

Human β -Globin Locus Control Region HS5 Contains CTCF- and Developmental Stage-Dependent Enhancer-Blocking Activity in Erythroid Cells

Keiji Tanimoto,^{1*} Akiko Sugiura,¹ Akane Omori,¹ Gary Felsenfeld,² James Douglas Engel,³ and Akiyoshi Fukamizu¹

Center for Tsukuba Advanced Research Alliance, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan¹; National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0540²; and University of Michigan Medical School, Ann Arbor, Michigan 48109-0616³

Received 14 April 2003/Returned for modification 7 July 2003/Accepted 10 September 2003

The human β -globin locus contains five developmentally regulated β -type globin genes. All five genes depend on the locus control region (LCR), located at the 5' end of the locus, for abundant globin gene transcription. The LCR is composed of five DNase I-hypersensitive sites (HSs), at least a subset of which appear to cooperate to form a holocomplex in activating genes within the locus. We previously tested the requirement for proper LCR polarity by inverting it in human β -globin yeast artificial chromosome transgenic mice and observed reduced expression of all the β -type globin genes regardless of developmental stage. This phenotype clearly demonstrated an orientation-dependent activity of the LCR, although the mechanistic basis for the observed activity was obscure. Here, we describe genetic evidence demonstrating that human HS5 includes enhancer-blocking (insulator) activity that is both CTCF and developmental stage dependent. Curiously, we also observed an attenuating activity in HS5 that was specific to the ϵ -globin gene at the primitive stage and was independent of the HS5 CTCF binding site. These observations demonstrate that the phenotype observed in the LCR-inverted locus was in part attributable to placing the HS5 insulator between the LCR HS enhancers (HS1 to HS4) and the promoter of the β -globin gene.

During the normal process of mammalian development, proper temporal and spatial expression of genetic information must be achieved. To this end, gene expression must be tightly regulated through *cis*-DNA elements: promoters, enhancers, and silencers. After the human genome had been completely sequenced, it became clear that the average size of intergenic regions is roughly 100 kbp. However, substantial evidence argues that enhancers can modulate promoter activity from very long distances, exceeding 100 kbp (19, 20). In such a circumstance, one can envision an activity that protects a locus from neighboring gene regulatory elements to prevent improper gene regulation. Indeed, such an activity has been described: DNA insulators were first identified in *Drosophila melanogaster*, and similar activities were subsequently found in mammals.

Insulators can protect a locus by two distinguishable means; one is a chromatin barrier function, while a second is referred to as enhancer-blocking activity. The most extensively characterized vertebrate insulator was originally identified in the chicken β -globin LCR, which consists of four DNase I-hypersensitive (HS) sites; the 5'-most is HS4. After stable transformation of K562 cells, a 1.2-kbp DNA fragment containing HS4 was found to interfere with enhancer/promoter interactions, but only when it was placed between them (10). Further analysis revealed two separable categories of activity within this 1.2-kbp DNA fragment, one of which protects stably integrated

transgenes from silencing after long-term cell culture (28). This property of an insulator is referred to as its chromatin barrier activity.

Chung et al. developed an enhancer-blocking assay to refine the position of a 250-bp core region within the 1.2-kbp chicken HS4 insulator fragment that still retained both activities (9). They then mapped four footprints (FI to FIV) within the core region to identify DNA binding proteins (2). Following column and subsequent DNA affinity purification of chicken red blood cell nuclear extracts, they found that transcription factor CTCF, known for its ubiquitous expression and both transcriptional activator and repressor properties (23), bound to the FII region. Subsequently, it was reported that the chromatin barrier activity in the 250-bp core of chicken HS4 does not require CTCF binding to the FII sequences and that this activity was clearly distinguishable from enhancer blocking (28). Despite a considerable number of observations analyzing insulator function, no conclusive evidence has shown how an insulator works.

The human β -globin genes are organized within a 70-kbp region, with the embryonic ϵ -globin gene located most 5', followed by the two fetal γ -globin genes ($G\gamma$ and $A\gamma$), while the adult δ - and β -globin genes are at the 3' end of the locus (Fig. 1A) (31). High-level human β -globin transcription is regulated by the LCR, consisting of erythroid cell-specific HS1 to HS4 and developmentally ubiquitous HS5. We have previously shown that inverting the LCR (HS1 to HS5) in human β -globin yeast artificial chromosome (YAC) transgenic mice significantly diminished expression of all the β -like globin genes (*LCR-inv*) (32). We proposed several interpretations for this observed phenotype, including the possibility that higher-order chromatin architecture might have been disrupted by the in-

* Corresponding author. Mailing address: Center for Tsukuba Advanced Research Alliance, Institute of Applied Biochemistry, University of Tsukuba, Tennoudai 1-1-1, Tsukuba, Ibaraki 305-8577, Japan. Phone and fax: 81 (29) 853-6070. E-mail: keiji@tara.tsukuba.ac.jp.

