Published online February 23, 2004

1354–1362 *Nucleic Acids Research, 2004, Vol. 32, No. 4*
DOI: 10.1093/nar/gkh301

CTCF-dependent enhancer blockers at the upstream region of the chicken α-globin gene domain

Viviana Valadez-Graham, Sergey V. Razin[1] and Félix Recillas-Targa*

Instituto de Fisiología Celular, Departamento de Genética Molecular, Universidad Nacional Autónoma de México, Apartado Postal 70-242, México D.F. 04510, México and [1]Laboratory of Structural and Functional Organization of Chromosomes, Institute of Gene Biology of the Russian Academy of Sciences, Vavilov Str. 34/5, 119334 Moscow, Russia

Received January 6, 2004; Revised and Accepted February 3, 2004

ABSTRACT

The eukaryotic genome is partitioned into chromatin domains containing coding and intergenic regions. Insulators have been suggested to play a role in establishing and maintaining chromatin domains. Here we describe the identification and characterization of two separable enhancer blocking elements located in the 5′ flanking region of the chicken α-globin domain, 11–16 kb upstream of the embryonic α-type π gene in a DNA fragment harboring a MAR (matrix attachment region) element and three DNase I hypersensitive sites (HSs). The most upstream enhancer blocking element co-localizes with the MAR element and an erythroid-specific HS. The second enhancer blocking element roughly co-localizes with a constitutive HS. The third erythroid-specific HS present within the DNA fragment studied harbors a silencing, but not an enhancer blocking, activity. The 11 zinc-finger CCCTC-binding factor (CTCF), which plays an essential role in enhancer blocking activity in many previously characterized vertebrate insulators, is found to bind the two α-globin enhancer blocking elements. Detailed analysis has demonstrated that mutation of the CTCF binding site within the most upstream enhancer blocking element abolishes the enhancer blocking activity. The results are discussed with respect to special features of the tissue-specific α-globin gene domain located in a permanently open chromatin area.

INTRODUCTION

The eukaryotic genome is partitioned into active and inactive chromatin domains that are likely to constitute targets for special domain-level regulatory mechanisms (1,2). It is becoming increasingly evident that maintaining chromatin domain configuration with the possible regulatory consequences should be understood in the context of nuclear compartmentalization (1–5). Chromatin insulators are considered to be among the key players in this regulation (2,6,7). Insulators are functionally defined by two properties: (i) the ability to interfere with enhancer–promoter communication and (ii) the capacity to protect a transgene against position effects caused by local chromatin structure at some of the randomly chosen integration sites.

In vertebrates, the enhancer blocking activity of insulators is largely dependent on CCCTC-binding factor (CTCF), an 11 zinc-finger nuclear factor (8,9). One of the clearest examples of CTCF-dependent enhancer blocking activity is in the imprinted *Igf2*/*H19* loci, where the conditional (regulated by CpG methylation) binding of CTCF to the imprinting-choice region determines the imprinting status of paternal and maternal alleles via allele-specific enhancer blocking activity and the chicken β-globin cHS4 insulator (2,7,10,11).

In contrast to the domains of β-globin genes, the domains of α-globin genes (in chicken and other vertebrates studied) remain in an open chromatin configuration in both erythroid and non-erythroid cells (1,7,12,13). This might be explained by localization of α-globin gene domains in gene-dense genomic regions containing many housekeeping genes, including the so-called '–14' gene overlapping the upstream area of the α-globin domain (14–16).

Two clusters of DNase I hypersensitive sites (HSs) (each of them including a constitutive HS and several erythroid-specific ones) were found previously in the upstream area of the chicken α-globin domain (15). Several lines of indirect evidence suggested that the first of these groups, located 11–15 kb upstream of the embryonic α-type globin gene, could represent the upstream boundary of the chicken α-globin domain (15). Previous observations using transient transfections demonstrated that the DNA fragments harboring these HSs possess neither promoter nor enhancer activity (15). At this point it is important to recall that the majority of the insulator elements are neutral in terms of their transcriptional influence (6). Moreover, placing some of these fragments between the SV40 promoter and a reporter gene caused strong repression of reporter gene activity, which could suggest an enhancer blocking activity, characteristic of one of the previously defined insulator properties (6,15). In addition, the most upstream erythroid-specific HS was found to possess properties of a matrix attachment region (MAR), these elements are frequently present at the borders of different

