassette 234-β-EGFPinv) or the mini-LCR (cassette 234inv-β-EGFP) were inverted and monitored their expression by FACS for 3 months (Fig. 2a and b). In the permissive orientation, the mean fluorescence levels of cassettes 234inv-β-EGFP and 234-β-EGFPinv were stable over time and were, respectively, 110% and 40% that of the control 234-β-EGFP cassette (Fig. 2c and d).

In the nonpermissive orientation, the cassette with the inverted mini-LCR silenced, albeit a little more slowly than the control 234-β-EGFP cassette (Fig. 2c). By contrast, the cassette in which the gene and promoter are inverted was almost stable and silenced extremely slowly, with more than 80% of the cells still expressing the transgene 3 months after integration in every clone tested (Fig. 2d). We conclude that the orientation of the promoter and coding sequence relative to the chromosomal context is a major factor in the silencing, since silencing seems to require that the promoter and the EGFP coding sequence point toward the centromeric side and be juxtaposed with the flanking sequence on the centromeric side.

Silencing at RL4 requires the mini-LCR. We then created and inserted at RL4 cells containing three new cassettes, as follows: cassette 234, which contained the mini-LCR alone; cassette β-EGFP, which contained the β-globin promoter and the EGFP coding sequence; and cassette EGFP, which contained only the EGFP coding sequence (Fig. 2a). As expected, cassette EGFP was not expressed in either orientation since it contains no promoter and no enhancer (Fig. 2e). Importantly, cassette β-EGFP was expressed at low but easily detected

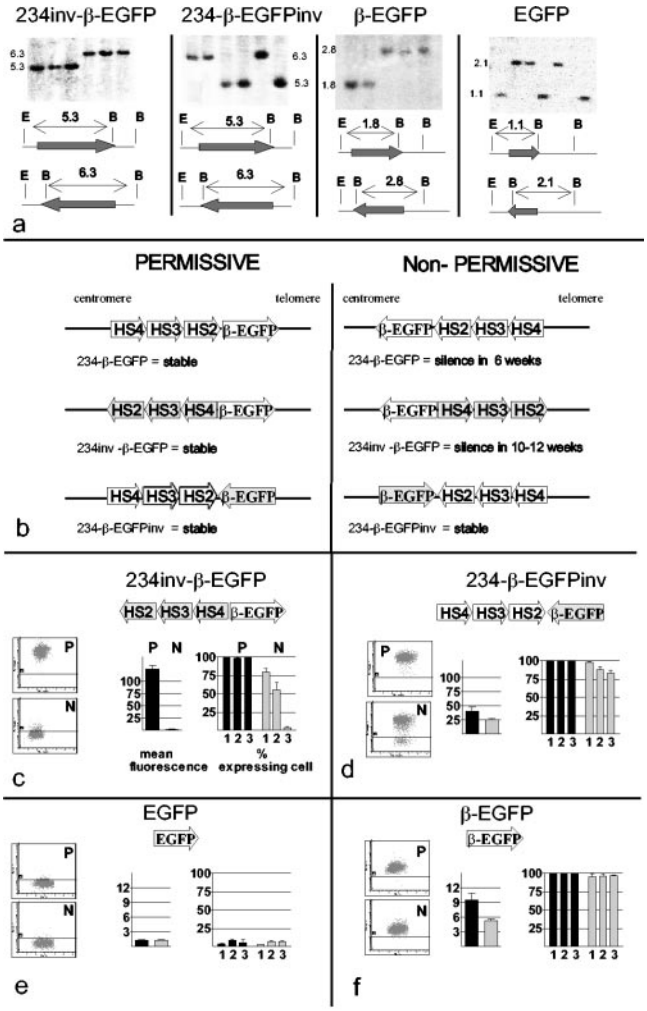

FIG. 2. The LCR has a chromatin context-dependent silencer activity. (a) Southern blot demonstrating the integration of cassettes 234inv-β-EGFP, 234-β-EGFPinv, β-EGFP, and EGFP at RL4. The upper band corresponds to the nonpermissive (N) orientation; the lower band corresponds to the permissive (P) orientation. In cassette 234inv-β-EGFP, the mini-LCR is inverted relative to the 234-β-EGFP cassette; in cassette 234-β-EGFPinv, the β-EGFP transcription unit is inverted; in cassette β-EGFP, the LCR is deleted; in cassette EGFP, both the LCR and the β-globin promoter are deleted. (b) Summary of the structure of the RL4 site after integration of cassettes 234-β-EGF, 234inv-β-EGFP, and 234-β-EGFPinv at RL4. (c to f) Dot plots illustrate the mean green fluorescence of each cassette in both orientations as determined by FACS analysis 3 months after integration. The *x* axis shows the forward scatter; the *y* axis shows EGFP fluorescence (FL-1). Cells below the horizontal bar do not express any EGFP (as determined by analysis of untransfected cells). Dead cells were gated out as in Fig. 1. Histograms in the middle of each panel summarize the results. The means of the linearized mean fluorescence levels in the FL1 (EGFP) channel (± standard deviations) for at least three clones in each orientation (permissive [P] and nonpermissive [N]) 3 months after integration is depicted. The levels of green fluorescence are normalized to that of the 234-β-EGFP control cassette integrated on the same day at RL4. The *y* axis values are 100 × (average FL1 fluorescence of cassette)/(average FL1 fluorescence of 234-β-EGFP cassette). Histograms on the right summarize the mean percentages of cells expressing the EGFP transgene 1, 2, and 3 months after integration at RL4. The cutoff between expressing and nonexpressing cells was determined for each experiment by analysis of untransfected control cells. Cassette 234inv-β-EGFP silences a little bit slower than the control cassette (Fig. 1a). Cassette 234-β-EGFPinv silences dramatically slower than the control cassette, suggesting that the orientation of the gene in the locus plays a critical role in silencing. Cassette β-EGFP does not silence in either orientations, demonstrating that the LCR has a chromatin context-dependent silencing activity. As expected, cassette EGFP does not express at RL4.

levels in both orientations (Fig. 2f). In the permissive orientation, cassette β-EGFP was expressed at about 10% of the control 234-β-EGFP cassette. In the nonpermissive orientation, the mean fluorescence level of cassette β-EGFP was about twofold lower than in the permissive orientation, but silencing did not occur, even after more than 6 months in culture. These experiments demonstrate that in the permissive orientation the mini-LCR acts, as expected, as a strong enhancer but that in the nonpermissive orientation, it acts as a silencer of the β-globin promoter. To determine if the mini-LCR-induced silencing is associated with the formation of a typical condensed chromatin structure, we then performed a series of experiments to characterize the silenced transgene.