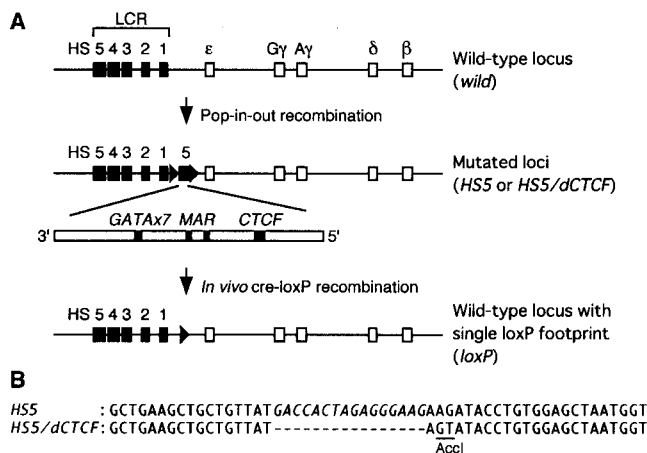


FIG. 1. Schematic representation of the experimental system. (A) An extra copy of the HS5 fragment (2.6-kbp, inverted orientation, middle) was floxed (solid triangles) and introduced between HS1 and the ϵ -globin gene in human β -globin YAC (A201F4.3; wild type, top) by homologous recombination in *S. cerevisiae*. Putative cis-DNA elements within LCR-HS5 are schematically represented (*GATAx7*, *MAR*, and *CTCF*). The YAC DNA was purified and used to generate transgenic mice (TgM). A single-copy, intact YAC transgenic mouse (mutant) was then mated with Cre-expressing transgenic mice to recreate the wild-type locus by deleting the ectopic HS5 fragment (wild-type with single *loxP* footprint) (B) Wild-type (*HS5*, upper) and mutant (*HS5/dCTCF*, lower) HS5 sequences surrounding the putative CTCF binding site are in italics. To facilitate screening for proper homologous recombination in *S. cerevisiae*, two nucleotides (underlined) outside the CTCF consensus motif were also mutated, which created an *AccI* restriction enzyme site.

version, since it appears that individual HS sites structurally and functionally constitute a holocomplex, which activates the globin genes as a single unit (6, 21). If this indeed is the case, dividing the holocomplex into two domains, HS1 to HS5 and HS6-7, might attenuate LCR activity (5). Another possible interpretation was related to intergenic transcription (1), wherein transcripts originating either 5' to the LCR (24) or within the HS2/HS3 region (20, 35) were always oriented toward the β -type globin genes (20, 29).

It was hypothesized that intergenic transcription might somehow be involved in the chromatin-opening activity of the LCR (12). In this scenario, inversion of the LCR would direct HS2/HS3-initiated intergenic transcription (but not ERV-9-initiated transcription) (24) to the wrong strand and thus could negatively affect globin gene transcription. A third interpretation of the phenotype was that the putative insulator activity in HS5 (22) interfered with LCR enhancer (HS1 to HS4)-promoter interactions. Since chicken HS4 bears strong insulator activity, it would not be surprising if its ortholog (human HS5) bore similar activity. In fact, Farrell et al. searched for CTCF binding sites in the human and mouse HS5 regions and found reasonably conserved sites (14). These motifs bound the CTCF factor in vitro and showed enhancer-blocking activity in cell culture assays.

In order to evaluate the insulator hypothesis and to clearly distinguish it from other possible interpretations (32), we established multiple transgenic mouse lines bearing a mutant β -globin locus on a YAC and analyzed globin gene transcription. Here we demonstrate that human HS5 contains CTCF-

dependent enhancer-blocking activity in vivo and that this activity is developmental stage specific. We also demonstrate that there may be an additional activity in HS5 that specifically attenuates ϵ -gene transcription independent of the CTCF binding site.

MATERIALS AND METHODS

Targeting vectors for homologous recombination in *Saccharomyces cerevisiae*. A backbone plasmid for targeting constructs, pHS1/*loxPw+*, was generated as follows. Human β -globin HS1 DNA fragment (from nucleotides 13299 to 14250; HUMHBB, GenBank) was PCR amplified with the primer set HS1-HS5S (5'-ATTTCTCGAGCAACTAACTCATG-3') and HS1-3A (5'-GAATCTCATGCTAGAAATTTTG-3'), digested with *XhoI* and *XbaI* restriction enzymes, subcloned into *XhoI*- and *XbaI*-digested pRS306 yeast targeting vector, and verified by sequencing (pHS1). Artificially introduced *XhoI* or native *XbaI* sites on each oligonucleotide are in italics.

HS1-LX1-5S (5'-AGCTGGATCCATAACTTCGTATAGCATACATTATACGAAGTTATC-3') and HS1-LX1-3A (5'-AGCTGATAACTTCGTATAATGTATGCTATACGAAGT TATGGATCC-3') oligonucleotides were annealed, which makes *HindIII*-compatible (but not redigestible) ends and ligated to *HindIII* (at nucleotide position 13769)-digested pHS1 in the forward (*loxP* sense) orientation. The resultant plasmid was then digested with *BamHI* at a site that was artificially introduced into the *loxP* oligonucleotide (italics) and then ligated to the next *loxP* oligonucleotides: HS1-LX2-5S (5'-GATCTTGAAATTCATAAC TTCGTATAGC ATACATTATACGAAGTTATG-3') and HS1-LX2-3A (5'-G ATCCATAACTTCGTATAATGTATGCTATACGAAGTTATGAAATTCAA-3') after annealing. This second oligonucleotide has *BglII*- and *BamHI*-compatible ends, and a *BamHI* site is generated between the two *loxP* sites after ligation into the *BamHI* site in the forward orientation (this plasmid is hereafter referred to as pHS1/*loxPw+*). For convenience, we digested pHS1/*loxPw+* with *BamHI*, blunt ended it, and ligated to a phosphorylated *HindIII* linker (5'-CAAGCTTG-3'), which created a unique *HindIII* site between the *BamHI* sites in the tandemly duplicated *loxP* sites.

LCR-HS5 DNA was excised as a 2.6-kbp *HindIII* fragment (from nucleotides 1 to 2647; HUMBGLIBC, GenBank) (36) from the human β -globin YAC (A201F4.3) and subcloned into *HindIII*-digested pHS1/*loxPw+* to generate pHS1/*loxPw+*/*HS5*. Deletion of the putative CTCF binding motif from within the 2.6-kbp HS5 fragment was done by PCR-directed mutagenesis (Fig. 1B), and successful mutagenesis was verified by sequencing. This mutant HS5 was used for constructing pHS1/*loxPw+*/*HS5dCTCF*. The targeting plasmid DNAs were linearized by digestion with *SpeI* (at nucleotide position 13670) and used for human β -globin YAC mutagenesis by homologous recombination (6).