*To whom correspondence should be addressed. Tel: +52 55 5622 5674; Fax: +52 55 5622 5630; Email: frecilla@ifc.unam.mx

tissue-specific gene domains, some of which possess insulator features, such as in the chicken lysozyme domain (15,17–19). In the present study we tested the DNA fragments harboring the above three HSs for the presence of an enhancer blocking activity. This activity was found in DNA fragments harboring the two most upstream of the three HSs studied. Furthermore, the corresponding DNA fragments were found to bind CTCF *in vitro*, and at least one of these sites interacts with CTCF *in vivo* in both erythroid and non-erythroid cells. Mutant analysis carried out on the corresponding enhancer blocking element has demonstrated that binding of CTCF is essential for the enhancer blocking activity. These findings strongly suggest that the most upstream group of HSs present in the 5′ flank of the chicken α-globin domain could separate distinct chromatin subdomains in a constitutively open genomic region.

MATERIALS AND METHODS

Plasmids and constructs

The recombinant clone containing the 6.6 kb BamHI fragment harboring the three most upstream HSs of the chicken α-globin domain has been described previously (15). Four subfragments of the insertion were PCR-amplified: a 1.4 kb DNA fragment harboring the 5′ erythroid-specific HS (αEHS-1.4), 1.2 and 0.6 kb DNA fragments harboring the constitutive HS with flanking sequences of different lengths (αCHS-1.2 and αCHS-0.6) and a 1 kb DNA fragment harboring the 3′ erythroid-specific HS (αEHS-1.0). Each pair of primers contained SalI and HindIII sites necessary for cloning the PCR-amplified DNA fragments in the pAcatE vector (20,21). Oligonucleotides used for amplification of the regions αEHS-1.4, αCHS-1.2, αCHS-0.6 and αEHS-1.0 were as follows. For αEHS-1.4: a, 5′-CCGGGGAAGCTTGTCGACTGTTTGCACCTCTAA-3′; b, 5′-CCGGGGAAGCTTGTCGACTTTCATGAAGTT-3′; αCHS-1.2: a, 5′-CCGGGGAAGCTTGTCGACGTCCCTCCTTTTCAAATCTAAC-3′; b, 5′-CCGGGGAAGCTTGTCGACCTATTTCAGATGTGATCA-3′; αCHS-0.6: a, 5′-CCGGGGAAGCTTGTCGACGGGTTCACAGAGAAGACTTG-3′, b, 5′-CCGGGGAAGCTTGTCGACAGCCACAAAATAGGA-3′; and αEHS-1.0: a, 5′-CCGGGGAAGCTTGTCGACGTGCTCAGCATA-3′; b, 5′-CCGGGGAAGCTTGTCGACAATTGCCTTCTGA-3′. The new pALucE luciferase reporter vector was constructed by digesting the pGL3-basic vector (Promega) with the BglII and BamHI restriction enzymes to obtain the luciferase gene with the poly(A) signal. This 2214 bp DNA fragment was subcloned in the pAcatE vector, which was digested with BamHI to eliminate the chloramphenicol acetyltransferase (*CAT*) gene. The pALucE vector retains the SalI site located between the gene and the β/ε enhancer. All tested fragments were cloned in this site.

Cell culture

Chicken cell lines DT40 [pre-B cells transformed by avian leukosis virus (ALV)], 6C2 (transformed pre-erythroblasts) and AEV-transformed erythroblasts [line LSCCHD3 (HD3)] were grown as described previously (22,23). HeLa and COS-1 cells (African green monkey kidney cells transformed by SV40) were grown in DMEM supplemented with 10% fetal bovine serum. Ten-day chicken embryonic red blood cells (RBCs) were cultivated as described previously (20).

Transient enhancer blocking assay

RBCs were obtained directly from the apical veins of chicken embryos. 1×10^8 cells were used per transfection with 6 μg of each linearized plasmid and 6 μg of pTkβ-gal plasmid (Stratagene). The latter served as an internal marker to normalize the efficiency of transfection. The RBCs were resuspended in 500 μl of PBS buffer and electroporated at 500 μF with 450 V. CAT activities were measured as described previously (20,21). In experiments with constructs bearing firefly luciferase and *Renilla* luciferase gene reporters, 1×10^8 cells were transfected with 1 μg of each plasmid and with 200 ng of the *Renilla* plasmid (pRL-CMV). Cells were resuspended in 500 μl of PBS buffer and electroporated at 500 μF and 450 V, luciferase activities were determined 48 h later using the Dual luciferase kit (Promega). Relative luciferase units were measured in a TD 20/20 luminometer (Turner Designs).

