MOLECULAR AND CELLULAR BIOLOGY, June 2002, p. 3820–3831 Vol. 22, No. 11
0270-7306/02/$04.00+0 DOI: 10.1128/MCB.22.11.3820–3831.2002

Conserved CTCF Insulator Elements Flank the Mouse and Human β-Globin Loci

Catherine M. Farrell, Adam G. West, and Gary Felsenfeld*

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

Received 15 November 2001/Returned for modification 17 January 2002/Accepted 13 February 2002

A binding site for the transcription factor CTCF is responsible for enhancer-blocking activity in a variety of vertebrate insulators, including the insulators at the 5′ and 3′ chromatin boundaries of the chicken β-globin locus. To date, no functional domain boundaries have been defined at mammalian β-globin loci, which are embedded within arrays of functional olfactory receptor genes. In an attempt to define boundary elements that could separate these gene clusters, CTCF-binding sites were searched for at the most distal DNase I-hypersensitive sites (HSs) of the mouse and human β-globin loci. Conserved CTCF sites were found at 5′HS5 and 3′HS1 of both loci. All of these sites could bind to CTCF in vitro. The sites also functioned as insulators in enhancer-blocking assays at levels correlating with CTCF-binding affinity, although enhancer-blocking activity was weak with the mouse 5′HS5 site. These results show that with respect to enhancer-blocking elements, the architecture of the mouse and human β-globin loci is similar to that found previously for the chicken β-globin locus. Unlike the chicken locus, the mouse and human β-globin loci do not have nearby transitions in chromatin structure but the data suggest that 3′HS1 and 5′HS5 may function as insulators that prevent inappropriate interactions between β-globin regulatory elements and those of neighboring domains or subdomains, many of which possess strong enhancers.

There is much conservation of structure and function at the multigene β-globin locus in vertebrates (21, 39). In erythroid cells, the region upstream of the genes is marked by a series of DNase I-hypersensitive sites (HSs) that comprise the locus control region (LCR). The LCR can confer high-level, copy number-dependent, position-independent β-globin gene expression in erythroid cells of transgenic mice, and thus, it has a dominant chromatin-opening activity in transgenic lines. At its natural location in the mouse and human loci, the LCR clearly plays a role in conferring high-level expression on the β-globin genes, but some questions remain about its role in chromatin opening (12, 23).

In the case of the chicken β-globin domain, both the 5′ and 3′ chromatin boundaries are marked by constitutive HSs (35, 36). Beyond these HSs, a strong transition from generally DNase I-sensitive and hyperacetylated chromatin to generally DNase I-insensitive and underacetylated chromatin is seen in erythroid cells (22, 28, 36). Each HS can act as an enhancer-blocking element in transfected erythroid cells but only when positioned between the enhancer and promoter (9, 36). The 5′ insulator, 5′HS4, can also protect transgenes against chromatin position effects in cell culture, fruit flies, mice, and rabbits (3). The enhancer-blocking function of 5′HS4 has been studied in detail (2, 10), and the minimal element responsible for this activity is a binding site for the 11-zinc finger transcription factor CTCF (2). A CTCF-binding site is also responsible for the enhancer-blocking activity of the chicken 3′HS and other vertebrate insulators (3, 30, 36). CTCF has also been implicated in positive and negative regulation of transcription, and it is a highly conserved and ubiquitous protein in vertebrates (8, 15, 30).

The neighboring genes of the human, mouse, and chicken β-globin loci have also been identified (Fig. 1A). The upstream neighbor of the chicken locus is a pre-erythroid-specific folate receptor gene that is separated from the β-globin domain by a 16-kb region of condensed chromatin (32). At least two olfactory receptor (OR) genes are present at the 3′ side of the chicken domain and are presumably expressed in olfactory tissue (6, 36, 38). Thus, the 5′ and 3′ boundary elements may function to prevent interactions between the β-globin locus and its neighboring genes. In the case of the human and mouse β-globin loci, a somewhat different scenario exists. The functional boundaries and the limits of the LCRs of these mammalian loci have not yet been defined, and general sequence homology between these loci exists further upstream of the major 5′ HSs of the LCR (6, 7). Both loci are also embedded within an array of functional OR genes (Fig. 1A), and several of these genes are included in the same "open" chromatin domain with the β-globin genes in erythroid cells (7). The relationship, if any, between the mammalian β-globin locus and its surrounding OR gene clusters is not known, and it is not understood how independent regulation of these genes is achieved in different tissues. No abrupt changes in chromatin structure have been defined immediately upstream of 5′HS5 or downstream of 3′HS1 in the mammalian loci (7, 37). Additional erythroid 5′ HSs have been found upstream of both loci, with the most 5′ sites being HS−62.5, located near the *MOR5′β4* gene promoter in the mouse locus (7, 14), and a homologous site that is near the orthologous human *HOR5′β7* gene (7).

In an attempt to define insulator elements that could function to avoid cross talk between these gene clusters in mammals, we searched for CTCF-binding sites at the most distal

* Corresponding author. Mailing address: Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892. Phone: (301) 496-4173. Fax: (301) 496-0201. E-mail: gary.felsenfeld@nih.gov.

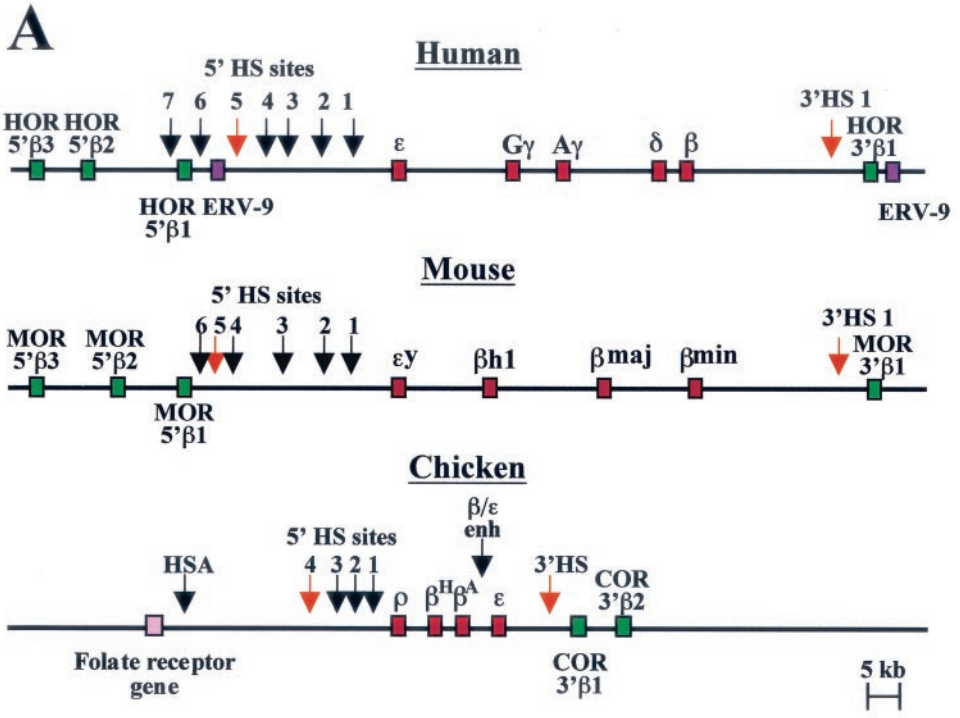


FIG. 1. Conserved CTCF sites flanking β-globin loci. (A) Extended structures of the human, mouse, and chicken β-globin loci. The human, mouse, and chicken β-globin loci are drawn approximately to scale. β-Globin genes are indicated by red boxes, and OR genes are indicated by green boxes. A pre-erythroid-specific folate receptor gene that exists upstream of the chicken β-globin locus is indicated by a pink box. The purple boxes in the human locus indicate ERV-9 retroviral elements (29). HSs are indicated by arrows. The rust-colored arrows indicate HSs where CTCF-binding sites were found. (B) Alignments of conserved CTCF sites at 5′HS5 and 3′HS1. Sixteen bases of the chicken 5′HS4 FII CTCF site (shown in red) are aligned with conserved human and mouse 5′HS5 sequences or in the opposite orientation with conserved human and mouse 3′HS1 sequences. Vertical lines indicate matches with both the human and mouse sequences, while dots indicate a match with only one of these species.

HSs of the human and mouse loci. The CTCF recognition sequence from the chicken 5′ insulator was used in this search. We found that conserved CTCF-binding sequences are present at 5′HS5 of both mammalian loci, while conserved sequences in the opposite orientation are also present at 3′HS1 of both loci. In vitro binding studies are presented showing the various abilities of these conserved sites to bind the CTCF factor, which are generally correlated with their various abilities to function in enhancer-blocking assays. We suggest that these conserved sites function as insulators at the mouse and human

β-globin loci, and models are presented indicating possible interactions that these insulators may prevent between regulatory elements in the β-globin cluster and those in neighboring domains.