Transgenic mice. The generation and structural analysis of human β -globin YAC transgenic mice has been described elsewhere (34). Removal of the ectopic HS5 copy was conducted by mating intact, single-copy human β -globin YAC transgenic mice with transgenic mice ubiquitously expressing Cre recombinase (37).

Semiquantitative RT-PCR analysis. RNA expression analysis of human β -like globin transgenes has been described elsewhere (33). In short, total RNA from two individuals from each transgenic line was extracted from yolk sac (9.5 days post coitum), fetal liver (14.5 days post coitum) or anemic adult spleen with Isogen (Nippon Gene). First-strand cDNA was synthesized from 2.5 μ g of RNA with Moloney murine reverse transcriptase (Gibco-BRL). A twentieth of the reaction was used for PCR amplification with the following parameters: 94°C for 30 s, 58°C for 1 min, and 72°C for 1 min. Cycle numbers used for PCR analyses, which are within the linear amplification range, were as follows: 12 for both the β and α cDNAs in spleen; 18 and 12 for the γ and β/α cDNAs, respectively, in the liver; and 18 and 12 for the ϵ and γ/α cDNAs, respectively, in the yolk sac. An aliquot of each PCR was electrophoresed on 8% polyacrylamide gels, dried, and subjected to X-ray autoradiography and phosphorimaging for quantitative analysis.

RESULTS

Generation of YAC transgenic mice. Since several activities which could potentially attenuate globin gene transcription have been reported within the vicinity of HS5, we used a 2.6-kbp DNA fragment to represent HS5 (*HS5*; Fig. 1A, middle). This fragment includes seven tandem copies of GATA

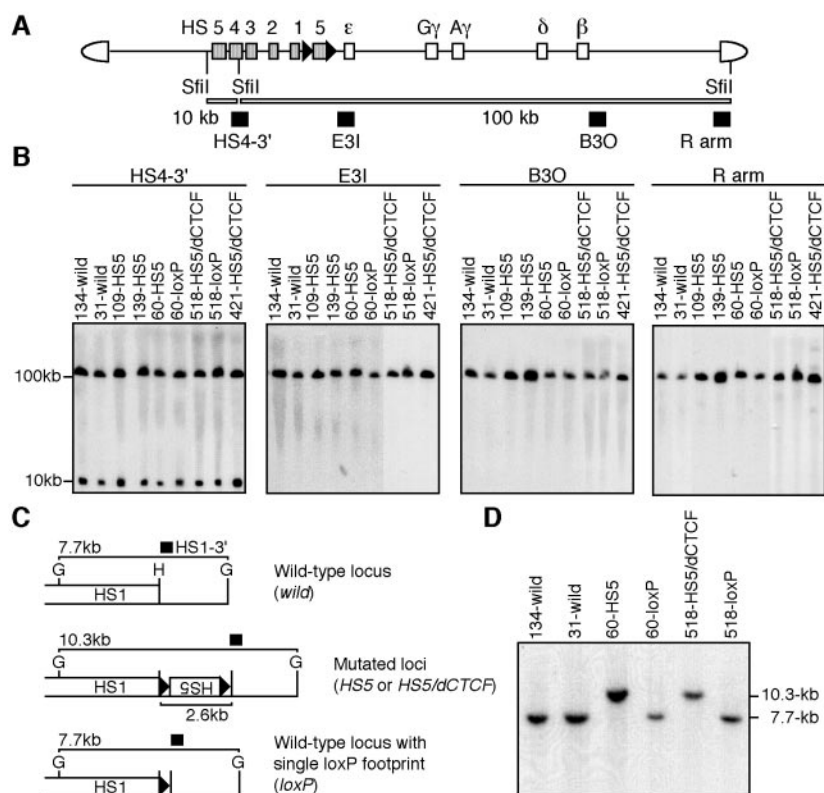


FIG. 2. Structural analysis of the β -globin YAC in transgenic mice. (A) Schematic representation of the mutant human β -globin YAC. The positions of the β -like globin genes are shown relative to the LCR. *Sfi*I restriction enzyme sites are located 5' to HS5, between HS4 and HS3, and in the right arm of the YAC. Probes (solid rectangles) used for long-range fragment analysis and expected restriction enzyme fragments with their sizes are shown. (B) Long-range structural analysis of transgenes. The whole β -globin locus is contained within two *Sfi*I fragments (10 and 100 kbp, as in A). DNA from thymus cells was digested with *Sfi*I in agarose plugs, separated by pulsed-field gel electrophoresis, and hybridized separately to probes (shown in A) from the β -globin locus or from the right YAC vector arm. (C) Schematic representation of the transgene locus around HS1. Cre-*loxP*-mediated HS5 deletion removes the 2.6-kbp insert from the mutated loci (middle), which creates a 7.7-kbp *Bgl*II fragment in the locus (bottom). Solid arrowheads, *loxP* sequences. G, *Bgl*II; H, *Hind*III. (D) Tail DNAs from wild-type (lines 134 and 31), mutant (lines 60-HS5 and 518-*HS5/dCTCF*), and *loxP* footprint (lines 60- and 518-*loxP*) transgenic mouse lines was digested with *Bgl*II, separated on an agarose gel, and transferred to a nylon membrane. Hybridization was performed with the HS1-3' probe (solid rectangle in C). The bands representative of wild-type (7.7 kbp), mutated (10.3 kbp), and footprint (7.7 kbp) loci are indicated.

motifs (*GATAx7* [27]), a matrix attachment region (*MAR* [36]), and putative enhancer-blocking activity, i.e., a transcription factor CTCF binding site (*CTCF* [14]). At the same time, to test the contribution of *CTCF* to the insulator activity within HS5, we deleted the CTCF binding site from the 2.6-kbp DNA fragment and used it for YAC mutagenesis (*HS5/dCTCF* for deleted CTCF; Fig. 1B). Both the 2.6-kbp *HS5* and *HS5/dCTCF* mutant fragments were inserted 3' to HS1 in reverse orientation to mimic the situation (32) in which the whole LCR was inverted (thereby localizing HS5 between LCR HS1 to HS4 and the globin gene promoters in the antisense orientation).