Chicken CTCF over-expression

Plasmid pSG5-CTCF (kindly provided by Dr Elena Klenova) was transfected into COS-1 cells by electroporation at 100 V and 950 μF. 24 h later, cells were lysed by freezing and thawing for 30 min at 4°C in RIPA buffer (150 mM NaCl, 1% Nonidet P-40, 0.5% deoxycholate, 0.1% SDS and 50 mM Tris–HCl pH 8.0) with 10 μg/ml of protease inhibitors (leupeptin, PMSF and pepstatin A from Sigma). After centrifugation, lysates were stored at –70°C. This same protocol was used for preparation of HeLa total cell extracts.

In vitro transcription/translation

The full-length chicken CTCF cDNA was *in vitro* transcribed and translated from the pSG5-cCTCF vector, using the TnT reticulocyte lysate-coupled *in vitro* transcription–translation system (Promega).

Electrophoretic mobility shift assay

The EMSA assay was performed as previously described (8). Competitions were carried out with 100 pmol of cold oligonucleotide. The following oligonucleotides (all purified by PAGE) were used in different mobility shift experiments (in all cases only one of the complementary chains is shown): M9, 5′-TTTCCATATAGAACAATGTAAATTCTTTCCTAGAGTTAAGCAGGGCTTTTTCATAAGCTG-3′; C10, 5′-CATGGCCAGCAGGCCCTCCTTGGTGCCCACCTTGCTGCAGTGGGACTGCGCTGCTCCGGA-3′; C11, 5′-CCTCAGTGAGCTGCCAGGGGAAGCTCATAGTTGCAGGATGGGAGCAGTGTGAGGTTTGAT-3′; C12, 5′-TGTGTGCCTCCATCAGACCACCCCCTCCACCATGCAGGTACCCCAGATAACACAGTATCA-3′; C13, 5′-TGGGTGTGTTCCACACCCACCTTATCTCTCGGCCACAAGCAGGCCCTCAGTGCCCACAGA-3′; C14, 5′-CCAAACAGGATGCCCAGCACAGCTGAGGGCTAGCAGAGAGGAGGGGAGAGT-3′; FII, 5′-CCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGGCAGCAGGCGCGCCT-3′ (8); and DMD, 5′-TTGGTTGTAGTTGTGGAATCGGAAGTGGCCGCGCGGCGGCAGTGCAGGCTCACACATCACAGCCCGAGC-3′ (10). The following oligonucleotides have SalI and

HindIII restriction sites. M9Δcag, 5′-TTTCCATATAGAACAATGTAAATTCTTTCCTAGAGTTAATCTGGGCTTTTTCATAAGCTG-3′; M9Δctag, 5′-TTTCCATATAGAACAATGTAAATTCTTTCACTTAGTTAAGCAGGGCTTTTTTCATAAGCTG-3′; M9ΔcagΔctag, 5′-TTTCCATATAGAACAATGTAAATTCTTTCACTTAGTTAATCTGGGCTTTTTCATAAGCG-3′.

Western blot

Cells were harvested with lysis buffer containing 20 mM HEPES pH 7.9, 400 mM KCl, 20% glycerol, 2 mM DTT, plus protease inhibitors, freeze–thawed twice, and proteins were quantified by the Bradford assay (BioRad). After separation by SDS–PAGE, the resolved proteins were transferred to PVDF membranes (Pharmacia). Membranes were blocked with 5% non-fat milk for 2 h and then incubated with the appropriate antibody for different time intervals (from 2 h to overnight) at 4°C. After washing, the blots were incubated for 1 h with an appropriate HRP-conjugated secondary antibody (Santa Cruz Biotechnology), washed and antibody-reactive proteins were detected with a chemiluminescence substrate (Pierce) according to manufacturer's instructions.

Generation of chicken CTCF antibody

The part of the chicken CTCF gene encoding the N-terminal region of the protein comprising amino acids 86–233 was cloned in the pET-28a$^{(+)}$ vector (Novagen). Protein expression was induced with IPTG and purification was carried out through a Ni$^{(+)}$ charged column. Once the protein was obtained, antibodies were generated by immunizing male New Zealand white rabbits. Antibody titers were estimated by ELISA.