Mini-LCR-induced silencing is associated with late replication. We have recently reported that DNA replication of the RL4 site on chromosome 15 occurs late in S phase and that insertion of the HYTK-selected marker advances replication to early S phase. We have also shown that after replacement of that marker by the 234-β-EGFP cassette in the permissive orientation, replication continues to occur early in S phase, but that the timing of replication is delayed to late S phase when the cassette is integrated in the nonpermissive orientation (22).

To determine if these changes in timing of replication are mini-LCR-dependent, we compared the replication timing of the β-EGFP and 234-β-EGFP cassettes inserted at RL4 in both orientations by measuring the abundance of sequences from the transgene in BrdU-labeled, newly replicated DNA from asynchronous cells fractionated by FACS according to their position in the cell cycle. As controls, we also determined the timing of replication of sequences in the murine β-globin and amylase loci, which are, respectively, known to replicate early and late in S phase in these cells (22). Figure 3a shows that EGFP sequences from cells containing the β-EGFP cassette in both orientations at RL4 were abundant in the early S-phase fractions. By contrast, in cells containing the 234-EGFP cassette, the EGFP sequences were enriched in the early S-phase cells when the cassette was in the permissive orientation but were most abundant in the late S-phase cells when the cassette was in the nonpermissive orientation. We conclude that the switch to late replication at RL4, which is associated with silencing, is mini-LCR-dependent.

Chromatin condensation at RL4 requires the mini-LCR. We have previously shown that silencing is also associated with chromatin condensation at the RL4 locus (22). To determine whether this chromatin condensation depends on the mini-LCR, we used a ChIP assay to measure the levels of acetylation and methylation of histones H3 and H4 in chromatin derived from cells containing the β-EGFP and 234-EGFP cassette insertions. As shown in Fig. 3b, sequences from the inserted β-EGFP transgene were abundant in chromatin containing acetylated histones H3 and H4 when the β-EGFP cassette was inserted into RL4 in both orientations. By contrast, sequences from the 234-β-EGFP cassette were abundant in acetylated

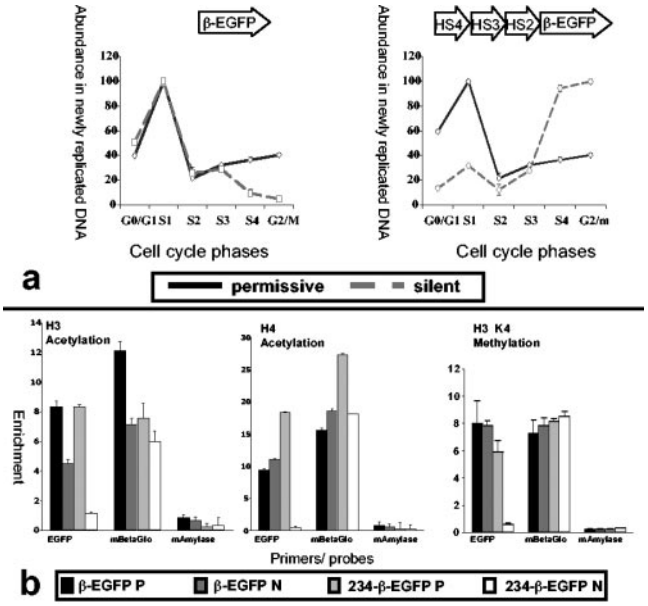

FIG. 3. The LCR affects replication timing and chromatin. (a) Replication timing of cassettes β-EGFP and 234-β-EGFP inserted into the RL4 site. The β-EGFP cassette replicates early in both orientations, while the 234-β-EGFP replicates early in the permissive and late in the nonpermissive orientation. As expected, a murine β-globin locus control exhibited early replication in both cell lines, while an amylase locus control replicated late (data not shown). Solid line, permissive orientation; dashed line, nonpermissive orientation. (b) ChIP assays. Right, middle, and left panels, enrichment of the cassettes and controls in chromatin immunoprecipitation using, respectively, an anti-acetylated histone H3 antibody, an anti-acetylated histone H4 antibody, and an antibody against methylated lys4 on histone H3. β-EGFP, coding sequence of the β-EGFP gene; mAmylase, murine amylase locus; mβGlo, sequences from the murine β-major globin locus; P, permissive orientation; N, nonpermissive orientation.