MATERIALS AND METHODS

Sequences and CTCF site alignments. The sequences of the distal HSs flanking the human and mouse β-globin loci have been reported previously (6, 7, 16). The GenBank accession numbers for these sequences are as follows: human 5′ HSs, AF137396; human 3′HS1, X54282; mouse upstream HSs, AF071080; mouse 3′HS1, AF133300. The sequence 5′-CCGCTAGGGGGCAGCA-3′ or its reverse counterpart from footprint II (FII) of the chicken 5′HS4 insulator (10; GenBank accession number U78775) was used to search for conserved CTCF sites in the mouse and human sequences with the alignment program MultAlin (http://www.toulouse.inra.fr/multalin.html). Only potential sites that were conserved and had at least 11 bp of identity to the above FII sequence were chosen for study.

In vitro binding assays. Nuclear extracts from human K562 cells and partially purified chicken CTCF (pooled fraction following SP Sepharose chromatography) were prepared as described previously (2). Complementary single-stranded oligonucleotides were purified by denaturing polyacrylamide gel electrophoresis, annealed, and ^{32}P end labeled. Proteins were incubated with 20 to 40 fmol of labeled DNA probe in the presence of 50 ng of poly(dG-dC) competitor DNA per μl in 20 mM HEPES (pH 7.9)–150 mM KCl–5 mM $MgCl_2$–5% glycerol–1 mM dithiothreitol–0.5% Triton X-100 for 45 min at room temperature. Supershifts were carried out with anti-CTCF antibody as described previously (2). Complexes were resolved on 5% (29:1 acrylamide-bisacrylamide ratio) gels in 1× Tris-borate-EDTA at 150 V for 2 h. Relative binding strengths (see Table 1) were determined from the intensities of shifted bands in direct labeling gel shift assays, where 40 fmol of each labeled oligonucleotide was incubated with 200 ng of partially purified CTCF under the conditions described above.

Oligonucleotides. Complementary oligonucleotides were generated on an ABI 394 DNA synthesizer. The top strands of sequences used for gel retardation analyses were mouse 5′HS5-72 (TGACAAAACTAGAGAAAAAGAATGAGG CGTTTTCACCACTAGAGGGAAGGCAATTATTATGGAGCCTACAGA), mouse 3′HS1-72 (GGAGAGGAGGGCGGAAATCAGTGGAACACTTCTGC CCCCTACTGGTATGCAACAGGATCATTAGAGAAATGA), human 5′HS5-72 (CATCTTGGACCATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGG TCATAACAGCAGCTTCAGCTACCTCTC), and human 3′HS1-72 (TAGGG AGGGAAGAAAGTTTGATGAACTACTTCTGACCCCTAGTGGTGTCCA GAAAAGACCATTAAAGGAATG). The sequences of mutant sites are identical to those above, except for those bases in bold lowercase letters in Fig. 3. The full sequences of chicken 5′HS4FII and its mutants X3′ and ctt were described previously (2). The sequence of chicken 3′HS-A was described previously (36). Fragments used in enhancer-blocking assays were those used in gel retardation analyses, except that 102-mers with additional flanking sequences containing *Asc*I/*Sal*I restriction sites were used to make the mouse 5′HS5, human 5′HS5, and mouse 3′HS1 constructs. The top strands of these 102-mers were mouse 5′HS5-102 (AGGCGCGCCGTCGACTGACAAAACTAGAGAAAAAGAAT GAGGCGTTTTCACCACTAGAGGGAAGGCAATTATTATGGAGCCTAC AGAGTCGACGGCGCGCCT), mouse 3′HS1-102 (AGGCGCGCCGTCGAC GGAGAGGAGGGCGGAAATCAGTGGAACACTTCTGCCCCCTACTGG TATGCAACAGGATCATTAGAGAAATGAGTCGACGGCGCGCCT), and human 5′HS5-102 (AGGCGCGCCGTCGACCATCTTGGACCATTAGCTCC ACAGGTATCTTCTTCCCTCTAGTGGTCATAACAGCAGCTTCAGCTAC CTCTCGTCGACGGCGCGCCT).

Plasmids and constructs. pJCbasic was made by removing the chicken 5′HS4 insulator sequences from the *Sac*I and *Xba*I sites of previously described p137 (10), thereby leaving the ^{A}γ-globin promoter-neomycin reporter gene in the *Bam*HI site and the mouse 5′HS2 enhancer in the *Eco*RI site. pJCΔE was made by subsequently removing the enhancer from the *Eco*RI site of pJCbasic. The construct pJC5-4 was described previously (9). In order to make the remaining constructs, the oligonucleotides (described above) and their complementary strands were annealed. The double-stranded 102-mers were then restricted with *Hin*cII, followed by *Xba*I or *Sac*I linker ligation. One copy of the mouse 5′HS5, human 5′HS5, or mouse 3′HS1 sequence was subsequently subcloned into the *Xba*I and *Sac*I sites of pJCbasic to make the construct pmHS5Sac, phHS5Sac, or pm3′HS1Sac, respectively. In the case of pmHS5Sac-2X, phHS5Sac-2X, or pm3′HS1Sac-2X, a subclone with two copies of the mouse 5′HS5, human 5′HS5, or mouse 3′HS1 sequence in the *Sac*I site was selected, respectively. To make ph3′HS1Sac, ph3′HS1Sac-2X, and ph3′HS1Sac-4X, *Sac*I linkers were ligated onto the human 3′HS1 72-mer and one, two, or four copies, respectively, of this fragment were used to replace the *Sac*I fragment in pm3′HS1Sac. To make pmHS5KpnE, phHS5KpnE, pm3′HSKpnE, and ph3′HSKpnE, the HS2 enhancers were removed from the *Eco*RI sites of pmHS5Sac, phHS5Sac, pm3′HS1Sac, and ph3′HS1Sac, respectively, and the enhancers were replaced with *Kpn*I linker ligations into the *Kpn*I sites of these constructs. To make constructs with mutant test fragments, *Sac*I linkers were ligated onto the mouse 3′HS1, human 3′HS1, and human 5′HS5 aact mutant 72-mers. In the cases of pm3′HS-AACT and ph3′HS-AACT, the appropriate mutant *Sac*I fragment replaced the wild-type *Sac*I fragment in pm3′HS1Sac. In the case of phHS5-AACT, the human 5′HS5 mutant *Sac*I fragment replaced the wild-type *Sac*I fragment in phHS5Sac.

Enhancer-blocking assays. Colony assays with K562 cells were carried out as described previously (2, 9), with the recovery period before plating of the cells in selective medium being 26 h instead of 48 h. A concentration of 750 μg of active G418 per ml was used to select for neomycin-resistant colonies. In determining the relative number of neomycin-resistant colonies, the construct pJCbasic was used as a reference for the other constructs.

RESULTS

Conserved CTCF-binding sites flanking the mouse and human β-globin loci. The sequences containing the most distal HSs of the mouse and human β-globin loci were searched for the presence of CTCF-binding sites. Since CTCF is known to bind a variety of sequences (30), the sequence 5′-CCGCTAG GGGGCAGCA-3′ from FII (10) of the chicken 5′HS4 insulator was used for this search. It has been shown previously that the FII sequence is a useful indicator of CTCF-binding sites at other vertebrate insulators (2, 36). The domain organizations of the human, mouse, and chicken β-globin loci are shown in Fig. 1A, indicating the major HSs found at each locus. At the distal 5′ HSs, the best matches to the FII sequence were found at 5′HS5 of both the mouse and human loci (Fig. 1B). Over this region, 14 of the 16 bases are identical in the mouse and human sequences and 12 bases of the human site and 11 bases of the mouse site are identical to the chicken FII site. This conserved sequence was identified previously as a phylogenetic footprint in mice, galagos, and humans and was originally thought to be a potential binding site for the AML-1 factor (27). Likewise, another good match to the FII sequence is present at 3′HS1 of both mammalian loci (Fig. 1B), but this time the sequence is in the reverse orientation. Over this 16-bp stretch, 14 of the 16 bases are identical between the mouse and human sequences and 13 bases of each mammalian sequence are identical to the reverse FII sequence.