Chromatin immunoprecipitation assay (ChIP)

The ChIP assay was performed as described by Weinmann *et al*. (24) with slight modifications. After formaldehyde cross-linking and sonication, chromatin corresponding to 2×10^7 cells was incubated overnight at 4°C with 20 µl of preimmune serum, and 5, 10, 15 and 20 µl of a 1/10 000 titer anti-cCTCF(86–233) or 10 µg of anti-CTCFN17 [shown as anti-CTCF(sc)] and 20 µl of protein A/G plus agarose (Santa Cruz Biotechnology) were added per reaction and incubated for 5 h at 4°C. An IgG (anti-rabbit IgG) was used as an irrelevant antibody. The mixture was centrifuged and the supernatant was saved as the unbound fraction. The immunoprecipitated fraction was washed as in the original protocol (24). After RNase I and proteinase K treatments and reversal of cross-linking, the DNA was precipitated, dissolved in 40 µl of deionized water and radioactive PCRs were then performed. Two pairs of oligonucleotides were designed: the first pair, 5′-CAGTCTGTCAGGAAGAAAGA-3′ and 5′-TGTGGCCTGGTTTTAGCTG-3′ amplifies a 120 bp region of the chicken αEHS-1.4 containing the M9 site; the second pair, 5′-CTGGGTGGGGGCAGGT-3′ and 5′-CTGCTTTTGCTGCCCTGTG-3′, used as a negative control, amplifies a 301 bp sequence containing a portion of the chicken β-globin enhancer [110–558 bp after the poly(A) signal of the β^A gene (24)]. Three independent ChIP experiments were carried out.

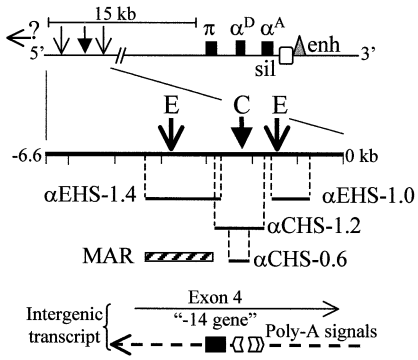

Figure 1. Scheme of the upstream non-coding region of the chicken α-globin domain. Positions of globin genes are indicated by black rectangles. The gray triangle and open rectangle show the positions of the enhancer (12,28) and silencer (12), respectively, located downstream of the gene cluster. Arrows show the positions of the upstream HSs studied in the present work. A detailed map of the area under study is shown below the map of the domain. The fragments bearing erythroid-specific HSs are named αEHS-1.4 and αEHS-1.0. The fragments bearing the constitutive HSs are named αCHS-1.2 and αCHS-0.6. The dashed box shows the position of a MAR element (15). At the bottom bi-directional intergenic transcript overlapping the studied area is shown (14,16).

RESULTS

The two most upstream HSs are enhancer blockers

The scheme of the chicken α-globin domain and a detailed map of the area under study are shown in Figure 1. The previously cloned 6.6 kb BamHI fragment harbors a constitutive HS flanked on each side by an erythroid-specific HS (15). The downstream part of the 6.6 kb DNA fragment was previously found to exhibit a moderate silencing activity, while the most upstream erythroid-specific HS was found to co-localize with a MAR element (15). Here the four sub-fragments harboring the individual HSs with flanking sequences were tested for enhancer blocking activity using the previously established (20) episomal enhancer blocking assay. The fragments to be tested were cloned into the pAcatE vector (20; Fig. 2, scheme at top) in both orientations at the SalI site (between the enhancer and promoter) and at the HindIII site (downstream of the enhancer) as shown in Figure 2. To avoid bi-directional action of the enhancer, all the constructs were linearized and transiently transfected in 10-day-old chicken embryo RBCs. The pAcatE-(II/III-Ins)Q plasmid (20) containing four copies of the FII/FIII footprints from the chicken 5′cHS4 β-globin insulator that possesses strong enhancer blocking activity was used as a positive control (Fig. 2A, construct 5). The pAcatE and the enhancer-less pAcat vectors were also used for transfection in each set of experiments in order to calibrate the system (Fig. 2A, constructs 6 and 7). The results of transient transfection experiments are shown in Figure 2. A significant reduction in reporter activity (relative to pAcatE, construct 6) was seen when the 1.2 kb DNA fragment was located between the enhancer and the promoter, independently of the orientation of the fragment (Fig. 2A, constructs 3 and 4). Consistent with a classic enhancer blocking activity, the same sequences did not significantly reduce the activity of the reporter gene when they were cloned outside the enhancer (Fig. 2A, constructs 1 and 2).