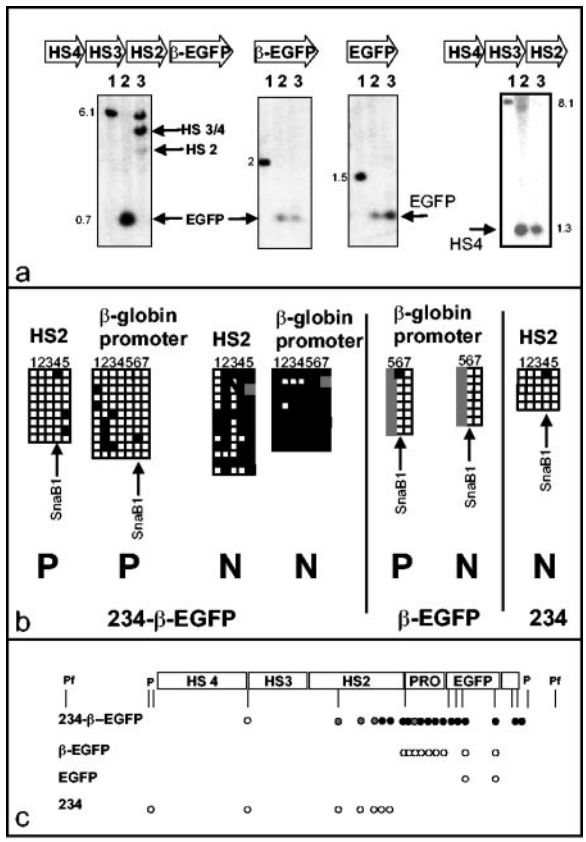

FIG. 4. DNA methylation analysis. (a) Genomic DNAs from cells with cassettes 234-β-EGFP (control), β-EGFP (LCR deletion), and EGFP (coding sequence only), integrated at RL4 in either orientations, were extracted 3 months posttransfection and digested with either PvuII alone or PvuII and HpaII and probed with the EGFP coding sequence. Genomic DNA from cells with cassette 234 at RL4 were digested with PflMI alone or with PflMI and HpaII and probed with a fragment of HS4 of the LCR. Lanes 1, genomic DNA with cassette in P orientation digested with PvuII (or PflMI); lanes 2, same DNA digested with PvuII (or PflMI) plus HpaII; lanes 3, genomic DNA with cassette in N orientation digested with PvuII (or PflMI) and HpaII. (b) Methylation of the promoter and HS2 in cassettes 234-β-EGFP, β-EGFP, and 234 was analyzed by subcloning of PCR products obtained from bisulfite-converted genomic DNA followed by sequencing of subclones. Each row represents a single molecule. Black squares represent methylated CpGs; white squares, unmethylated CpG; grey squares, nonanalyzed CpGs. CpGs 1, 2, 3, 4 in the promoter fragment do not exist in the genomic sequences and come from residual multiple cloning sites in cassette 234-β-EGFP. CpG 4 of HS2 and CpG 6 of the promoter fragment correspond to SnaBI sites in the corresponding genomic DNA. The control 234-β-EGFP cassette is lightly methylated in the permissive orientation (P) and heavily methylated in the nonpermissive orientation (N). Very little methylation could be detected in the β-EGFP or in the 234 cassette, suggesting that methylation at RL4 requires the combined presence of the mini-LCR and of β-EGFP. (c) Summary of methylation analysis for cassettes in the nonpermissive orientation. Vertical bars below the map represent HpaII restriction sites. Circles not aligned with an HpaII site correspond to CpG dinucleotides analyzed by bisulfite sequencing. Open circles, unmethylated CpG; grey circles, partially methylated CpG; black circles, methylated CpG.

histone-containing chromatin only when the cassette was inserted in the permissive orientation. As controls, we performed the same experiments on the mouse β-globin gene and amylase gene, which are known to be associated with condensed and decondensed chromatin, respectively, in these cells. As expected, sequences from the mouse β-globin gene but not from the amylase gene were enriched in chromatin precipitated with anti-acetyl H3 and H4 antibodies (Fig. 3b). Similar results were observed with antibodies against methylated lysine 4 of histone H3 (Fig. 3b).

DNA methylation at RL4 requires the mini-LCR. Silencing of cassette 234-β-EGFP at RL4 is also associated with DNA methylation of HS2 of the mini-LCR, the β-globin promoter, and the EGFP coding sequence (16).

DNA methylation at RL4 in the nonpermissive orientation also requires the mini-LCR. We used methylation-sensitive restriction enzymes (Fig. 4a) and bisulfite sequencing (Fig. 4b) to determine the DNA methylation status of cassettes 234-β-EGFP, β-EGFP, and EGFP 3 months postintegration. As previously reported, the coding sequence, the promoter, and HS2 of cassette 234-β-EGFP were heavily methylated in the nonpermissive orientation. By contrast, little or no methylation could be detected for cassettes β-EGFP and EGFP, demonstrating that DNA methylation at RL4 in the nonpermissive orientation is induced by the presence of the mini-LCR. These results are summarized in Fig. 4c.

Presence of a transcription unit is necessary for DNA methylation at RL4. In order to test whether the presence of the

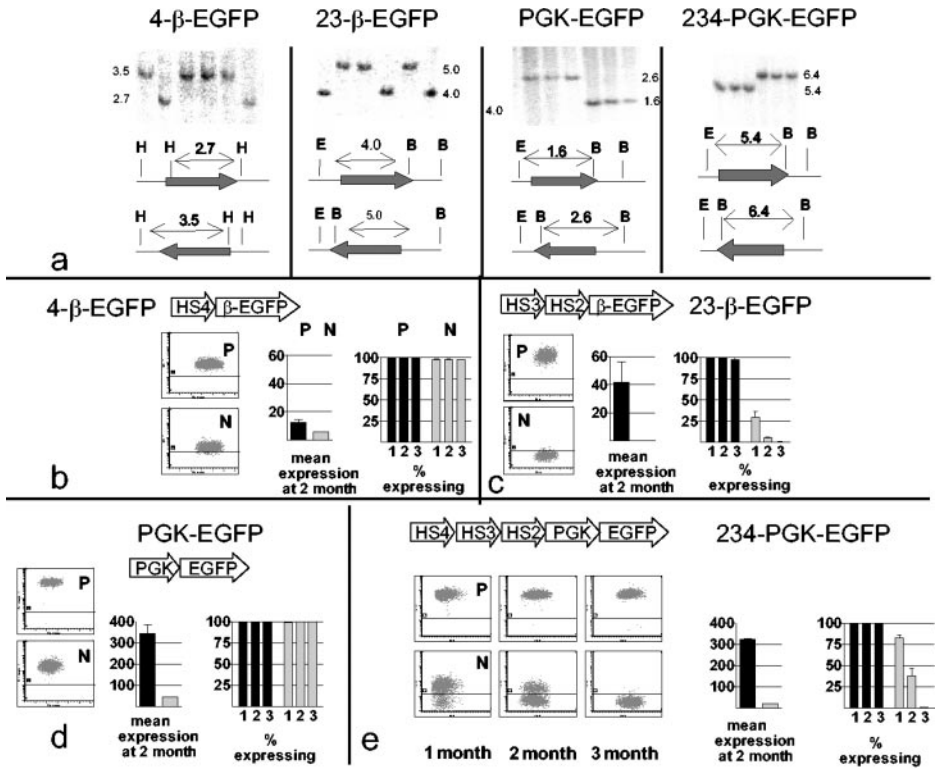

FIG. 5. Silencing at RL4 is not related to expression level. (a) Southern blots demonstrating insertion of cassettes 4-β-EGFP, 23-β-EGFP, PGK-EGFP, and 234-PGK-EGFP at RL4 in both orientations. H, HindIII; E, EcoRI; and B, BglII. Panels b to e are as in Fig. 2. In the permissive orientation (P), cassettes 4-β-EGFP and PGK-EGFP are expressed at about 10 and 320%, respectively, of the control 234-β-EGFP cassette and do not silence in the nonpermissive orientation (N), suggesting that silencing is not linked to expression level. Silencing of cassette 23-β-EGFP but not of 4-β-EGFP demonstrates that HS3 and HS2, but not HS4, are required for silencing. Silencing of cassette 234-PGK-EGFP confirms the observation that the LCR causes silencing (Fig. 2) and demonstrates that the silencing activity of the LCR is not specific to the β-globin promoter.