In vitro binding of the conserved sequences to CTCF. In order to see if the conserved CTCF sequences are capable of binding to CTCF, gel retardation assays were carried out (Fig. 2). Seventy-two-base-pair genomic fragments, with each containing the appropriate CTCF site in its center (see oligonucleotides in Materials and Methods), were first tested for the ability to compete for binding of CTCF to the chicken FII element (Fig. 2A). All of these sites were able to compete somewhat for binding, although not as well as the FII element itself. This suggested that the mammalian sites were bona fide CTCF-binding sites, which was shown by direct binding studies either with nuclear extract from human erythroleukemia K562 cells or with partially purified chicken CTCF (Fig. 2B). In each case, CTCF-DNA complexes were detected, and these complexes could be supershifted by incubation with a CTCF antibody (Fig. 2B, lanes 3, 6, 9, 12, and 15). Since the oligonucleotides used were quite large and given the wide sequence variation among CTCF-binding sites (30), it was then necessary

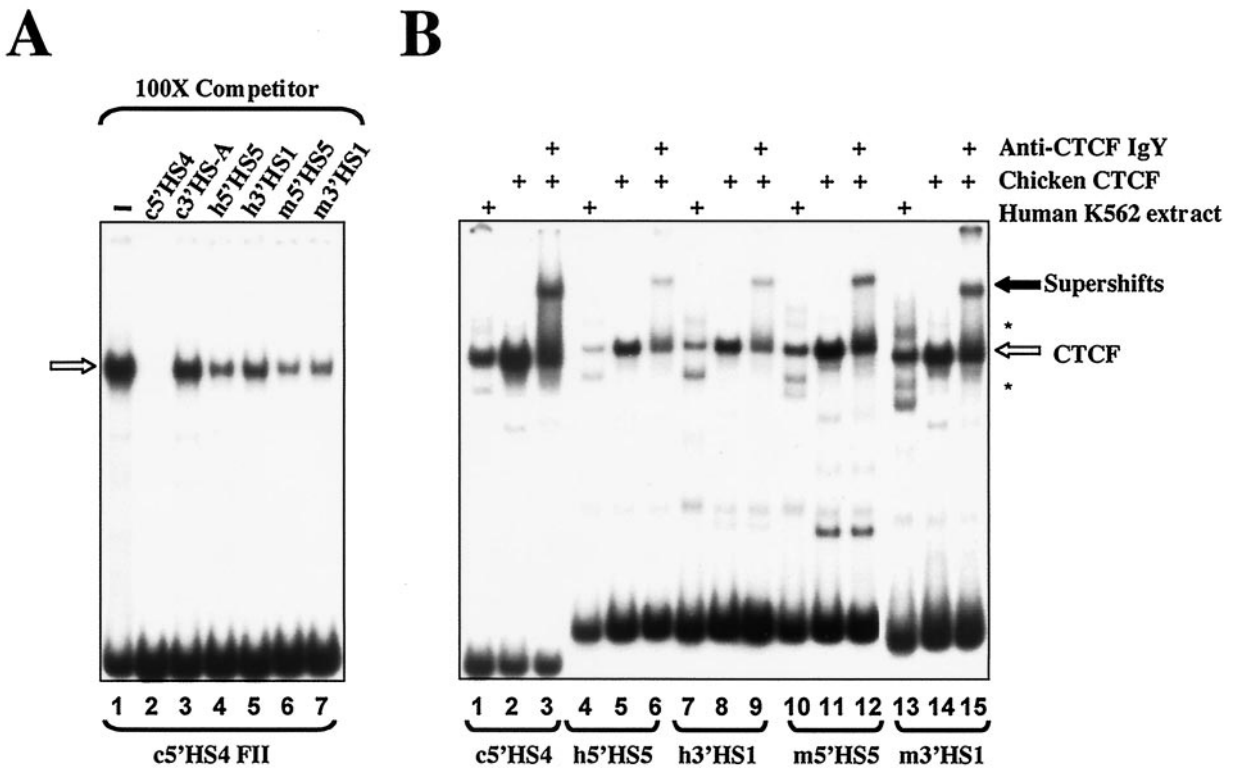
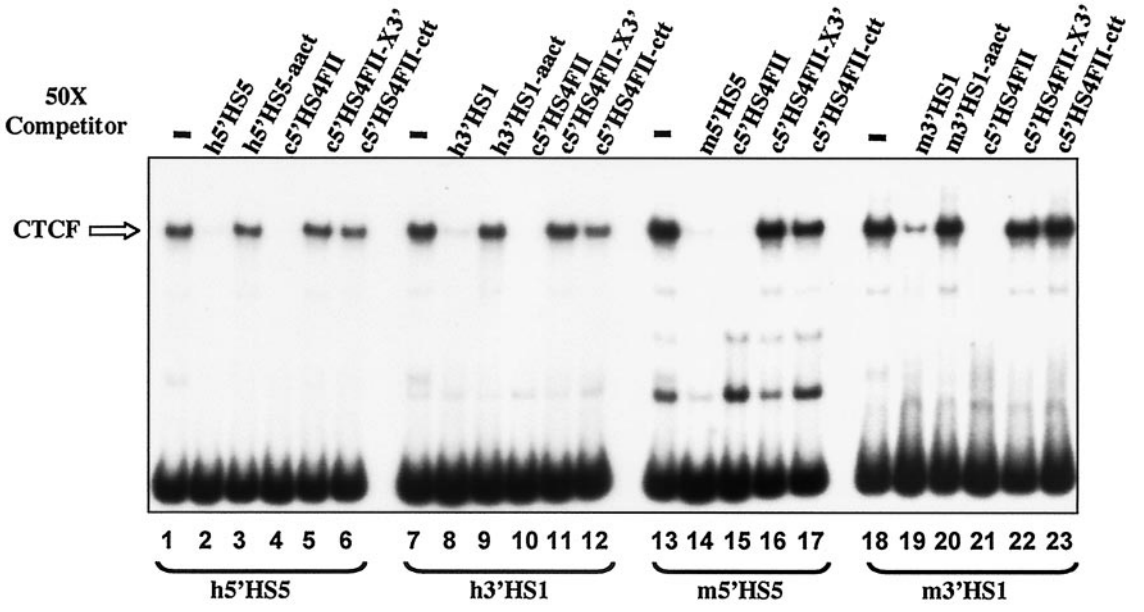


FIG. 2. CTCF binds to sequences from HSs that flank the human and mouse β-globin loci. (A) Competition of binding of CTCF to the chicken 5′HS4 FII enhancer-blocking element with putative CTCF-binding sequences from the human and mouse β-globin loci. Gel retardation analysis of complexes between labeled FII (40 fmol) and purified chicken CTCF (200 ng; indicated by an open arrow) following competition with a 100-fold excess (4 pmol) of unlabeled competitor duplexes. Lanes: 1, no competition; 2, chicken 5′HS4FII; 3, chicken 3′HS-A; 4, human 5′HS5; 5, human 3′HS1; 6, mouse 5′HS5; 7, mouse 3′HS1. (B) Gel retardation analysis of complexes formed between labeled chicken 5′HS4FII (40 fmol; lanes 1 to 3), human 5′HS5 (20 fmol; lanes 4 to 6), human 3′HS1 (20 fmol; lanes 7 to 9), mouse 5′HS5 (20 fmol; lanes 10 to 12), or mouse 3′HS1 (40 fmol; lanes 13 to 15) site and either 2 μg of human K562 nuclear extract (lanes 1, 4, 7, 10, and 13) or 200 ng of purified chicken CTCF (lanes 2, 3, 5, 6, 8, 9, 11, 12, 14, and 15). CTCF-DNA complexes are indicated by an open arrow. Asterisks indicate nonspecific (data not shown) complexes formed with K562 nuclear extract. CTCF-DNA complexes were supershifted with the addition of anti-CTCF (C-terminal) immunoglobulin Y (IgY) (lanes 3, 6, 9, 12, and 15). Supershifted complexes are indicated by the filled arrow.

to see if binding of CTCF to the mammalian sites was specific for the conserved bases (Fig. 3). This was shown by competition assays with oligonucleotides that were mutated at conserved bases, which were shown previously to be important for binding of CTCF to the chicken enhancer-blocking elements (2, 36). For each mammalian CTCF-DNA complex, competition was evident with a wild-type version of the appropriate binding site (Fig. 3, lanes 2, 8, 14, and 19) and with the wild-type chicken FII sequence (Fig. 3, lanes 4, 10, 15, and 21). However, when aact mutant oligonucleotides were used (Fig. 3, lanes 3, 9, and 20) or when mutant versions of the chicken FII element were used (Fig. 3, lanes 5, 6, 11, 12, 16, 17, 22, and 23), little or no competition was seen. This indicates that binding of CTCF to the mammalian β-globin sites is specific for bases conserved with the chicken FII enhancer-blocking element.

Even though all of these sites could bind specifically to CTCF, some differences in their relative binding affinities were detected (Table 1 and data not shown). The strongest binding affinities were seen with the chicken FII element and the mouse 3′HS1 sequence, while moderately good binding was seen with the human 3′HS1 and 5′HS5 and mouse 5′HS5 sites. These binding affinities were somewhat stronger than that seen with the chicken 3′HS-A site, which was shown previously to be a weaker binding element than the chicken FII site (36) (Table 1). These differences in binding strength indicate that while the central conserved sequences of these sites are important, other nonconserved genomic sequences that flank the conserved region also play a role in protein binding. It is known that CTCF binds to ~50 bp of sequence and that it does so by using different combinations of its 11 zinc fingers (15, 30). Thus, the different binding strengths that we observed among the conserved β-globin CTCF sites are likely to represent variations in combinatorial use of the zinc fingers.