We flanked both ectopic HS5 fragments by *loxP* sites so that they could be removed *in vivo* to regenerate a wild-type locus (*loxP*; Fig. 1A, bottom). We mutated a 150-kbp human β -globin YAC (A201F4.3; Fig. 1A, top) by homologous recombination in yeast cells (6, 17), purified the YAC DNA by standard methods (6), and used the recovered DNA for microinjection into fertilized mouse embryos. Tail DNA prepared from founder offspring (F_0) was screened first with PCR primers

specific for the human β -globin gene and then by Southern blot.

Transgene structural analysis. We generated independent transgenic lines for YAC *HS5* (lines 60, 109, and 139) and YAC *HS5/dCTCF* (lines 421 and 518) (Fig. 2). Combined end fragment and copy number analyses (with a fragment from the endogenous angiotensinogen locus as an internal control) (11) of thymus DNA from F_1 animals revealed that three of the lines (60, 421, and 518) carried single-copy YAC transgenes and that the others (109 and 139) bore two copies (Fig. 2B and not shown).

High-molecular-weight DNA recovered from thymuses was embedded in agarose plugs and then analyzed by pulsed-field gel electrophoresis followed by Southern blot hybridization. We used probes spanning the locus and the YAC vector arms, HS4-3', ϵ -globin gene (E3I), β -globin (B3O) gene, and the right arm (R arm) (Fig. 2A) to detect either the 10- or 100-kbp *Sfi*I restriction fragments predicted from integration of an intact human β -globin gene locus. All of the probes detected bands of the expected sizes in all the lines (Fig. 2B), indicating

that each transgenic mouse line carried intact, unfragmented copies of the transgenes.

Following these structural analyses, we focused on two transgenic lines (60 and 518), both of which carried single-copy, intact transgenes. F_1 animals from both lines were subjected to Cre-mediated HS5 deletion by mating to ubiquitously expressed Cre recombinase transgenic mice to remove the ectopic HS5 fragments. Offspring from this mating were analyzed for proper fragment excision by Southern blot analysis (Fig. 2C and D) and those that had undergone Cre-mediated recombination (Cre- F_0) were again mated with wild-type animals to remove the Cre recombinase allele. transgenic mice of fixed genotype (Cre- F_1) were analyzed for transgene RNA expression.

Expression analysis in erythroid cells of YAC HS5 transgenic mice. Two animals (1 to 2 months old) from each YAC HS5 transgenic mouse line were made anemic by phenylhydrazine injection 5 days prior to analysis, when spleen samples were collected for RNA extraction. Semiquantitative RT-PCR was performed to compare the levels of globin mRNA from wild-type (lines 134- and 31-*wild*, both single-copy) (32) and mutant (109-, 139- and 60-*HS5*) transgenic mice. As shown in Fig. 3A, human β -globin gene expression in the mutant lines was consistently 50 to 70% lower compared with the wild-type transgenic mouse lines. To confirm that the phenotype in YAC HS5 transgenic mice was due to insertion of the HS5-containing 2.6-kbp DNA fragment, we regenerated a wild-type transgenic mouse line (60-*loxP*) from the mutant parental line (60-*HS5*) by *in vivo* Cre-*loxP* recombination. We found that excision of the 2.6-kbp HS5 fragment resulted in increased expression in this Cre-treated line, close to the levels observed in control wild-type YAC transgenic mouse lines (134- and 31-*wild*).

Next, we analyzed expression of the human γ - and β -globin genes in the fetal liver (14.5 days post coitum; Fig. 3B). RT-PCR analysis revealed from 50% to more than 90% reduction in the expression of the β - and γ -globin genes, respectively, in the HS5 transgenic mice (109- and 60-*HS5*). Once again, this reduced expression was ameliorated after site-specific deletion of the ectopic HS5 fragment (60-*loxP*).

We then analyzed expression of the human ϵ - and γ -globin genes in the yolk sac (9.5 days post coitum; Fig. 3C). Surprisingly, γ -globin expression did not differ significantly between wild-type (*wild*) and mutant (*HS5*) animals at this stage of erythropoiesis (Fig. 3C, bottom). Since the neighboring ϵ -globin expression was severely affected (>90% reduction) in the same samples (Fig. 3C, top), the observed phenotype is gene specific. Furthermore, γ -globin expression was severely attenuated in definitive erythroid cells (Fig. 3B), suggesting that an activity within HS5 may also confer developmental-stage specificity.

Expression analysis in erythroid cells of YAC HS5/dCTCF transgenic mice. As noted earlier, there are several *cis*-DNA elements that have been reported to lie within the 2.6-kbp fragment containing human LCR HS5, any of which could potentially attenuate the expression of *cis*-linked genes (14, 27, 36). In previous work (32), we inserted an ectopic ϵ -globin gene 5' to HS5 and observed virtually no expression of this gene, while the resident copy of ϵ in its normal position within the locus was expressed at normal levels. We interpreted these