mini-LCR alone at RL4 is sufficient to induce the chromatin changes in the nonpermissive orientation at RL4, we determined the DNA methylation status of cassette 234 when it is integrated at RL4. Expression cannot be measured with this cassette, because the cassette does not include any gene. Analysis by methylation-sensitive restriction enzymes and bisulfite sequencing 3 months posttransfection revealed that insertion of the mini-LCR did not lead to any detectable DNA methylation in either orientation (Fig. 4a and b). We conclude that the mini-LCR is necessary but not sufficient to induce methylation at RL4. DNA methylation is therefore triggered by interactions between the mini-LCR, the β-EGFP cassette, and some feature of the integration sites (to account for the orientation dependence of the silencing).

HS2 and HS3 but not HS4 of the LCR mediate silencing. HS2 or HS3 linked to the β-globin promoter is sufficient to cause rapid silencing at RL4 in the nonpermissive orientation (28). To determine if all hypersensitive (HS) sites of the mini-LCR can induce silencing, we integrated at RL4 a new cassette, 4-β-EGFP, in which HS4 alone is inserted upstream of the β-EGFP gene (Fig. 5a). In the permissive orientation, the 4-β-EGFP cassette was expressed at about 10% of the 234-β-EGFP control cassette (Fig. 5b). In the nonpermissive orientation, the 4-β-EGFP cassette did not silence and was stably expressed at about 50% of the level of the same cassette in the permissive orientation. This result suggested that HS4 is not required for silencing. To confirm this result, we created cassette 23-β-EGFP, a cassette that contains HS2 and 3 but lacks HS4, and inserted it at RL4 (Fig. 5a). As expected, cassette 23-β-EGFP silenced as fast as cassette 234-β-EGFP in the nonpermissive orientation and was expressed at about 40% of the level of cassette 234-β-EGFP in the permissive orientation (Fig. 5c). We conclude that the silencing activity of the mini-LCR involves HS2 and HS3 but not HS4. These results also confirmed our previous observation that HS4 does not have enhancer activity when it is alone but that it has some enhancer activity in synergy with HS2 and HS3 (28).

Silencing at RL4 does not depend on the level of expression of the transgene. β-EGFP and 4-β-EGFP, the two cassettes that did not silence at RL4 in the nonpermissive orientation, are expressed at low levels (about 5% of the 234-β-EGFP cassette in the permissive orientation), suggesting that the silencing might occur only in the presence of strong expression. To test this hypothesis, we constructed and inserted at RL4 a cassette driven by a highly expressed promoter in MEL cells, the mouse PGK1 promoter, which is a housekeeping promoter embedded in a CpG island (26) (Fig. 5a). In the permissive orientation, the PGK-EGFP cassette was expressed at about 300% of the level of the control 234-β-EGFP cassette (Fig. 5d). Importantly, in the nonpermissive orientation, the same cassette did not silence for at least 6 months after the integra-