Enhancer-blocking activities of the 5′HS5 CTCF sites. In order to see if the 5′HS5 CTCF sites display enhancer-blocking activity, colony assays with stably transfected K562 cells were carried out (Fig. 4). The test fragments in each case were 72-bp genomic fragments containing the appropriate conserved CTCF site, as used in the binding assays. In each colony assay, a construct containing the reporter gene (γ-Neo) and the mouse 5′HS2 enhancer (E) was used as a reference to determine the relative colony number (construct 1), while a construct with the reporter gene alone was used as an enhancerless negative control (construct 2). In addition, the construct pJC5-4 (construct 3) (9), which contains one copy of the 1.2-kb




```
c5′HS4FII WT   CCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGGCAGCAGCGA
c5′HS4FII X3′  CCAGGGATGTAATTACGTCCCTCCaaatagctttttCAGCAGCGA
c5′HS4FII ctt  CCAGGGATGTAATTACGTCCCTCCCCCGCTcttGGGCAGCAGCGA

h5′HS5 WT      CATTAGCTCCACAGGTATCTTCTTCCCTCTAGTGGTCAT-AACAG
h5′HS5 aact    CATTAGCTCCACAGGTATCTTCTTCCaTaTctTGGTCAT-AACAG

h3′HS1 WT      AGTTTGATGAACT----ACTTCTGACCCCTAGTGGTGTCCAGAAA
h3′HS1 aact    AGTTTGATGAACT----ACTTCTGACaCaTctTGGTGTCCAGAAA

m5′HS5 WT      AGAGAAAAAGAATGAGGCGTTTTCACCACTAGAGGGAAG--GCAA

m3′HS1 WT      AATCAGTGGAAC-----ACTTCTGCCCCCTACTGGTATGCAACAG
m3′HS1 aact    AATCAGTGGAAC-----ACTTCTGCCaCaTctTGGTATGCAACAG
```

FIG. 3. Binding of CTCF to mammalian β-globin sites is specific for bases conserved with the chicken FII enhancer-blocking element. (A) Gel retardation analysis of complexes formed between the labeled human 5′HS5 (20 fmol; lanes 1 to 6), human 3′HS1 (20 fmol; lanes 7 to 12), mouse 5′HS5 (20 fmol; lanes 13 to 17), or mouse 3′HS1 (40 fmol; lanes 18 to 23) site and 300 ng of purified chicken CTCF. CTCF-DNA complexes are indicated by an open arrow. Complexes were competed with a 50-fold excess of the unlabeled human 5′HS5 (lane 2), human 5′HS5-AACT (lane 3), human 3′HS1 (lane 8), human 3′HS1-AACT (lane 9), mouse 5′HS5 (lane 14), mouse 3′HS1 (lane 19), mouse 3′HS1-AACT (lane 20), chicken 5′HS4FII (lanes 4, 10, 15, and 21), chicken 5′HS4FII-X3′ (lanes 5, 11, 16, and 22), or chicken 5′HS4FII-ctt (lanes 6, 12, 17, and 23) competitor site. Lanes 1, 7, 13, and 18, no competition. (B) Partial sequences of binding sites used in competition assays. Mutations are indicated in bold lowercase. WT, wild type.

5′HS4 chicken insulator sequence on each side of the reporter gene, was used as a control for enhancer-blocking activity. The test constructs also contained a copy of the appropriate 5′HS5 CTCF site on the other side of the reporter gene to block any enhancer-promoter activity between adjacent copies of the transgene in multicopy lines (Fig. 4A and B, constructs 4 to 7). When the human 5′HS5 CTCF site was tested in these assays (Fig. 4A), a 2.5-fold reduction in colony number was seen when one copy of this site was placed between the enhancer and promoter (construct 4). This reduction in colony number was further increased to 5.1-fold when two copies of this site were placed between the enhancer and promoter (construct 5). However, when a mutant version of this site was tested between the enhancer and promoter (construct 6), the same mutation described in Fig. 3, enhancer-blocking activity was lost. In addition, enhancer-blocking activity was mostly lost when a wild-type human 5′HS5 CTCF site was placed on the other side of the enhancer, such that it was no longer located between the enhancer and promoter (construct 7). Therefore, these data indicate that the human 5′HS5 CTCF site acted as a positional enhancer-blocking element in these assays.

The same enhancer-blocking assays performed with the mouse 5′HS5 CTCF site are shown in Fig. 4B. In this case, however, enhancer-blocking activity was considerably weaker than that obtained with the human 5′HS5 sequence. When one copy of the mouse 5′HS5 CTCF site was placed between the enhancer and promoter (construct 4), only 1.3-fold enhancer-blocking activity was seen. Enhancer-blocking activity increased to 1.8-fold when two copies of this fragment were placed between the enhancer and promoter (construct 5) but was lost when the fragment was placed on the other side of the enhancer (construct 6). This weak enhancer-blocking activity may reflect the

TABLE 1. Comparison of CTCF-binding affinity and enhancer-blocking activity

CTCF site	Binding affinity[a]	Enhancer-blocking activity[b]
Chicken 5′HS4 FII	+++++	+++++[c]
Mouse 3′HS1	+++++	++++
Human 3′HS1	+++	+++
Human 5′HS5	+++	+++
Chicken 3′HS	++	++[d]
Mouse 5′HS5	+++	+

[a] Relative DNA-binding affinities were determined from the intensities of shifted bands in gel shift assays (data not shown) using 200 ng of partially purified CTCF and 40 fmol of labeled duplexes. Under these conditions, >50% (+++++), 20 to 50% (++++), 5 to 2% (+++), or 1 to 5% (++) of the labeled duplexes were shifted.

[b] Relative enhancer-blocking activities were determined from colony assays. The highest level of activity (+++++) is 6-fold for the chicken 5′HS4 FII site, while the lowest level of activity (+) is 1.3-fold for the mouse 5′HS5 CTCF site. ++, 2-fold; +++, 2.5- to 2.9-fold; ++++, 3.6-fold.

[c] Data derived from Bell et al. (2).

[d] Data derived from Saitoh et al. (36) (3′HS-A site).

fact that mouse 5′HS5 does not appear to play a significant role in vivo (see Discussion).

Enhancer-blocking activities of the 3′HS1 CTCF sites. Colony assays were performed with 72-bp test fragments containing either the mouse 3′HS1 (Fig. 5A) or the human 3′HS1 (Fig. 5B) CTCF sequence. All of the test constructs contained a copy of the mouse 3′HS1 fragment on one side of the reporter gene to avoid cross talk between different copies of the transgene in multicopy lines. When one copy (construct 4) or two copies (construct 5) of the mouse 3′HS1 site were placed between the enhancer and promoter, 3.6- and 4.1-fold enhancer-blocking activities were seen, respectively (Fig. 5A). However, this enhancer-blocking activity was lost when either a mutant version of the CTCF site was used (construct 6) or when the site was no longer positioned between the enhancer and promoter (construct 7).

Enhancer-blocking activity was also tested with the human 3′HS1 CTCF site (Fig. 5B). When one copy (construct 4), two copies (construct 5), or four copies (construct 6) of this fragment were placed between the enhancer and promoter, 2.9-, 4.2-, and 5.1-fold enhancer-blocking activities were seen, respectively (Fig. 5B). As was the case for the equivalent mouse site, this enhancer-blocking activity was lost when a mutant version of the site was tested (construct 7) or when the site was no longer positioned between the enhancer and promoter (construct 8). Therefore, these results indicate that both mouse and human 3′HS1 CTCF sites acted as insulators in these assays.

Comparison of in vitro binding and enhancer-blocking activities. A general, though not complete, correlation can be made between the ability of these conserved sites to bind to CTCF in vitro and their ability to function in enhancer-blocking assays (Table 1), similar to the correlation made previously with the chicken FII site and its mutant derivatives (2). In this comparison, the chicken FII site shows the strongest binding affinity and enhancer-blocking activity. This 72-bp FII fragment was previously shown to give sixfold insulation (2) or approximately fivefold insulation when this site was included in a 1.2-kb fragment (Fig. 4 and 5, construct 3). The mouse 3′HS1 site, which bound to CTCF with an affinity similar to that of the FII site, also showed good enhancer-blocking activity (3.6-fold; Fig. 5A). Both the human 3′HS1 and 5′HS5 sites showed moderate binding affinities, and they each showed moderately good enhancer-blocking activity (2.9-fold for human 3′HS1 [Fig. 5B] and 2.5-fold for human 5′HS5 [Fig. 4A]). The comparative binding affinity of the chicken 3′HS site was somewhat weaker than that seen with the mammalian 3′HS1 sites, and this was reflected by a more modest enhancer-blocking activity, which was shown previously to be about twofold when tested as a 100-bp sequence (3′HS-A site) or about threefold when included in a larger fragment (36). However, an exception to the binding affinity–enhancer-blocking activity correlation was seen in the case of the mouse 5′HS5 site, which showed moderately good binding affinity (Table 1) but only weak enhancer-blocking activity (1.3-fold for one copy and 1.8-fold for two copies; Fig. 4B), suggesting that this site does not play a significant functional role at the β-globin locus.