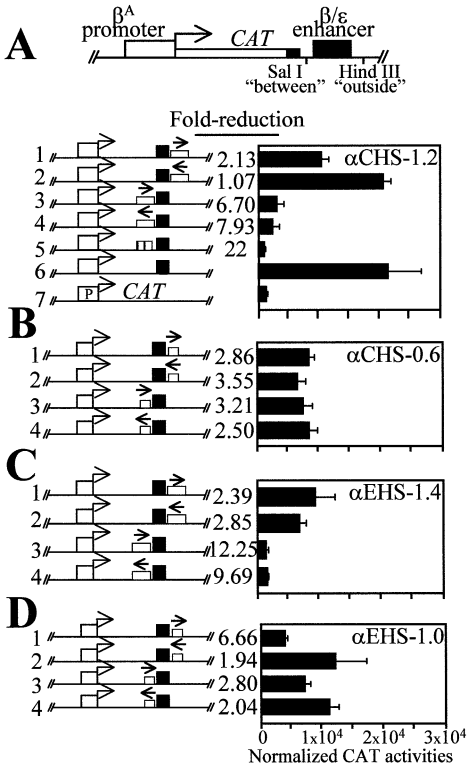

Figure 2. Transient enhancer blocking activity of the HSs located at the 5′ upstream region of the chicken α-globin locus. Transient transfections of linearized plasmids in chicken 10-day-old embryonic erythrocytes. The scheme of the basic construct with the β^A promoter and β^A/ε enhancer is shown at the top of the figure. (**A**) The constitutive HS (αCHS-1.2) tested 'between' and 'outside' in both orientations. Positions of the tested fragments relative to the enhancer are shown schematically at the left of the histogram. Construct 5 contains four copies of the FII/FIII sites of the β-globin insulator (striped box). Constructs 6 and 7 represent pAcatE and enhancer-less pAcat used to normalize the results. (**B**–**D**) Results obtained correspondingly with the αCHS-0.6, αEHS-1.4 and αEHS-1.0 fragments. All designations are the same as in (A). Data shown in histograms represent the average of seven independent experiments and error bars show the SEM in all cases.

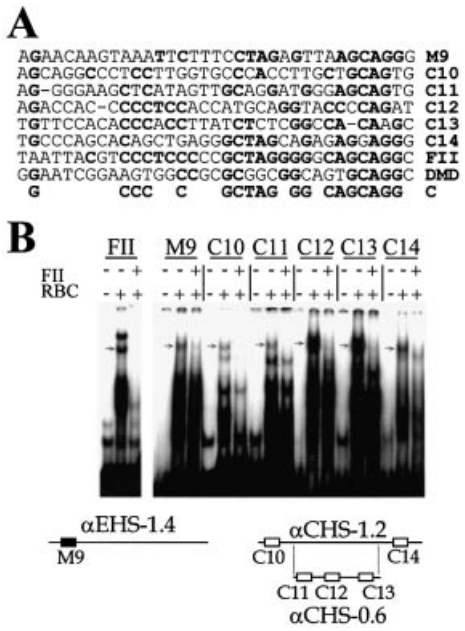

Figure 3. Sequence identification of conserved CTCF sites and binding activities. (**A**) Sequences of putative DNA binding sites for CTCF clustered in the αEHS-1.4 and αCHS-1.2 elements. Conserved nucleotides are shown in bold. The DMD sequence corresponds to a CTCF binding site of the differentially methylated domain from the mouse *Igf2*/*H19* imprinted locus (10). (**B**) Electrophoretic mobility shift assay using nuclear extracts from chicken RBCs. Each retarded complex was competed with 100-fold molar excess of FII unlabeled DNA. The location of each binding site is shown below.