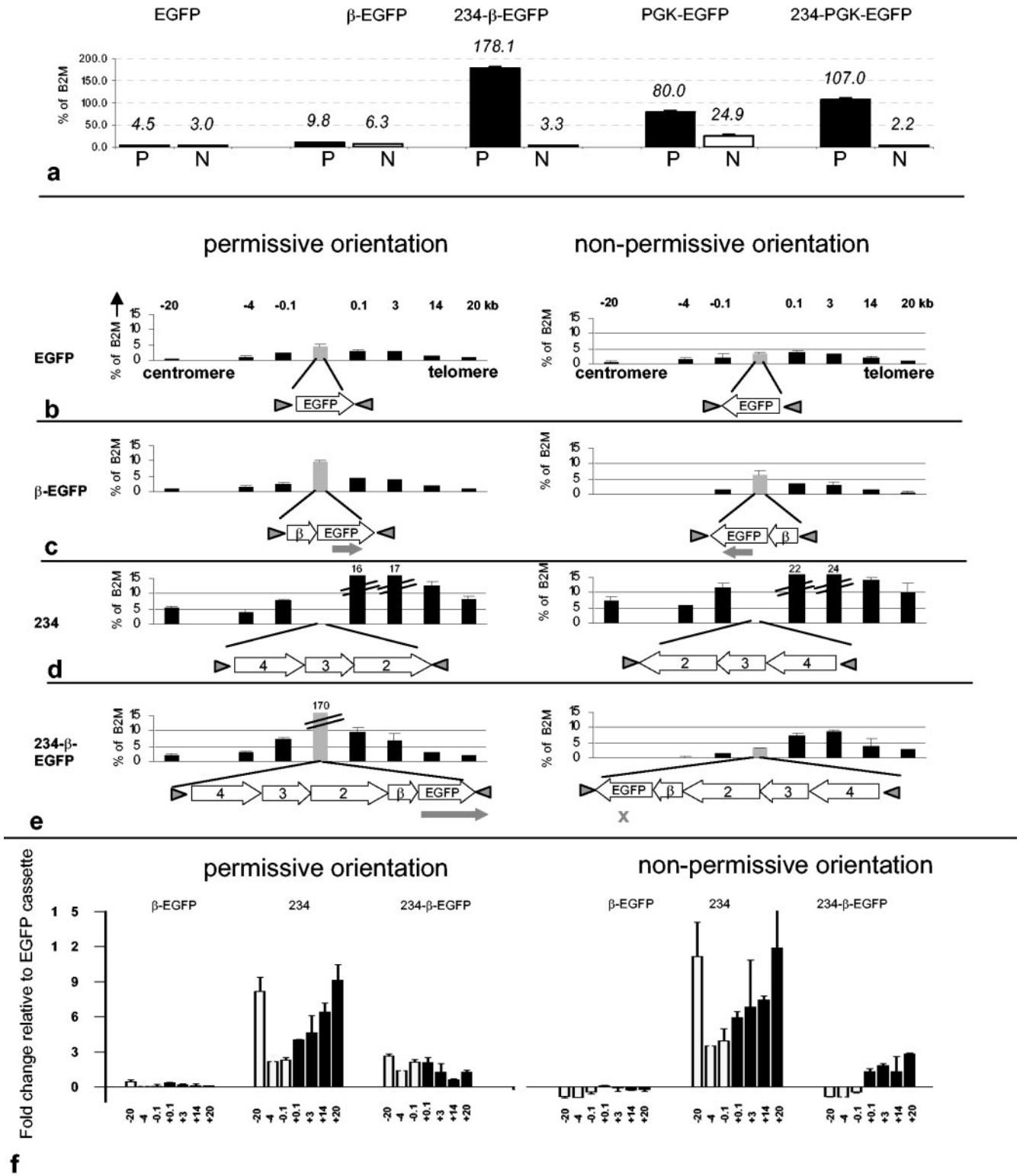

FIG. 6. Quantitative RT-PCR analysis of nuclear transcription of the EGFP coding sequence and of the flanking sequences. (a) Histograms depict the percentage of EGFP transcripts relative to B2M transcripts in nuclear RNAs extracted from cells with cassettes, EGFP, β-EGFP, 234-β-EGFP, PGK-EGFP, or 234-PGK-EGFP at RL4 3 months postintegration (*y* axis, 100 × number of EGFP transcripts/number of B2M transcripts). Results were obtained by quantitative RT-PCR and are expressed as means ± standard deviations. Levels of EGFP expression detected at the RNA level coincide largely with the levels detected by FACS. P, permissive orientation; N, nonpermissive orientation. (b to e) Histograms summarizing Q-RT-PCR analyses on nuclear RNAs extracted from cells containing cassettes EGFP, β-EGFP, 234, and 234-β-EGFP inserted using seven primer pairs located at −20, −4, −0.1, 0.1, 3, 14 and 20 kb, respectively, from the RL4 integration site. All clones were analyzed at about 3 months postintegration. Results are normalized to the levels of B2M. The grey bars summarize results with a primer pair

tion. Again, expression in the nonpermissive orientation was lower than in the permissive orientation.

The mini-LCR induces silencing of nonglobin promoters. To determine if the mini-LCR could induce silencing of nonglobin promoters, we then constructed 234-PGK-EGFP, in which the mini-LCR is placed upstream of the PGK-EGFP cassette, and inserted it at the RL4 cassette (Fig. 5a). FACS analysis revealed that the mini-LCR did not increase the level of expression of the PGK promoter in the permissive orientation (Fig. 5e). Interestingly, in the nonpermissive orientation, the presence of the mini-LCR caused silencing of the PGK-EGFP cassette in less than 3 months. These results demonstrate that the silencing activity of the mini-LCR is not restricted to the β-globin promoter.

The results shown in Fig. 5d suggest that silencing at RL4 is not a general property of highly expressed cassettes, since the PGK-EGFP cassette is expressed at a very high level when inserted at RL4. However, levels of EGFP expression determined by FACS do not necessarily reflect levels of transcription, since small differences in the structure of the transcripts might lead to differences in posttranscriptional processing, mRNA stability, or translation. We therefore performed a quantitative PCR analysis to determine the levels of EGFP transcripts on nuclear RNAs extracted from cells containing cassettes EGFP, β-EGFP, 234-β-EGFP, PGK-EGFP, and 234-PGK-EGFP integrated at RL4 (Fig. 6a). This revealed that in the permissive orientation, the β-EGFP and the PGK-EGFP cassettes, respectively, were expressed at about 10 and 80% of the levels of B2M determined by the same method as an internal standard. Adding the mini-LCR to these cassettes leads to a 17-fold increase in expression of the β-globin promoter (comparison of cassettes β-EGFP and 234-β-EGFP) but has little or no effects on the PGK promoter (comparison of cassette PGK-EGFP and 234-PGK-EGFP). As expected from the FACS results, in the nonpermissive orientation, cassette β-EGFP and PGK-EGFP were expressed, but at lower levels than in the permissive orientation.

This analysis therefore confirms the FACS data and demonstrates that the differences of expression observed at the protein levels largely reflect differences at the transcriptional level. We conclude that silencing at RL4 does not correlate with transcription levels and that the mini-LCR causes silencing of the PGK promoter in the nonpermissive orientation even though it does not activate it in the permissive orientation.

The mini-LCR induces transcription in the flanking sequences. Our findings that silencing and formation of a closed chromatin structure at RL4 occurs only when the mini-LCR is present and when the β-EGFP transgene faces toward the centromeric side and is juxtaposed to the centromeric sequences suggested that the mechanism of orientation-dependent silencing might involve transcriptional interferences that would be caused by activation of a promoter in the flanking sequence. Since there is no gene nearby, we hypothesized that this promoter might be within the numerous repetitive sequences on both sides of the RL4 integration site.

To test this hypothesis, we designed a series of primer pairs complementary to unique sequences spaced 3 to 10 kb apart on both sides of the integration site and performed quantitative RT-PCR on nuclear RNAs extracted from cells with various cassettes at RL4 (Fig. 6b to e). All measurements were made when silencing was complete, at least 3 months postintegration. As above, all results were normalized to the levels of B2M RNAs. A total of 40 kb of sequences flanking RL4 was scanned for nuclear transcription. All samples were treated with DNase I to avoid amplification of contaminating nuclear DNA. Control experiments in which the reverse transcriptase was not included in the PCR mixture were performed with every RNA sample tested to ascertain that the DNase I treatment had been effective.