DISCUSSION

In this study, we show that conserved CTCF-binding sequences are present at 5′HS5 and 3′HS1 of the mouse and human β-globin loci. This is reminiscent of the CTCF enhancer-blocking elements present at 5′HS4 and 3′HS of the chicken β-globin locus. We show that the conserved mammalian sequences can bind to CTCF in vitro and that all of them, except the mouse 5′HS5 site, can function reasonably well in enhancer-blocking assays. The conservation of these CTCF sites flanking the β-globin loci suggests that these elements may function as insulators. In Fig. 6, models are presented showing possible functions of these potential insulators, and these are discussed in detail below, along with other information pertaining to each HS.

Potential insulating role for human 5′HS5. No clear function in vivo has been demonstrated for human 5′HS5, which is present in some nonerythroid cells (HS V) (41) but is not ubiquitous (43). This HS is not necessary for LCR function in transgenic mice, and it does not have any activation properties in transiently or stably transfected cells (24) or transgenic mice (43). However, this HS is a multipartite site and various potential regulatory elements have been detected in the region surrounding it. In addition to the presence of the conserved CTCF site, the core of human 5′HS5 includes a dyad of CACC motifs, which may bind Krüppel-like zinc finger proteins (21), as well as consensus sequences for the erythroid-specific NF-E2 and GATA-1 factors (27). A retroviral element, named ERV-9, is present upstream of the 5′HS5 core (29). This retroviral element has a U3 enhancer region rich in GATA, CACC, and CCAAT motifs and is transcribed preferentially in erythroid tissue. The ERV-9 transcripts may be the originating point of the exclusively nuclear LCR transcripts that are detected at different erythroid developmental stages, and it has been suggested that these transcripts may play a role in creating and/or maintaining an open chromatin configuration (1, 20, 31). A silencing element consisting of seven tandem GATA repeats is present in the downstream region of human 5′HS5 (33). In addition, the presence of a matrix attachment region (MAR) has been reported in a 3-kb fragment encompassing 5′HS5 (25) and two topoisomerase II recognition sites, thought to contain this MAR activity, are present just downstream of the 5′HS5 core (42). This MAR contributes to position effect

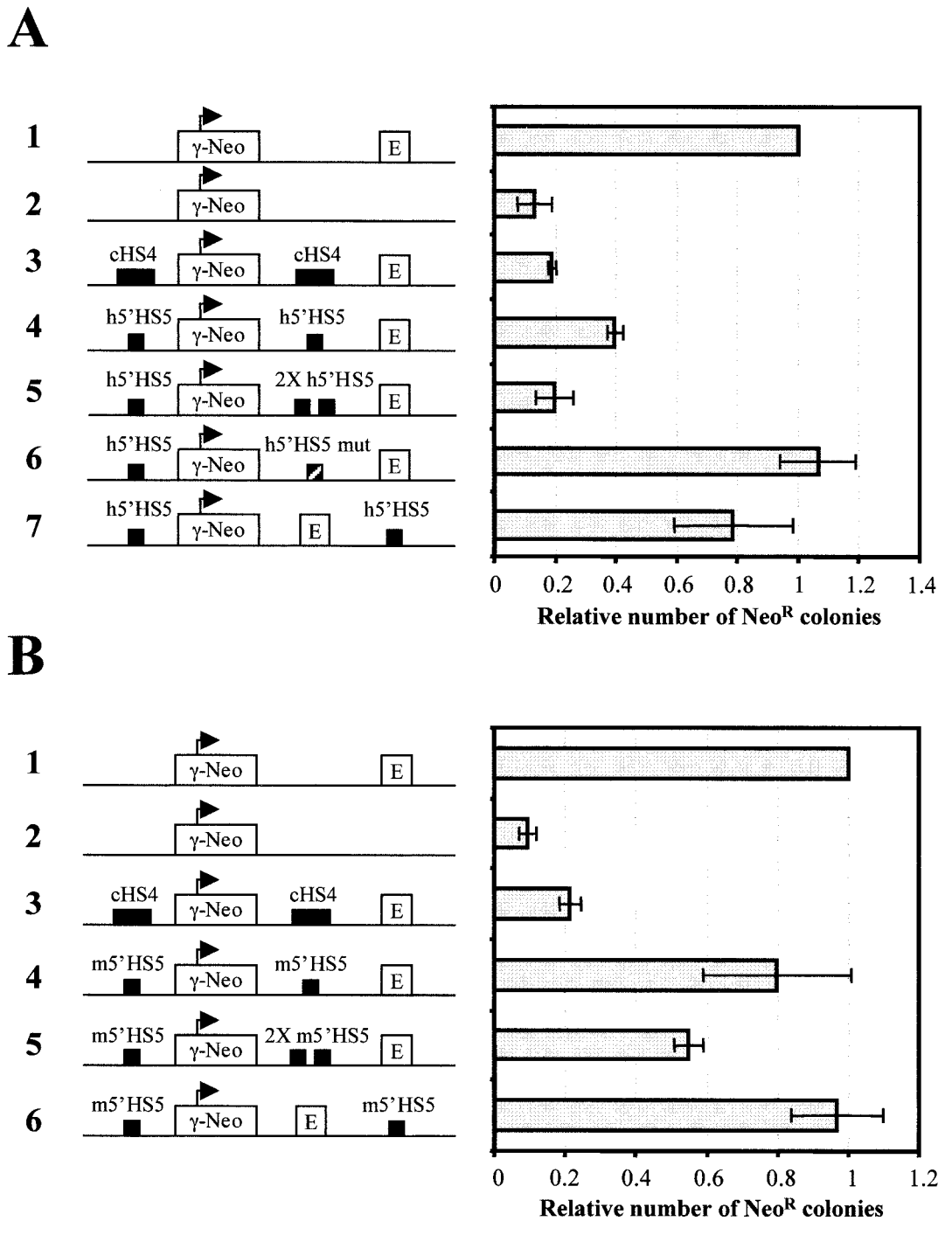


FIG. 4. Enhancer-blocking activities of the 5′HS5 CTCF sites. Colony assays were performed with transfected K562 cells. Constructs contained a neomycin reporter gene driven by the human $^{A}\gamma$-globin promoter (γ-Neo) and the mouse 5′HS2 enhancer (E). Shaded boxes in the construct diagrams indicate the positions of the test fragments, while striped boxes indicate those of the mutant test fragments. The average colony numbers obtained relative to construct 1 are plotted. Standard errors are indicated. (A) Enhancer-blocking assays with the human 5′HS5 CTCF site. The following constructs were used: 1, pJCbasic; 2, pJCΔE; 3, pJC5-4; 4, phHS5Sac; 5, phHS5Sac-2X; 6, phHS5-AACT; 7, phHS5KpnE. (B) Enhancer-blocking assays with the mouse 5′HS5 CTCF site. The following constructs were used: 1, pJCbasic; 2, pJCΔE; 3, pJC5-4; 4, pmHS5Sac; 5, pmHS5Sac-2X; 6, pmHS5KpnE.

protection when linked to the polyomavirus enhancer (42). Another study has also implicated human 5′HS5 in position effect protection in stably transfected cell lines, but the insulation effect was not complete in this case (24).

Of particular interest are reports showing that human 5′HS5 has enhancer-blocking activity. In two studies carried out with transfected erythroid cells (9, 26), positional enhancer-blocking activity was observed with 5′HS5 fragments of ~3 kb. Thus,