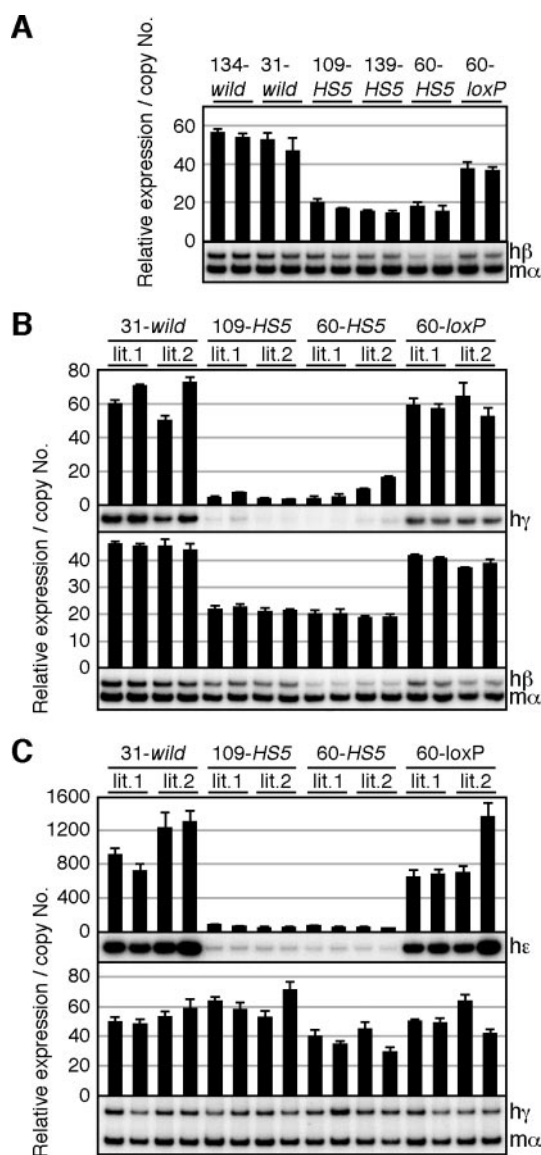


FIG. 3. RT-PCR analysis of β -like globin expression in erythroid cells of HS5 YAC transgenic mice. (A) Total RNA was prepared from the spleens of 1- to 2-month-old anemic mice and subjected to cDNA synthesis by reverse transcriptase (RT). The relative expression levels of the β -like globin genes, after normalization to that of the endogenous mouse α -globin gene, were determined by RT-PCR analysis. Data were collected from two individuals for each line and the average and standard deviation, determined by three sets of PCRs, are graphically depicted. Representative results of RT-PCR for human β ($h\beta$) and mouse α ($m\alpha$) amplicons are shown below each panel. (B and C) Total RNA was prepared from the liver (E14.5) or yolk sacs (E9.5) of two fetuses in two litters (lit.1 and lit.2) derived from the intercross of male transgenic and female wild-type animals. Representative results of RT-PCR for human ϵ ($h\epsilon$), γ ($h\gamma$), β ($h\beta$), and mouse α ($m\alpha$) are shown below each panel.

data, in combination with those from the LCR inversion experiment, to mean that HS5 must be positioned between the ϵ -globin gene promoter and enhancer activities within the LCR (HS1 to HS4) for the gene to be suppressed. This strict contextual dependence was reminiscent of the enhancer-blocking activity of an insulator. Hence, we chose to test the contribu-

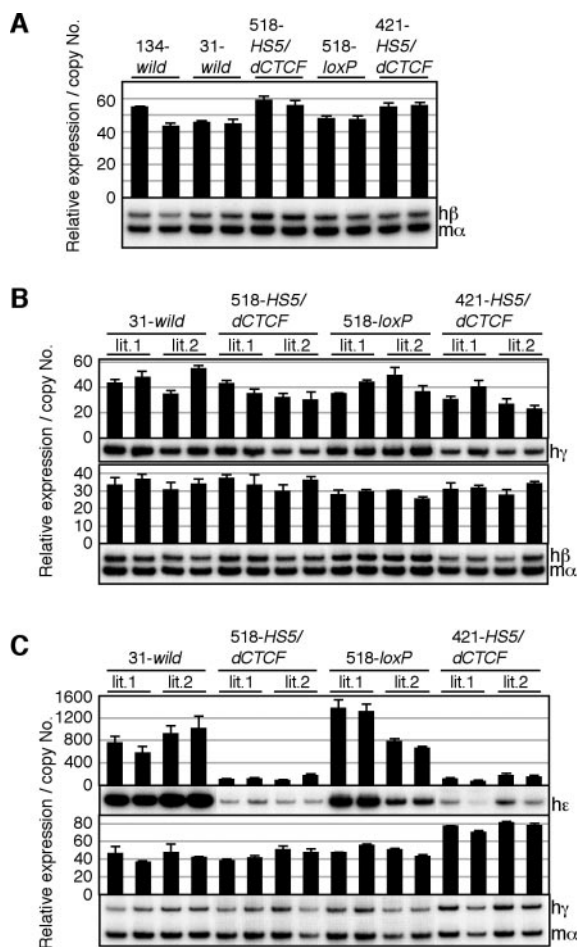


FIG. 4. RT-PCR analysis of β -like globin expression in erythroid cells of *HS5/dCTCF* YAC transgenic mice. Total RNA was prepared and analyzed as in Fig. 3.

tory activity of the CTCF binding motif (14) among all the elements present within the 2.6-kbp HS5 fragment to the phenotype observed in YAC *HS5* transgenic mice.

In definitive erythroid cells (adult spleen or fetal liver) of *HS5/dCTCF* transgenic mice, no significant change in expression of the β -like globin was observed in comparison to wild-type transgenic mice (134- and 31-*wild*; Fig. 4A and B), suggesting that HS5 was devoid of the enhancer-blocking activity observed in the YAC *HS5* transgenic mice at the same stages. To determine whether another suppressing activity might reside within HS5, the mutant HS5 fragment bearing the deleted CTCF binding site was removed by *in vivo* Cre-*loxP* recombination (Fig. 2C and D) to recreate pseudo-wild-type loci (518-*loxP*). Again, there was no significant difference in the expression of β -like globin genes before or after the removal of the fragment (518-*HS5/dCTCF* and 518-*loxP*; Fig. 4A and B), indicating that a putative CTCF binding motif in the human LCR-HS5 is solely responsible for the phenotype observed in definitive erythroid cells of *HS5* YAC transgenic mice (Fig. 3A and B). Surprisingly, however, the level of ϵ -globin gene expression in the embryonic yolk sac of *HS5/dCTCF* transgenic mice was significantly lower than that observed in wild-type control mice (Fig. 4C). This phenotype was not due to a posi-

tion-of-integration site effect, since the expression was completely restored after removing the HS5 fragment (Fig. 4C). The level of γ -globin expression was once again unaffected at this stage, as it was similarly unaffected in the *HS5* transgenic mice (Fig. 4C).