Analysis of cells containing the EGFP cDNA alone at RL4 in both orientations (cassette EGFP) revealed that in the absence of any added promoter, the integration site was not completely silent, since we could detect expression ranging from 1 to 3% of B2M in the 40 kb tested (Fig. 6b). In the presence of the β-globin promoter-EGFP transgene, but in the absence of the mini-LCR (cassette β-EGFP), the level of transcription of EGFP was low and there was no detectable increase or decrease of transcription in the flanking sequences regardless of the cassette orientation (Fig. 6c), suggesting that insertion of an enhancerless cassette has little or no effect on transcription of the flanking sequences. An examination of cells with the mini-LCR alone inserted at RL4 (cassette 234) revealed a large increase in transcriptional activity in the flanking sequences, with transcript levels reaching almost 25% of B2M levels on the telomeric side and 12% on the centromeric side of the integration site (Fig. 6d), suggesting the existence of promoters responsive to the mini-LCR in the flanking sequences. In the case of cells containing the mini-LCR-driven β-EGFP transcription unit (cassette 234-β-EGFP) the results were dependent on the orientation of the cassette. In the permissive orientation, transcription could be detected on both sides of the cassette (Fig. 6e). In the nonpermissive orientation, transcription was present on the telomeric side and undetectable on the centromeric side. The nuclear transcription data, which are summarized in Fig. 6f, therefore support our hypothesis that silencing might be caused by transcriptional interference since, as predicted, we observed transcripts in the flanking sequences. However, the pattern of nuclear transcription is quite complex.

Mechanism of silencing at RL4. Our results clearly demonstrate that at site RL4, silencing is an active process that is

located within EGFP (see panel a). Presence of the LCR in the locus causes activation of transcription on both sides of the integration site (cassette 234). In the presence of the LCR plus the β-globin promoter and the EGFP coding sequence (cassette 234-β-EGFP) transcription can be detected on both sides when the cassette is in the permissive orientation but only on the telomeric side when the cassettes are in the nonpermissive orientation. (f) Summary of changes in levels of transcription in the flanking sequences when cassettes β-EGFP, 234, and 234-β-EGFP are inserted at RL4. Changes in transcript levels are relative to levels of transcription at RL4 in the presence of cassette EGFP. Grey bars represent changes (*n*-fold) in transcription levels at (from left to right) 20, 4, and 0.1 kb from the integration site on the centromeric side. Black bars represent changes (*n*-fold) in transcription levels at (from left to right) 0.1, 3, 14, and 20 kb from the integration site on the telomeric side.

induced by insertion of the LCR. The mini-LCR must interact with an element in the flanking sequences, since silencing is orientation dependent. This putative mini-LCR-activated "silencer" could be either an unknown type of element, a late-firing replication origin (22), or a cryptic promoter whose activation leads to transcriptional interference. The preponderance of our observations favor the latter hypothesis, since the orientation of the β-EGFP gene appears critical and since we detected transcription in the flanking sequences.

We propose the following model to explain the results in this and previous studies. Prior to insertion of any cassette, the RL4 site is transcriptionally inactive and replicates late in S phase; insertion of the CMV-HYTK cassette in the permissive orientation leads to a switch to early replication (22). Insertion of cassette β-EGFP (which lacks the mini-LCR) leads to low-level expression in both orientations; silencing does not occur, because no transcription is induced in the flanking sequences. When cassette 234-β-EGFP is inserted at RL4, the presence of the mini-LCR initially induces transcription on both sides of the site. However, transcriptional interference leads to the silencing of both the β-globin promoter and a cryptic promoter on the centromeric side. Hence, when nuclear transcription is assessed 3 months postintegration, no transcription is detected on the centromeric side. In the permissive orientation, the silencing of cassette 234-β-EGFP does not occur, because the β-EGFP gene faces towards the telomeric side and does not cause transcriptional interference. The transcriptional interference in the nonpermissive orientation leads to late replication, DNA methylation, and loss of histone acetylation and H3-K4 methylation. Together, these epigenetic changes condense the chromatin structure and completely silence the reporter.

The finding that the mini-LCR induces transcription on both sides of the site when it is integrated alone supports the hypothesis that there are cryptic promoters in the flanking sequences. As discussed earlier, there are no known genes in the vicinity of the RL4 site; we therefore propose that the promoter induced by the mini-LCR is within a repetitive sequence. Many potential cryptic promoters are present in the vicinity of RL4. Further experiments in which nuclear transcription in the flanking sequences is measured shortly after integration will be required to test this model and identify this putative promoter.

A recent report that transcriptional interference caused by a naturally occurring deletion induces silencing and methylation of the α2-globin gene supports this model (41). Recent experiments with *Arabidopsis* and yeast, suggesting that silenced chromatin structures can be initiated by the local production of double-stranded RNA, provide a potential mechanism for the initiation of silencing (20, 39, 42).

The most striking observation in this study is that silencing at RL4 is an active process that requires the mini-LCR. Previous studies had shown at other sites of integration that HS2 of the LCR can suppress silencing by preventing localization of the transgene close to heterochromatin (16). While the mini-LCR causes rather than prevents silencing at RL4, our results do not exclude the possibility that the mini-LCR acts by changing the nuclear localization of the locus. A possible mechanism for silencing at RL4 is that interactions between the LCR and the putative cryptic promoter localize the transgene in a nuclear compartment where the β-globin promoter cannot be transcribed.

In summary, we have shown that rapid silencing and chromatin condensation associated with DNA methylation, deacetylation of histone H3 and H4, methylation of lysine 4 of histone H3, and replication late in S phase at RL4 require the combined presence of a transcription unit facing toward the centromeric side of the integration site and of DNase I HS2 or HS3, but not of HS4, of the LCR. We have also shown that the LCR can induce transcription on both sides of the integration sites but that when silencing of cassette 234-β-EGFP occurs in the nonpermissive orientation, no nuclear transcription on the centromeric side can be detected 3 months postintegration.

Transcriptional interference: a general trigger for silencing? Silencing of transgenes in cell culture and transgenic mice occurs frequently. A generally admitted model for transgene variegation of expression is that integration next to heterochromatin, associated with variable spreading of the condensed chromatin state, leads to silencing of the transgenic sequences (14, 19). While this model explains silencing at some sites of integration, it does not appear to be true in general for mammalian cells, since transgenes inserted into pericentromeric heterochromatin are sometimes expressed at high levels (25, 35), since the same transgene inserted at the same position can be expressed in one orientation but silenced in the other in fly, mouse, and cultured cells (1, 16, 34), and since retroviral vectors which have a strong propensity for integration into active chromatin and within genes (44) are often silenced (32).