Since several putative protein-binding sites were identified around the middle of the 1.2 kb sequence in the vicinity of the constitutive HS, we decided to test a shorter DNA fragment representing the central part of the αCHS-1.2 DNA fragment (see Fig. 1). Unexpectedly, this shorter 600 bp DNA fragment (αCHS-0.6) showed a moderate (3-fold) reduction in *CAT* reporter activity in both orientations and locations (Fig. 2B). Taken together, these data show that the αCHS-1.2 element possesses an enhancer blocking activity that is lost when the fragment is delimited to a 600 bp central part (see below). Instead, the αCHS-0.6 fragment shows a moderate silencing activity.

In the next set of experiments the fragments harboring the two erythroid-specific HSs (αEHS-1.4 and αEHS-1.0; see Fig. 1) were studied. The αEHS-1.4 DNA fragment was found to possess the enhancer blocking activity, in addition to a modest silencing activity. Indeed, a strong reduction in the reporter *CAT* gene activity was observed when the αEHS-1.4 DNA fragment was placed between the β^A promoter and the β^A/ε enhancer (Fig. 2C, constructs 3 and 4). The enhancer blocking effect in this case was even more prominent than that observed for the αCHS-1.2 DNA fragment. In contrast, the αEHS-1.0 fragment harboring the 3′ erythroid-specific HS behaves like a classical silencer, since the activity of the *CAT* gene was similarly reduced whether this fragment was placed between the enhancer and promoter or in the 'outside' position (Fig. 2D). Thus, the most upstream HSs, αEHS-1.4 and αCHS-1.2, are enhancer blocker elements, and the αEHS-1.4 shows a stronger ability to interfere with enhancer–promoter action.

CTCF interacts with the αEHS-1.4 and the αCHS-1.2 enhancer blocking elements

Since enhancer blocking activity is a property of insulators and it has been previously demonstrated that CTCF interacts with many known vertebrate insulators, we decided to analyze the DNA sequence of the region under study for the presence of potential CTCF-binding motifs (6,8,9,25). Identification of such motifs represents a difficult task, since the differential use of the 11 zinc-fingers allows CTCF to bind divergent sequences having no close similarity (8,9,25). We fixed a CTAG sequence as the starting motif of similarity with upstream and downstream CG-rich sequences and compared our sequences with the FII CTCF-binding site present in the chicken β-globin insulator (8,25). Computer-assisted analysis allowed us to identify six potential binding sites for CTCF (named M9 and C10–14; Fig. 3A). The M9 site was found in the αEHS-1.4 element while the αCHS-1.2 DNA fragment harbored several putative CTCF-binding motifs (C10–C14).

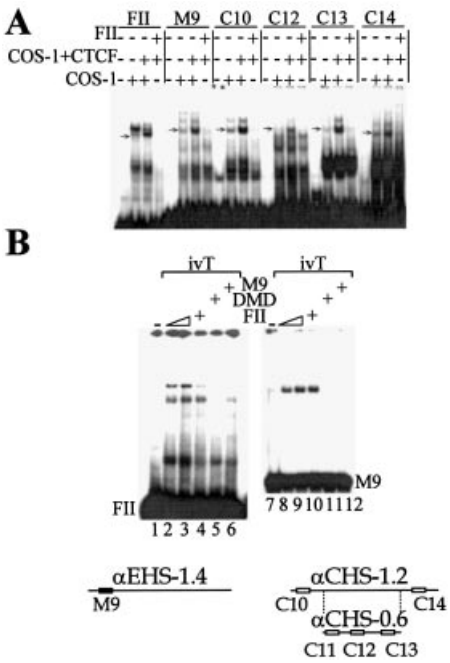

Figure 4. DNA–protein interactions of overexpressed and *in vitro* transcription/translated chicken CTCF. (**A**) The gel shift experiments show binding of nuclear extracts from cells transfected with the empty (COS-1) and the pSG5-CTCF vectors (COS-1+CTCF) to the FII, M9 and C10–C14 labeled oligonucleotides. The retarded complexes were competed with 100-fold molar excess of the chicken β-globin cHS4 FII insulator site. (**B**) Gel-shift experiments using CTCF obtained by *in vitro* transcription/translation (ivT). Labeled FII and M9 sites were used as probes and 25-fold molar excesses of the FII, M9 and mouse DMD sites were used as cold competitors as indicated above the slots. Open triangles over lanes 2, 3 and 8, 9 represent increasing amounts of *in vitro* transcription/translated CTCF.

Most importantly, no apparent CTCF-binding motifs were found in the αEHS-1.0 fragment that did not show enhancer blocking activity (Fig. 2D).