DISCUSSION

To test whether or not HS5 carries an activity that can insulate a gene(s) from environmental effects (referred to here as enhancer-blocking activity), Zafarana et al. (38) introduced a copy of the adult β -globin gene 5' to the LCR in transgenic mice. Inasmuch as the adult β -globin gene was expressed at normal levels in this mouse line, they concluded that HS5 does not bear enhancer-blocking activity *in vivo*. We also tested this hypothesis by generating YAC transgenic mice, inserting a marked ϵ -globin gene 5' to HS5 within the YAC (5' ϵ transgenic mice (32), and found very little expression of the marked ϵ -globin gene in primitive erythroid cells of this mouse. We interpreted these data to mean that HS5 might bear intrinsic insulator activity.

What might account for the apparent discrepancy in the two reports? One possibility is that the two simply differ in regulatory information, although it is difficult to judge precisely the limits of the transgene construct examined by Zafarana et al. from the information provided. In general, smaller transgenes tend to suffer more significantly from position-of-integration effects, even when the LCR is joined to an expression construct, which complicates quantitative comparisons among transgenic lines. Alternatively, it is clear from the results presented here that the adult β -globin gene (at the definitive stage) is less severely affected than is the ϵ -globin gene (at the primitive stage), although the phenotype observed in the 5' ϵ transgenic mice is not generated solely by CTCF binding to HS5 (Fig. 4C). Regardless, we conclude that HS5 embodies intrinsic enhancer-blocking activity, one of three formal possibilities presented in previous work (32).

We found one intriguing phenotypic difference between our previous *LCR-inv* (32) and the present *HS5* YAC transgenic mouse lines, which was manifested as a difference in the expression of γ -globin genes in the yolk sac. When HS5 was misplaced by inverting the whole LCR (*LCR-inv*), we found >50% reduction in the expression of these genes. However, we did not observe a similar reduction in the mutant transgenic mice created by inserting a single ectopic HS5 in the locus (*HS5*), suggesting that important characteristics of the LCR other than HS5 position and identity were altered during the process of LCR inversion. One such feature may be a unidirectional intergenic transcription from the HS2/HS3 domain. Routledge et al. (29) performed a detailed analysis of LCR transcription and reported discrete initiation sites, in both HS2 and HS3, directionally towards globin genes. If this directional intergenic transcription were somehow involved in LCR function, inversion of the LCR would result in aberrant gene regulation. However, there is another example showing that HS2 transcription is bidirectional (18). Hence, the underlying mechanism accounting for the phenotypic difference and its relevance to β -globin locus gene regulation is currently unknown.

Farrell et al. reported a comparison of CTCF-dependent enhancer-blocking activities of sequences from the 5'- and

3'-flanking regions of the chicken, mouse, and human β -globin gene loci (14). In their transfection-based enhancer-blocking assay, human and mouse 5' HS5s conferred moderate or almost negligible activities, respectively, compared with the chicken 5' HS4 insulator sequence. These data thus complement the earlier observation that deletion of 5' HS5 from the endogenous murine β -globin locus produces no apparent phenotype (3, 13). Furthermore, deletion of single HSs from the endogenous mouse locus generally generates much more modest phenotypes than the deletion of a corresponding HS from a human cosmid or YAC transgenic globin locus (summarized in 7), suggesting that the chromatin environment surrounding the endogenous murine β -globin locus may be more permissive for establishing an open erythroid chromatin configuration and thus may not require powerful insulator activity. While murine HS5 may have lost its enhancer-blocking function during the process of evolution or may never have possessed the function, human HS5 may act either to protect β -like globin genes from the surrounding regulatory elements for olfactory receptor genes (4, 15) and/or to protect olfactory receptor genes from β -globin LCR enhancement (14), as proposed for the chicken locus (26). This hypothesis can be tested directly.

While the expression of γ -globin genes in primitive-stage erythroid cells was minimally affected (Fig. 3C), the same genes were severely affected during definitive erythropoiesis (Fig. 3B). Additionally, when the degree of phenotypic severity in the *HS5* transgenic mice on two different β -like globin genes is compared at a given stage, one gene is generally affected far more than the other. Both observations imply that the enhancer-blocking activity within HS5 may have gene and developmental stage specificity. When we examine these phenotypes carefully, we can generalize that the gene with lower transcriptional activity is always more severely affected (e.g., the γ gene is less abundantly transcribed [less than 10% of the β gene] [25] and more severely affected than the β gene at the fetal liver stage; Fig. 3B), precisely in accord with the observed transcriptional effects elicited by the gypsy insulator (30). When more than one gene can potentially interact with a shared enhancer in a competitive environment (8, 16) but only one of them can be preferentially expressed, such a feature of enhancer-blocking activity would be beneficial for finely tuning gene expression activity both for multigenic loci and also generally. We therefore propose that the original role of the enhancer-blocking function of insulators might have been to establish proper enhancer-promoter interactions by reducing inappropriate noise in surrounding gene activity. The facts that consensus CTCF binding motifs are not strictly conserved in sequence and that these are found at many sites throughout the genome are consistent with this hypothesis (2). Furthermore, the conclusion that a promoter with more feeble activity is always more severely affected is reminiscent of the promoter competition hypothesis (8). The enhancer-blocking insulator may thus be playing a role by participating in the competition as one of the competing elements.