Because genes and repetitive sequences are two of the most common elements in the genome, transgenes are likely to integrate near a promoter. We propose that transcriptional interference triggered either by transgene-induced activation of promoters or by insertion in already active transcription units may be the major cause of transgene silencing. Our previous report that an effect of transcriptional interference on transcription can be detected by a bioinformatics approach on the whole genome (7), and our demonstration in collaboration with the Fiering lab that transcriptional interference has dramatic effects on gene expression support this concept (10).

This model predicts that, paradoxically, the addition of strong enhancers to expression cassettes can induce their silencing and may also induce the silencing of genes in the flanking sequences near the site of integration. This has profound implications for the design of gene therapy cassettes.

The cassettes that do not silence in the nonpermissive orientation are invariably expressed at lower levels than when they are in the permissive orientation. A similar orientation-dependent decrease in expression can also be observed at RL5 and RL6, two sites of integration similar to RL4 (16; Eric E. Bouhassira, unpublished results). The mechanism for this decrease in expression is unknown. An interesting possibility would be that this difference in expression level is linked to the location of the closest origin of replication and therefore to whether the coding strand is replicated as a leading or lagging strand.

LCR-mediated silencing and regulation of the endogenous β-globin locus. As discussed in the introduction, the β-globin cluster is strongly developmentally regulated. Several studies have shown that the human β-globin gene has the potential to be active in embryonic erythroid cells and that its silencing

depends on *cis*-acting elements present in the construct and on the chromatin context (3, 9, 23, 33, 36).

Previous reports by Tanimoto et al. (40) and by our laboratory (1) that in vivo inversions of the LCR in large yeast artificial chromosome constructs has dramatic effects on gene expression at several chromosomal positions, suggests that this chromatin context-dependent silencing activity of the LCR is not specific for the locus RL4 or for small LCR derivatives. On the contrary, this activity might be a property of the complete LCR that occurs at many sites of integration. Our previous report that an 11-kb construct encompassing the β-EGFP reporter driven by a 9-kb genomic fragment containing HS2, HS3, and HS4 of the LCR with their normal spacing and flanking sequences also silenced in the nonpermissive orientation at RL4 support this hypothesis (16). The recent report that the LCR is required to keep the β-globin locus late replicating in nonerythroid cells (38) suggests that the finding that the LCR is required for the switch from early to late replication associated with silencing at RL4 reflects a general property of the LCR rather than an idiosyncrasy of the RL4 site.

We propose that the novel, chromatin context-dependent silencing activity of the mini-LCR on the β-globin promoter that we have demonstrated contributes to silencing of globin genes at their native locus during development or in nonerythroid cells and that transcriptional interference contributes to this regulation. The detection of transcripts initiating in repetitive sequences and running through the globin locus by several groups (2, 24, 30) supports this hypothesis.

ACKNOWLEDGMENTS

We thank Carl Schildkraut, Steve Fiering, and Mark Lalande for helpful discussions.

E.E.B. is supported by NIH grants DK56845, HL55435, and DK061799. M.I.A. is supported by the NIH intramural program. H.F. is supported by an NCI Intramural Research Award.