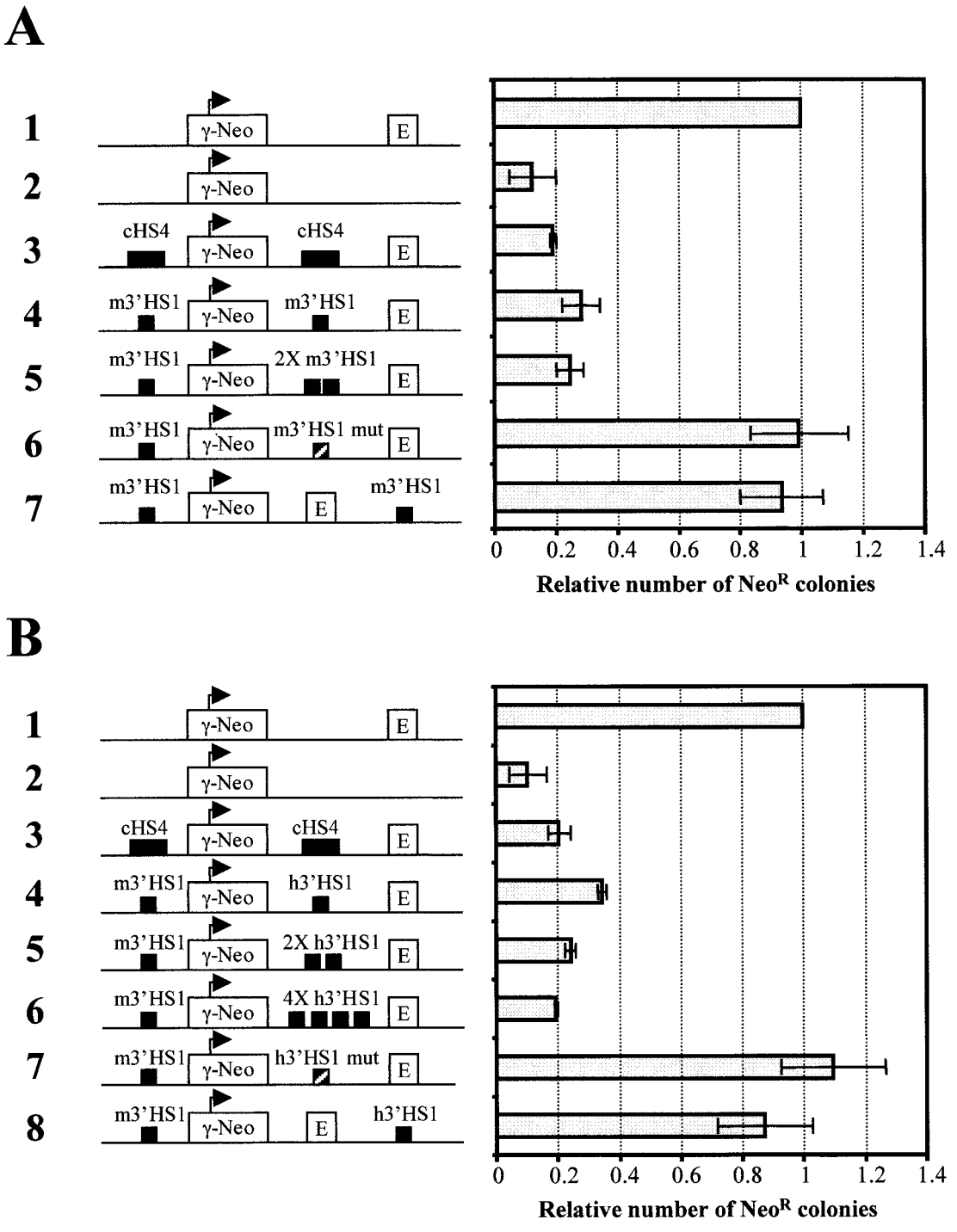


FIG. 5. The mouse and human 3′HS1 CTCF sites act as insulators in enhancer-blocking assays. (A) Enhancer-blocking assays with the mouse 3′HS1 CTCF site. The following constructs were used: 1, pJCbasic; 2, pJCΔE; 3, pJC5-4; 4, pm3′HS1Sac; 5, pm3′HS1Sac-2X; 6, pm3′HS-AACT; 7, pm3′HSKpnE. (B) Enhancer-blocking assays with the human 3′HS1 CTCF site. The following constructs were used: 1, pJCbasic; 2, pJCΔE; 3, pJC5-4; 4, ph3′HS1Sac; 5, ph3′HS1Sac-2X; 6, ph3′HS1Sac-4X; 7, ph3′HS-AACT; 8, ph3′HSKpnE.

it is likely that the CTCF sites present in these larger fragments are responsible for this enhancer-blocking activity. However, in yet another study with transfected erythroid cells, a 1.2-kb 5′HS5 fragment synergized with the ERV-9 enhancer to activate a transgene when it was placed between the enhancer and an ε-globin promoter (29). The discrepancies observed among these results may reflect the differences in the 5′HS5 fragments used, all of which contain multiple elements in addition to the CTCF site. It is possible that the 5′HS5 fragment used in the latter study included a particular combination of positive reg-

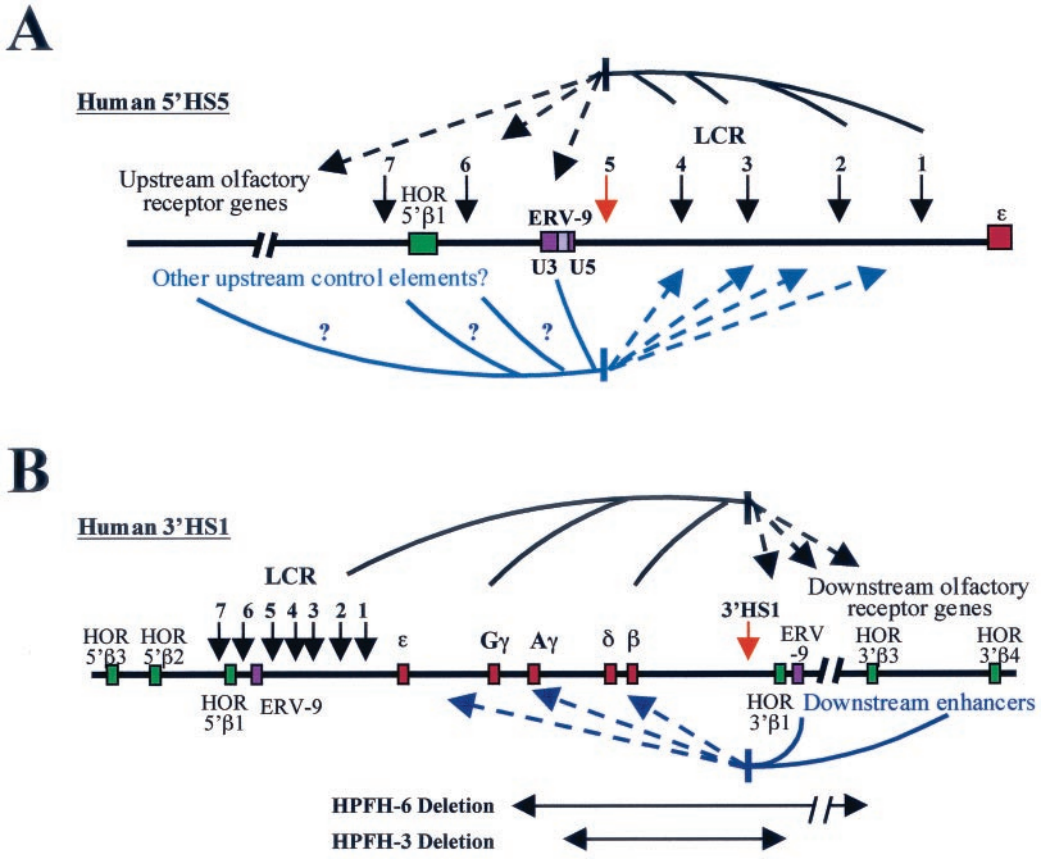


FIG. 6. Models indicating potential insulating functions of human 5′HS5 and 3′HS1. The boxes and arrows are described in the legend to Fig. 1A. The diagrams are not to scale relative to each other. (A) Model indicating possible insulating functions of the human 5′HS5 CTCF site. An insulating element at 5′HS5 could prevent inappropriate interactions between the LCR and upstream regions (black curved arrows shown on top) or vice versa (blue curved arrows shown underneath). (B) Possible insulating functions of human 3′HS1. 3′HS1 could prevent β-globin regulatory elements from inappropriately interacting with the downstream OR genes or other downstream elements (black curved arrows shown on top) or vice versa (blue curved arrows shown underneath). The extents of two HPFH deletions are shown underneath (17). A similar model can also be applied to mouse 3′HS1.

ulatory elements that negated the insulation effect, and this is consistent with the presence of a moderate enhancer component in the 5′HS5 insulator reported by Chung et al. (9).

There is also conflicting evidence for a human 5′HS5 insulator function in the context of the LCR in transgenic mice. In one such study, when a human β-globin gene was placed upstream of 5′HS5 with an additional copy of 5′HS5 on its other side, the β-globin gene was fully expressed, suggesting that 5′HS5 could not insulate against the activation properties of the LCR (43). In contrast, in another study with transgenic mice, when a marked ε-globin gene was placed upstream of 5′HS5, no expression of the ε-globin gene was seen, suggesting that 5′HS5 acted as an insulator (40). In the same study, the orientation of the LCR was reversed, such that 5′HS5 was closest to the β-globin genes. This caused a severe reduction in the transcription of all of the β-globin genes, suggesting that 5′HS5 is an enhancer-blocking element that prevents the LCR from interacting with the globin gene promoters.

In the present study, we demonstrated that the conserved CTCF site of human 5′HS5 has the ability to act as an enhancer-blocking element. Notwithstanding that human 5′HS5 is a multipartite and complex site associated with various regulatory elements, we propose a functional role for its CTCF component at the human β-globin locus (Fig. 6A). In this model, one possibility is that an insulating element at 5′HS5 could prevent inappropriate interactions between the LCR and upstream regions in erythroid cells. Upstream elements with which the LCR could interfere are the 5′ OR genes, 5′HS6 and -7, and the U3 enhancer of the ERV-9 retroviral element. Whether or not ERV-9 transcription plays a role in chromatin opening, an insulating element at 5′HS5 could be a means by which to prevent overactivation of this retroviral element by the strong LCR enhancers. Alternatively, 5′HS5 could prevent interactions between possible upstream regulatory elements and the downstream regions of the LCR. Such upstream elements, e.g., the ERV-9 enhancer, could potentially interfere with communication between LCR elements and the downstream β-globin gene promoters.