When we removed a CTCF binding motif from the ectopically positioned human HS5 fragment (*HS5/dCTCF*), enhancer-blocking activity was completely lost in definitive erythroid cells (in both the adult spleen and fetal liver). We concluded that this CTCF binding motif is solely responsible for the transcriptional attenuating activity of HS5 at this stage. This

observation also confirmed that the reduced expression of β -globin in *HS5* transgenic mice is not due to increasing the distance between the LCR and the promoter. However, in primitive erythroid cells (yolk sac), the situation was radically different: the level of ϵ -globin gene expression was still far lower than the abundance observed in the wild-type locus. Although γ -globin expression was not affected significantly in primitive erythroid cells of *HS5/dCTCF* transgenic mice, this was also true for *HS5* transgenic mice bearing an intact CTCF binding motif. The degree of repression in ϵ expression was around 94% and 86% in *HS5* and *HS5/dCTCF* transgenic mice, respectively. If we assume that this difference is statistically meaningless, we could draw an intriguing conclusion: that the enhancer-blocking activity associated with disrupted CTCF binding may be completely absent in primitive erythroid cells. Thus, the reduced level of ϵ -globin gene expression detected at this stage in both *HS5* and *HS5/dCTCF* transgenic mice may be governed by a ϵ -globin gene-specific activity within HS5 that is independent of CTCF binding. This hypothesis, if correct, could also explain why γ -globin expression was not affected in yolk sac erythroid cells of *HS5* or *HS5/dCTCF* transgenic mice. This hypothesis is challenged by the fact that most enhancer-blocking assays performed to date have been conducted in K562 cells, which presumably most closely approximate the primitive erythroid environment, with the γ -globin promoter as the reporter gene (14).

The CTCF-independent ϵ -globin gene-specific attenuating activity associated with the HS5 fragment described here could be a novel enhancer-blocking activity which has strict gene specificity, or it could represent a different class of activity contained within the seven tandem copies of GATA motifs, which were reported to repress ϵ -globin promoter activity in K562 cells (27). It is also possible that this phenotype may be simply due to extending the distance between the LCR and the ϵ -globin gene promoter because of insertion of the 2.6-kb DNA fragment (the distance between HS2 and the ϵ promoter expands from 10 kb to 13 kb after insertion of the ectopic HS5 fragment). If this is indeed the case, the ϵ -globin gene must be far more susceptible to distance effects than the γ -globin genes.

LCR HS5 is a complex molecular unit, composed of multiple activities that may individually or cooperatively modulate β -globin locus regulation. In this work, we showed that one such function is a CTCF-dependent enhancer-blocking activity that is active in definitive erythroid cells. We also observed a ϵ -globin gene-specific attenuating activity in HS5 that appears to function independently of CTCF binding; the origin of this activity with respect to previously identified *cis*-DNA elements within the 2.6-kbp HS5 fragment is currently under investigation.

ACKNOWLEDGMENTS

We thank K. Foley and R. Eisenman for providing the CMV-Cre transgenic mice and Y. Tanimoto and W. Song for outstanding technical assistance.

This work was supported by grants from the Inamori Foundation (K.T.), the NIH (HL 24415; J.D.E.), the 21st Century COE Program (A.F.), and Grants-in-Aid for Scientific Research (to A.F. and K.T.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