REFERENCES

1. **Alami, R., J. M. Greally, K. Tanimoto, S. Hwang, Y.-Q. Feng, J. D. Engel, S. Fiering, and E. E. Bouhassira.** 2000. Beta-globin YAC transgenes exhibit uniform expression levels but position effect variegation in mice. Hum. Mol. Genet. **9:**631–636.
2. **Ashe, H. L., J. Monks, M. Wijgerde, P. Fraser, and N. J. Proudfoot.** 1997. Intergenic transcription and transinduction of the human beta-globin locus. Genes Dev. **11:**2494–2509.
3. **Behringer, R. R., T. M. Ryan, R. D. Palmiter, R. L. Brinster, and T. M. Townes.** 1990. Human gamma- to beta-globin gene switching in transgenic mice. Genes Dev. **4:**380–389.
4. **Bender, M. A., M. Bulger, J. Close, and M. Groudine.** 2000. Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. Mol. Cell **5:**387–393.
5. **Bulger, M., T. Sawado, D. Schubeler, and M. Groudine.** 2002. ChIPs of the beta-globin locus: unraveling gene regulation within an active domain. Curr. Opin. Genet. Dev. **12:**170–177.
6. **Carter, D., L. Chakalova, C. S. Osborne, Y. F. Dai, and P. Fraser.** 2002. Long-range chromatin regulatory interactions in vivo. Nat. Genet. **32:**623–626.
7. **Chiaramonte, F., W. Miller, and E. E. Bouhassira.** 2003. Gene length and proximity to neighbors affect genome-wide expression levels. Genome Res. **13:**2602–2608.
8. **Engel, J. D., and K. Tanimoto.** 2000. Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. Cell **100:**499–502.
9. **Enver, T., N. Raich, A. J. Ebens, T. Papayannopoulou, and F. Costantini.** 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. Nature **344:**309–313.
10. **Eszterhas, S. K., E. E. Bouhassira, D. I. Martin, and S. Fiering.** 2002. Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. Mol. Cell. Biol. **22:**469–479.
11. **Felsenfeld, G., and M. Groudine.** 2003. Controlling the double helix. Nature **421:**448–453.
12. **Feng, Y. Q., J. Seibler, R. Alami, A. Eisen, S. N. Fiering, and E. E. Bouhassira.** 1999. Site-specific chromosomal integration in mammalian cells: highly efficient CRE recombinase-mediated cassette exchange. J. Mol. Biol. **292:**779–785.
13. **Feng, Y. Q., M. C. Lorincz, S. Fiering, J. M. Greally, and E. E. Bouhassira.** 2001. Position effects are influenced by the orientation of a transgene with respect to flanking chromatin. Mol. Cell. Biol. **21:**298–309.
14. **Festenstein, R., and D. Kioussis.** 2000. Locus control regions and epigenetic chromatin modifiers. Curr. Opin. Genet. Dev. **10:**199–203.
15. **Forrester, W. C., E. Epner, M. C. Driscoll, T. Enver, M. Brice, T. Papayannopoulou, and M. Groudine.** 1990. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. Genes Dev. **4:**1637–1649.
16. **Francastel, C., M. C. Walters, M. Groudine, and D. I. Martin.** 1999. A functional enhancer suppresses silencing of a transgene and prevents its localization close to centromeric heterochromatin. Cell **99:**259–269.
17. **Grosveld, F.** 1999. Activation by locus control regions? Curr. Opin. Genet. Dev. **9:**152–157.
18. **Grosveld, F., G. B. van Assendelft, D. R. Greaves, and G. Kollias.** 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell **51:**975–985.
19. **Henikoff, S.** 1992. Position effect and related phenomena. Curr. Opin. Genet. Dev. **2:**907–912.
20. **Jenuwein, T.** 2002. Molecular biology. An RNA-guided pathway for the epigenome. Science **297:**2215–2218.
21. **Kaufman, R. M., C. T. Pham, and T. J. Ley.** 1999. Transgenic analysis of a 100-kb human beta-globin cluster-containing DNA fragment propagated as a bacterial artificial chromosome. Blood **94:**3178–3184.
22. **Lin, C. M., H. Fu, M. Martinovsky, E. E. Bouhassira, and M. I. Aladjem.** 2003. Dynamic alterations of replication timing in mammalian cells. Curr. Biol. **13:**1019–1028.
23. **Lloyd, J. A., J. M. Krakowsky, S. C. Crable, and J. B. Lingrel.** 1992. Human γ- to β-globin gene switching using a mini construct in transgenic mice. Mol. Cell. Biol. **12:**1561–1567.
24. **Long, Q., C. Bengra, C. Li, F. Kutlar, and D. Tuan.** 1998. A long terminal repeat of the human endogenous retrovirus ERV-9 is located in the 5′ boundary area of the human beta-globin locus control region. Genomics **54:**542–555.
25. **Lundgren, M., C. M. Chow, P. Sabbattini, A. Georgiou, S. Minaee, and N. Dillon.** 2000. Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell **103:**733–743.
26. **McBurney, M. W., L. C. Sutherland, C. N. Adra, B. Leclair, M. A. Rudnicki, and K. Jardine.** 1991. The mouse Pgk-1 gene promoter contains an upstream activator sequence. Nucleic Acids Res. **19:**5755–5761.
27. **McMorrow, T., A. van den Wijngaard, A. Wollenschlaeger, M. van de Coreut, M. Monkhorst, T. Trimborn, P. Fraser, M. van Lohuizen, T. Jenuwein, M. Djabali, S. Philipsen, F. Grosveld, and E. Milot.** 2000. Activation of the beta globin locus by transcription factors and chromatin modifiers. EMBO J. **19:**4986–4996.
28. **Molete, J. M., H. Petrykowska, E. E. Bouhassira, Y. Q. Feng, W. Miller, and R. C. Hardison.** 2001. Sequences flanking hypersensitive sites of the beta-globin locus control region are required for synergistic enhancement. Mol. Cell. Biol. **21:**2969–2980.
29. **Palstra, R. J., B. Tolhuis, E. Splinter, R. Nijmeijer, F. Grosveld, and W. de Laat.** 2003. The beta-globin nuclear compartment in development and erythroid differentiation. Nat. Genet. **35:**190–194.
30. **Plant, K. E., S. J. Routledge, and N. J. Proudfoot.** 2001. Intergenic transcription in the human beta-globin gene cluster. Mol. Cell. Biol. **21:**6507–6514.
31. **Porcu, S., M. Kitamura, E. Witkowska, Z. Zhang, A. Mutero, C. Lin, J. Chang, and K. M. L. Gaensler.** 1997. The human beta globin locus introduced by YAC transfer exhibits a specific and reproducible pattern of developmental regulation in transgenic mice. Blood **90:**4602–4609.
32. **Rivella, S., and M. Sadelain.** 1998. Genetic treatment of severe hemoglobinopathies: the combat against transgene variegation and transgene silencing. Semin. Hematol. **35:**112–125.
33. **Roberts, N. A., J. A. Sloanestanley, J. A. Sharpe, S. J. Stanworth, and W. G. Wood.** 1997. Globin gene switching in transgenic mice carrying HS2-globin gene constructs. Blood **89:**713–723.
34. **Sabl, J. F., and S. Henikoff.** 1996. Copy number and orientation determine the susceptibility of a gene to silencing by nearby heterochromatin in *Drosophila*. Genetics **142:**447–458.
35. **Saffery, R., H. Sumer, S. Hassan, L. H. Wong, J. M. Craig, K. Todokoro, M. Anderson, A. Stafford, and K. H. Choo.** 2003. Transcription within a functional human centromere. Mol. Cell **12:**509–516.
36. **Sargent, T. G., and J. A. Lloyd.** 2001. The human gamma-globin TATA and CACCC elements have key, distinct roles in suppressing beta-globin gene expression in embryonic/fetal development. J. Biol. Chem. **276:**41817–41824.
37. **Siebert, P. D., A. Chenchik, D. E. Kellogg, K. A. Lukyanov, and S. A. Lukyanov.** 1995. An improved PCR method for walking in uncloned genomic DNA. Nucleic Acids Res. **23:**1087–1088.

38. **Simon, I., T. Tenzen, R. Mostoslavsky, E. Fibach, L. Lande, E. Milot, J. Gribnau, F. Grosveld, P. Fraser, and H. Cedar.** 2001. Developmental regulation of DNA replication timing at the human beta globin locus. EMBO J. **20:**6150–6157.
39. **Stevenson, D. S., and P. Jarvis.** 2003. Chromatin silencing: RNA in the driving seat. Curr. Biol. **13:**R13–R15.
40. **Tanimoto, K., Q. Liu, J. Bungert, and J. D. Engel.** 1999. Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. Nature **398:**344–348.
41. **Tufarelli, C., J. A. Stanley, D. Garrick, J. A. Sharpe, H. Ayyub, W. G. Wood, and D. R. Higgs.** 2003. Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. Nat. Genet. **34:**157–165.
42. **Volpe, T. A., C. Kidner, I. M. Hall, G. Teng, S. I. Grewal, and R. A. Martienssen.** 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science **297:**1833–1837.
43. **Walters, M. C., W. Magis, S. Fiering, J. Eidemiller, D. Scalzo, M. Groudine, and D. I. K. Martin.** 1996. Transcriptional enhancers act in *cis* to suppress position-effect variegation. Genes Dev. **10:**185–195.
44. **Wu, X., Y. Li, B. Crise, and S. M. Burgess.** 2003. Transcription start regions in the human genome are favored targets for MLV integration. Science **300:**1749–1751.