Mouse 5′HS5. Mouse 5′HS5 is in a region homologous with human 5′HS5, yet this is a very weak site in erythroid cells (4, 27). As in its human equivalent, an NF-E2 motif and a GATA-1 motif are present in the conserved core sequence but the CACC motifs present in human 5′HS5 are missing (4). In addition, mouse 5′HS5 does not have an element equivalent to ERV-9 to its 5′ side, nor are there topoisomerase II recognition sites to its 3′ side. Instead, an additional prominent HS,

termed 5′HS4.2 (4) or HS5A (27), is present downstream of mouse 5′HS5. Thus, there are considerable differences between mice and humans in this region of the LCR.

The weak enhancer-blocking activity we observe with the mouse 5′HS5 CTCF site is consistent with this HS being weak in erythroid cells and also with the phenotypes of mouse LCR deletions that include 5′HS5 (4, 5, 13, 14). Of these deletions, only a large deletion of 5′HS1 to -6 showed a significant β-globin phenotype (5, 13), which was apparently due to the absence of 5′HS1 to -4. The other smaller deletions of the 5′HS5 region gave a negligible β-globin phenotype (4, 14), suggesting that any putative boundary function at mouse 5′HS5 is not necessary for gene regulation at the β-globin locus.

While we do not believe that mouse 5′HS5 plays a significant functional role in erythroid cells, we cannot rule out the possibility that it plays a minor role in another tissue or developmental stage. Our in vitro binding assays indicate that CTCF can bind to this site specifically, and the enhancer-blocking activity that we observe, albeit quite weak, is positional. It may be relevant that mouse 5′HS5 is an easily detectable HS in L929 fibroblasts, much stronger than in murine erythroleukemia cells (C. Farrell, unpublished observations). A major subband of this HS, which coincides with the location of the CTCF sequence and a nearby NF-E2 motif, has also been detected by another group in the mouse spleen, brain, and thymus but not in the adult kidney or liver (27). This suggests that in some instances, the appearance of 5′HS5 may be due to the binding of a factor that is not erythroid specific, such as CTCF. This also applies to human 5′HS5 and chicken 5′HS4, which are not restricted to erythroid cells (35, 41).

Potential insulating role for 3′HS1. 3′HS1 is present as a strong HS in human erythroid cells, but its presence has also been shown in other cell types (11). This HS, which exists ~20 kb downstream of the β-globin gene, is characterized by three subbands that surround a very A/T-rich region (16), with the CTCF site corresponding to the 3′-most subband. Several GATA-1 sequences and an NF-E2 motif are present within this region, but 3′HS1 does not display enhancer activity. Two topoisomerase II recognition sites are also present at this HS, and this region was shown to serve as a scaffold-associated region in both erythroid and nonerythroid cells (16).

An enhancer-blocking element at human 3′HS1 could prevent the LCR and/or other strong positive elements of the β-globin locus from inappropriately activating the downstream OR genes in erythroid cells (Fig. 6B). Alternatively, and perhaps more interestingly, there is the possibility that a 3′HS1 insulator could prevent downstream enhancers from inappropriately activating the β-globin genes. It is known that enhancers exist in the downstream region from the analyses of the 3′ breakpoints of hereditary persistence of fetal hemoglobin (HPFH) deletions (17, 39). These enhancers are thought to be responsible for the expression of the fetal globin genes throughout adulthood, due to their juxtaposition next to these genes. The enhancer closest to 3′HS1 maps to the HPFH-3 3′ breakpoint, located just downstream of the *HOR3′β1* gene, and this also corresponds to an ERV-9 retroviral element (Fig. 6B) (29). All of the large HPFH deletions, except HPFH-5, include a deletion of 3′HS1, so it is possible that the absence of a 3′HS1 insulator would allow the downstream enhancers to activate the γ-globin genes. 3′HS1 is also missing in many deletion thalassemias, some of which display elevated γ-globin gene expression (39). However, HPFHs and thalassemias have a variety of causes, including point mutations in the β-globin genes or their promoters, so other elements can also be involved in the persistent expression of the fetal globin genes. In addition, there are variants of HPFH that have no known deletions or mutations within the β-globin gene cluster, but some of these are thought to be linked to the β-globin locus (39). It is not known if 3′HS1 is present in these HPFH variants.

In mice, 3′HS1 is present ~21 kb downstream of the β-minor gene in a region homologous with that of humans (6). Like its human counterpart, it consists of three hypersensitive subbands and the conserved CTCF site again corresponds to the 3′-most subband. Even though the role of mouse 3′HS1 has not been analyzed in detail, the conserved sequence and structure of this HS suggest that it plays a role similar to that of its human counterpart. Therefore, the model shown for human 3′HS1 (Fig. 6B) can also be applied to mouse 3′HS1. Even though no enhancers are known to exist in the region downstream of mouse 3′HS1, an insulator at this HS could prevent inappropriate interactions between possible downstream elements, either positive or negative, and the β-globin genes.

Significance of insulators flanking the β-globin domain. The presence of conserved enhancer-blocking elements flanking the human and mouse β-globin loci could be a means by which to separate β-globin control elements from neighboring control elements. Such a function has also been suggested for the insulators at 5′HS4 and 3′HS of the chicken β-globin domain (36). However, while the chicken insulators coincide with the 5′ and 3′ structural chromatin boundaries, the putative mammalian insulators at 5′HS5 and 3′HS1 do not. Since it is not necessary that all insulators mark a transition in chromatin structure, this suggests that the potential insulating function of the conserved mammalian elements may solely be to prevent cross interactions between control elements of neighboring domains. Thus, these insulators may not necessarily be able to protect against silencing by neighboring condensed chromatin structures. Not all insulators possess this property, e.g., the chicken 3′HS insulator or the differentially methylated domain boundary element of the mammalian *H19/Igf2* locus (36; F. Recillas-Targa and G. Felsenfeld, unpublished data). It is also known that this function is not attributable to the CTCF site in the case of the chicken 5′HS4 insulator (Recillas-Targa and Felsenfeld, unpublished). On the whole, the mammalian 5′HS5 and 3′HS1 sites did not resemble the chicken 5′HS4 insulator outside of the CTCF-binding site, but in mouse 3′HS1, we noted a sequence similar to footprint III (FIII) (10) of the chicken 5′ insulator. Like the CTCF site, this was also in the opposite orientation relative to the chicken sequence. While FIII of the chicken 5′ insulator is necessary for position effect protection, a fragment containing this sequence combined with its adjacent CTCF site functions as a silencing element in position effect assays (Recillas-Targa and Felsenfeld, unpublished).

If the mammalian CTCF sites do have a function, then these potential insulators would have to block interactions over very large distances. How such a function could be achieved is unknown. A simple model proposes that the binding of CTCF (and associated factors) interferes with enhancer-promoter loop formation, or else it prevents some sort of processive

mechanism carrying signals between enhancers and promoters (3). One possibility is that common proteins bound to 5′ and 3′ insulators could somehow interact with each other or provide an attachment point to the nuclear scaffold (19), thereby isolating the β-globin domain from neighboring regulatory elements. This model may be of more relevance to the human β-globin locus, where both the 5′HS5 and 3′HS1 CTCF sites displayed significant enhancer-blocking activity. Also relevant here is the presence of MARs or scaffold-associated regions at human 5′HS5 and 3′HS1 (discussed above). However, the formation of a discrete β-globin domain between 5′HS5 and 3′HS1 may not be feasible in all instances, especially with a 5′HS5 cutoff point, since large deletions at the human and mouse LCRs have suggested that sequences upstream of 5′HS5 are involved in chromatin opening (5, 13, 34). Thus, some level of communication between distant upstream elements and the β-globin genes may be necessary, at least at some developmental stage, and an insulator at 5′HS5 may only be required transiently. This could, in part, explain the weak enhancer-blocking activity of the mouse 5′HS5 CTCF site and the weak presence and the apparent dispensable function of this HS in erythroid cells. In addition, this also leaves room for the possibility that other boundary elements are present at more distal 5′ regions. Nonetheless, given the notion that the mammalian β-globin locus is divided into subdomains (18, 20), one could speculate that an insulator at 5′HS5 could function to separate control elements of these subdomains. The possibility that boundary elements may flank these subdomains was suggested previously (12). This idea also implies that enhancer-blocking elements can function to separate regulatory elements from one another within the same locus. This may be analogous to the *Drosophila Fab-7* and *Fab-8* insulators, which also disrupt long-range enhancer action but at the same time serve to separate adjacent enhancers of the *Abd-B* gene in the *Bithorax* complex (3).

In conclusion, the mammalian β-globin locus is a large and complex domain that is embedded within clusters of functional OR genes. There is considerable homology in the arrangements of regulatory elements in the mouse and human loci, but there is much less similarity between these loci and the chicken locus. It is therefore likely to be of functional significance that CTCF sites are located at conserved positions in all of these loci. The mouse and human β-globin loci are surrounded by other regulatory elements, including strong retroviral enhancers, from which the globin genes may need to be shielded. The properties of the enhancer-blocking sites, as demonstrated here, make them excellent candidates to serve that purpose, and some existing deletions are at least consistent with such a role. Further extensive deletion studies are necessary to define that role precisely.