1. 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.
2. Bell, A. C., A. G. West, and G. Felsenfeld. 1999. The protein CTCF is required for the enhancer-blocking activity of vertebrate insulators. *Cell* **98**:387–396.
3. Bender, M. A., A. Reik, J. Close, A. Telling, E. Epner, S. Fiering, R. Hardison, and M. Groudine. 1998. Description and targeted deletion of 5' hypersensitive site 5 and 6 of the mouse beta-globin locus control region. *Blood* **92**:4394–4403.
4. Bulger, M., M. A. Bender, J. H. van Doorninck, B. Wertman, C. M. Farrell, G. Felsenfeld, M. Groudine, and R. Hardison. 2000. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. *Proc. Natl. Acad. Sci. USA* **97**:14560–14565.
5. Bulger, M., J. H. van Doorninck, N. Saitoh, A. Telling, C. Farrell, M. A. Bender, G. Felsenfeld, R. Axel, M. Groudine, and J. H. van Doorninck. 1999. Conservation of sequence and structure flanking the mouse and human beta-globin loci: the beta-globin genes are embedded within an array of odorant receptor genes. *Proc. Natl. Acad. Sci. USA* **96**:5129–5134.
6. Bungert, J., U. Dave, K. C. Lim, K. H. Lieuw, J. A. Shavit, Q. Liu, and J. D. Engel. 1995. Synergistic regulation of human beta-globin gene switching by locus control region elements HS3 and HS4. *Genes Dev.* **9**:3083–3096.
7. Bungert, J., K. Tanimoto, S. Patel, Q. Liu, M. Fear, and J. D. Engel. 1999. Hypersensitive site 2 specifies a unique function within the human beta-globin locus control region to stimulate globin gene transcription. *Mol. Cell. Biol.* **19**:3062–3072.
8. Choi, O.-R., and J. D. Engel. 1988. Developmental regulation of beta-globin gene switching. *Cell* **55**:17–26.
9. Chung, J. H., A. C. Bell, and G. Felsenfeld. 1997. Characterization of the chicken beta-globin insulator. *Proc. Natl. Acad. Sci. USA* **94**:575–580.
10. Chung, J. H., M. Whiteley, and G. Felsenfeld. 1993. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74**:505–514.
11. Clouston, W. M., B. A. Evans, J. Haralambidis, and R. I. Richards. 1988. Molecular cloning of the mouse angiotensinogen gene. *Genomics* **2**:240–248.
12. Engel, J. D., and K. Tanimoto. 2000. Looping, linking, and chromatin activity: new insights into beta-globin locus regulation. *Cell* **100**:499–502.
13. Farrell, C. M., A. Grinberg, S. P. Huang, D. Chen, J. G. Pichel, H. Westphal, and G. Felsenfeld. 2000. A large upstream region is not necessary for gene expression or hypersensitive site formation at the mouse beta-globin locus. *Proc. Natl. Acad. Sci. USA* **97**:14554–14559.
14. Farrell, C. M., A. G. West, and G. Felsenfeld. 2002. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol.* **22**:3820–3831.
15. Feingold, E. A., L. A. Penny, A. W. Nienhuis, and B. G. Forget. 1999. An olfactory receptor gene is located in the extended human beta-globin gene cluster and is expressed in erythroid cells. *Genomics* **61**:15–23.
16. Foley, K. P., S. Pruzina, J. D. Winick, J. D. Engel, F. Grosveld, and P. Fraser. 1994. The chicken beta/epsilon-globin enhancer directs autonomously regulated, high-level expression of the chicken epsilon-globin gene in transgenic mice. *Proc. Natl. Acad. Sci. USA* **91**:7252–7256.
17. Gaensler, K. M., M. Burmeister, B. H. Brownstein, P. Taillon-Miller, and R. M. Myers. 1991. Physical mapping of yeast artificial chromosomes containing sequences from the human beta-globin gene region. *Genomics* **10**:976–984.
18. Kong, S., D. Bohl, C. Li, and D. Tuan. 1997. Transcription of the HS2 enhancer toward a cis-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. *Mol. Cell. Biol.* **17**:3955–3965.
19. Lakshmanan, G., K. H. Lieuw, K. C. Lim, Y. Gu, F. Grosveld, J. D. Engel, and A. Karis. 1999. Localization of distant urogenital system-, central nervous system-, and endocardium-specific transcriptional regulatory elements in the GATA-3 locus. *Mol. Cell. Biol.* **19**:1558–1568.
20. Leach, K. M., K. Nightingale, K. Igarashi, P. P. Levings, J. D. Engel, P. B. Becker, and J. Bungert. 2001. Reconstitution of human beta-globin locus control region hypersensitive sites in the absence of chromatin assembly. *Mol. Cell. Biol.* **21**:2629–2640.
21. Levings, P. P., and J. Bungert. 2002. The human beta-globin locus control region. *Eur. J. Biochem.* **269**:1589–1599.
22. Li, Q., M. Zhang, H. Han, A. Rohde, and G. Stamatoyannopoulos. 2002. Evidence that DNase I hypersensitive site 5 of the human beta-globin locus control region functions as a chromosomal insulator in transgenic mice. *Nucleic Acids Res.* **30**:2484–2491.
23. Ohlsson, R., R. Renkawitz, and V. Lobanikov. 2001. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.* **17**:520–527.
24. 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.
25. Porcu, S., M. Kitamura, E. Witkowska, Z. Zhang, A. Mutero, C. Lin, J. Chang, and K. M. 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.
26. Prioleau, M. N., P. Nony, M. Simpson, and G. Felsenfeld. 1999. An insulator element and condensed chromatin region separate the chicken beta-globin locus from an independently regulated erythroid-specific folate receptor gene. *EMBO J.* **18**:4035–4048.
27. Ramchandran, R., C. Bengra, B. Whitney, K. Lanclos, and D. Tuan. 2000. A (GATA)(7) motif located in the 5' boundary area of the human beta-globin locus control region exhibits silencer activity in erythroid cells. *Am. J. Hematol.* **65**:14–24.
28. Recillas-Targa, F., M. J. Pikaart, B. Burgess-Beusse, A. C. Bell, M. D. Litt, A. G. West, M. Gaszner, and G. Felsenfeld. 2002. Position-effect protection and enhancer-blocking by the chicken beta-globin insulator are separable activities. *Proc. Natl. Acad. Sci. USA* **99**:6883–6888.
29. Routledge, S. J., and N. J. Proudfoot. 2002. Definition of transcriptional promoters in the human Beta globin locus control region. *J. Mol. Biol.* **323**:601–611.
30. Scott, K. C., A. D. Taubman, and P. K. Geyer. 1999. Enhancer-blocking by the *Drosophila* gypsy insulator depends upon insulator anatomy and enhancer strength. *Genetics* **153**:787–798.
31. Stamatoyannopoulos, G., and A. W. Nienhuis. 1994. Hemoglobin switching, p. 107–155. *In* G. Stamatoyannopoulos et al. (ed.), *The molecular basis of blood diseases*, 2nd ed. W. B. Saunders Co., New York, N.Y.
32. 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.
33. Tanimoto, K., Q. Liu, J. Bungert, and J. D. Engel. 1999. The polyoma virus enhancer cannot substitute for DNase I core hypersensitive sites 2–4 in the human beta-globin LCR. *Nucleic Acids Res.* **27**:3130–3137.
34. Tanimoto, K., Q. Liu, F. Grosveld, J. Bungert, and J. D. Engel. 2000. Context-dependent EKLF responsiveness defines the developmental specificity of the human epsilon-globin gene in erythroid cells of YAC transgenic mice. *Genes Dev.* **14**:2778–2794.
35. Tuan, D., S. Kong, and K. Hu. 1992. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc. Natl. Acad. Sci. USA* **89**:11219–11223.
36. Yu, J., J. H. Bock, J. L. Slightom, and B. Villeponteau. 1994. A 5' beta-globin matrix-attachment region and the polyoma enhancer together confer position-independent transcription. *Gene* **139**:139–145.
37. Yu, R. N., M. Ito, T. L. Saunders, S. A. Camper, and J. L. Jameson. 1998. Role of A/Ch in gonadal development and gametogenesis. *Nat. Genet.* **20**:353–357.
38. Zafarana, G., S. Raguz, S. Pruzina, F. Grosveld, and D. Meijer. 1995. The regulation of human beta-globin gene expression: the analysis of hypersensitive site 5 (betaHS5) in the LCR, p. 39–44. *In* G. Stamatoyannopoulos (ed.), *Molecular biology of hemoglobin switching*. Intercept Ltd., Andover, United Kingdom.