ACKNOWLEDGMENTS

We are grateful to Cecelia Trainor, Vesco Mutskov, and Félix Recillas-Targa for critical reading of the manuscript and useful discussion.

REFERENCES

1. **Ashe, H. L., J. Monks, M. Wijgerde, P. Fraser, and N. J. Proudfoot.** 1998. Intergenic transcription and transinduction of the human β-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. **Bell, A. C., A. G. West, and G. Felsenfeld.** 2001. Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. Science **291:**447–450.
4. **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 β-globin locus control region. Blood **92:**4394–4403.
5. **Bender, M. A., M. Bulger, J. Close, and M. Groudine.** 2000. β-Globin gene switching and DNase I sensitivity of the endogenous β-globin locus in mice do not require the locus control region. Mol. Cell **5:**387–393.
6. **Bulger, M., J. H. von Doorninck, N. Saitoh, A. Telling, C. Farrell, M. A. Bender, G. Felsenfeld, R. Axel, and M. Groudine.** 1999. Conservation of sequence and structure flanking the mouse and human β-globin loci: the β-globin genes are embedded within an array of odorant receptor genes. Proc. Natl. Acad. Sci. USA **96:**5129–5134.
7. **Bulger, M., M. A. Bender, J. H. van Doorninck, B. Wertman, C. M. Farrell, G. Felsenfeld, M. Groudine, and R. C. Hardison.** 2000. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse β-globin gene clusters. Proc. Natl. Acad. Sci. USA **97:**14560–14565.
8. **Burcin, M., R. Arnold, M. Lutz, B. Kaiser, D. Runge, F. Lottspeich, G. N. Filippova, V. V. Lobanenkov, and R. Renkawitz.** 1997. Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF. Mol. Cell. Biol. **17:**1281–1288.
9. **Chung, J. H., M. Whitely, and G. Felsenfeld.** 1993. A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell **74:**505–514.
10. **Chung, J. H., A. C. Bell, and G. Felsenfeld.** 1997. Characterization of the chicken β-globin insulator. Proc. Natl. Acad. Sci. USA **94:**575–580.
11. **Dhar, V., A. Nandi, C. L. Schildkraut, and A. I. Skoultchi.** 1990. Erythroid-specific nuclease-hypersensitive sites flanking the human β-globin domain. Mol. Cell. Biol. **10:**4324–4333.
12. **Engel, J. D., and K. Tanimoto.** 2000. Looping, linking, and chromatin activity: new insights into β-globin locus regulation. Cell **100:**499–502.
13. **Epner, E., A. Reik, D. Cimbora, A. Telling, M. A. Bender, S. Fiering, T. Enver, D. I. K. Martin, M. Kennedy, G. Keller, and M. Groudine.** 1998. The β-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse β-globin locus. Mol. Cell **2:**447–455.
14. **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 β-globin locus. Proc. Natl. Acad. Sci. USA **97:**14554–14559.
15. **Filippova, G. N., S. Fagerlie, E. M. Klenova, C. Myers, Y. Dehner, G. Goodwin, P. E. Neiman, S. J. Collins, and V. V. Lobanenkov.** 1996. An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-*myc* oncogenes. Mol. Cell. Biol. **16:**2802–2813.
16. **Fleenor, D. E., and R. E. Kaufman.** 1993. Characterization of the DNase I hypersensitive site 3′ of the human β globin gene domain. Blood **81:**2781–2790.
17. **Forget, B. G.** 1998. Molecular basis of hereditary persistence of fetal hemoglobin. Ann. N. Y. Acad. Sci. **850:**38–44.
18. **Forsberg, C. M., and E. H. Bresnick.** 2001. Histone acetylation beyond promoters: long-range acetylation patterns in the chromatin world. Bioessays **23:**820–830.
19. **Gerasimova, T. I., K. Byrd, and V. G. Corces.** 2000. A chromatin insulator determines the nuclear localization of DNA. Mol. Cell **6:**1025–1035.
20. **Gribnau, J., K. Diderich, S. Pruzina, R. Calzolari, and P. Fraser.** 2000. Intergenic transcription and developmental remodeling of chromatin subdomains in the human β-globin locus. Mol. Cell **5:**377–386.
21. **Hardison, R., J. L. Slightom, D. L. Gumucio, M. Goodman, N. Stojanovic, and W. Miller.** 1997. Locus control regions of mammalian β-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. Gene **205:**73–94.
22. **Hebbes, T. R., A. L. Clayton, A. W. Thorne, and C. Crane-Robinson.** 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken β-globin chromosomal domain. EMBO J. **13:**1823–1830.
23. **Higgs, D. R.** 1998. Do LCRs open chromatin domains? Cell **95:**299–302.
24. **Jackson, J. D., H. Petrykowska, S. Philipsen, W. Miller, and R. Hardison.** 1996. Role of DNA sequences outside the cores of DNase hypersensitive sites (HSs) in functions of the β-globin locus control region: domain opening and synergism between HS2 and HS3. J. Biol. Chem. **271:**11871–11878.
25. **Jarman, A. P., and D. R. Higgs.** 1988. Nuclear scaffold attachment sites in the human globin gene complexes. EMBO J. **7:**3337–3344.
26. **Li, Q., and G. Stamatoyannopoulos.** 1994. Hypersensitive site 5 of the human β locus control region functions as a chromatin insulator. Blood **84:**1399–1401.
27. **Li, Q., M. Zhang, Z. Duan, and G. Stamatoyannopoulos.** 1999. Structural

analysis and mapping of DNase I hypersensitivity of HS5 of the β-globin locus control region. Genomics **61:**183–193.
28. **Litt, M. D., M. Simpson, F. Recillas-Targa, M.-N. Prioleau, and G. Felsenfeld.** 2001. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. EMBO J. **20:**2224–2235.
29. **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 β-globin locus control region. Genomics **54:** 542–555.
30. **Ohlsson, R., R. Renkawitz, and V. Lobanenkov.** 2001. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Gen. **17:**520–527.
31. **Plant, K. E., S. J. E. Routledge, and N. J. Proudfoot.** 2001. Intergenic transcription in the human β-globin gene cluster. Mol. Cell. Biol. **21:**6507–6514.
32. **Prioleau, M.-N., P. Nony, M. Simpson, and G. Felsenfeld.** 1999. An insulator element and condensed chromatin region separate the chicken β-globin locus from an independently regulated erythroid-specific folate receptor gene. EMBO J. **18:**4035–4048.
33. **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 β-globin locus control region exhibits silencer activity in erythroid cells. Am. J. Hematol. **65:**14–24.
34. **Reik, A., A. Telling, G. Zitnik, D. Cimbora, E. Epner, and M. Groudine.** 1998. The locus control region is necessary for gene expression in the human β-globin locus but not the maintenance of an open chromatin structure in erythroid cells. Mol. Cell. Biol. **18:**5992–6000.
35. **Reitman, M., and G. Felsenfeld.** 1990. Developmental regulation of topoisomerase II sites and DNase I-hypersensitive sites in the chicken β-globin locus. Mol. Cell. Biol. **10:**2774–2786.
36. **Saitoh, N., A. C. Bell, F. Recillas-Targa, A. G. West, M. Simpson, M. Pikaart, and G. Felsenfeld.** 1999. Structural and functional conservation at the boundaries of the chicken β-globin domain. EMBO J. **19:**2315–2322.
37. **Schübeler, D., C. Francastel, D. M. Cimbora, A. Reik, D. I. K. Martin, and M. Groudine.** 2000. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human β-globin locus. Genes Dev. **14:**940–950.
38. **Staines, D. M., and J. O. Thomas.** 1999. A sequence with homology to human HPFH-linked enhancer elements and to a family of G-protein linked membrane receptor genes is located downstream of the chicken β-globin locus. Gene **234:**345–352.
39. **Stamatoyannopoulos, G., and F. Grosveld.** 2001. Hemoglobin switching, p. 135–182. *In* G. Stamatoyannopoulos, P. W. Majerus, R. M. Perlmutter, and H. Varmus (ed.), The molecular basis of blood diseases, 3rd ed. W. B. Saunders Company, Philadelphia, Pa.
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 β-globin gene expression in mice. Nature **398:**344–348.
41. **Tuan, D., W. Solomon, Q. Li, and I. M. London.** 1985. The "beta-like-globin" gene domain in human erythroid cells. Proc. Natl. Acad. Sci. USA **82:**6384–6388.
42. **Yu, J., J. H. Bock, J. L. Slightom, and B. Villeponteau.** 1994. A 5′ β-globin matrix-attachment region and the polyoma enhancer together confer position-independent transcription. Gene **139:**139–145.
43. **Zafarana, G., S. Raguz, S. Pruzina, F. Grosveld, and D. Meijer.** 1995. The regulation of human β-globin gene expression: the analysis of hypersensitive site 5 (HS5) in the LCR, p. 39–44. *In* G. Stammatoyannopoulos (ed.), Molecular biology of hemoglobin switching. Intercept, Andover, Mass.