Based on the above analysis, we synthesized six pairs of 60 bp long oligonucleotides harboring the putative CTCF binding sites (Fig. 3A). These fragments were tested in gel-shift assays with nuclear extracts prepared from chicken RBCs (Fig. 3B). Using the FII site from the chicken β-globin insulator as a positive control and a specific CTCF competitor, we found that all the predicted sites are able to produce a retarded complex to different extents, similar to the one seen with FII. These retarded complexes were efficiently competed with 100-fold molar excess of the FII site (Fig. 3B). Similarly, the CTCF-binding motif from the DMD region of the mouse *Igf2/H19* locus competed the retarded bands formed by the M9 motif in a specific fashion, while the oligonucleotide bearing an Sp1 binding site behaved similarly to a non-specific competitor (data not shown).

To confirm our conclusions, the extracts with enriched CTCF concentrations [prepared from COS-1 cells transiently transfected with a chicken CTCF expressing plasmid (pSG5-CTCF expression vector)] were made. Two parallel sets of band-shift experiments were carried out using extracts from COS-1 cells transfected with empty pSG5 and pSG5-CTCF plasmids. The results of gel-shift experiments confirmed the binding of CTCF to the M9 and C10-C14 sequences (Fig. 4A). Indeed, with all binding sequences tested the intensity of one of the retarded bands (indicated by arrows in Fig. 4A) increased significantly when the CTCF-rich extracts were used, and consistently those bands were all effectively competed by the cold FII site. Based on our observation demonstrating stronger enhancer blocking activity of the αEHS-1.4 element (Fig. 2C), we decided to focus our attention on the characterization of the M9 binding site. Thus, to further confirm that CTCF binds to the M9 motif, we carried out a gel-shift assay using the FII and M9 sequences as probes and *in vitro* transcribed/translated full-length CTCF (Fig. 4B). A specific retarded complex was observed in each case and the retarded band was competed, although with different efficiencies, by the FII, DMD and M9 sequences. To estimate differences in CTCF-binding affinities between the FII and M9 sites, competitions were done at low molar excess of cold competitors (25-fold). These experiments demonstrated that, at low concentrations, FII was not able to efficiently compete CTCF bound to the M9 sequence. In contrast, M9 clearly competed CTCF bound to FII (Fig. 4B, compare lane 4 and 10). This observation requires further investigation but may reflect differences in binding affinities or in the mode of binding that can correlate with differences in enhancer blocking strength of different insulators, as recently observed for the human and mouse β-globin insulators (25). In those domains the enhancer blocking activity is, at least partially, dependent on CTCF affinity for these DNA binding sequences (25). Thus we confirmed the *in vitro* interaction of CTCF with the M9 site of the αEHS-1.4 enhancer blocking element.

Enhancer blocking activity depends on CTCF binding sites and additional sequences

To further study a possible role of CTCF in enhancer blocking activity, we tested the binding capacity of the M9 site alone and three mutant versions of this element lacking important conservative elements (Fig. 5A and B). The specificity of CTCF binding to M9 was confirmed by loss of CTCF interaction to the mutant M9 motif (M9Δcag, M9Δctag and the double mutant M9ΔcagΔctag) by gel-shift assay (Fig. 5A and data not shown). In order to functionally characterize the M9 CTCF binding site we used luciferase gene as a reporter, allowing better quantitation of the results. The β-globin insulator FII site known to bind CTCF, the FII/FIII sites in single and four copies and the mutant versions of those sites (Fig. 5B, constructs 6–10) were used as controls. The results presented in Figure 5B demonstrate that the M9 site alone (construct 5) possesses a prominent enhancer blocking effect, although less strong than the enhancer blocking effect of the FII site. Most importantly, mutations abolishing binding of CTCF to the M9 site (Fig. 5A) resulted in a complete loss of the enhancer blocking activity (constructs 2–4). Thus, the enhancer blocking activity of the αEHS-1.4 element is dependent at least in part on CTCF.

In the next set of experiments the enhancer blocking activity of the CTCF binding sites present in the 1.2 kb DNA fragment was studied (Fig. 5C). As expected (see Fig. 2B) the central 600 bp fragment did not possess any positional enhancer blocking activity (Fig. 5C, construct αCHS-0.6). In contrast, when either 5′ or 3′ flanking sequences were added to this fragment, we recovered a significant enhancer blocking

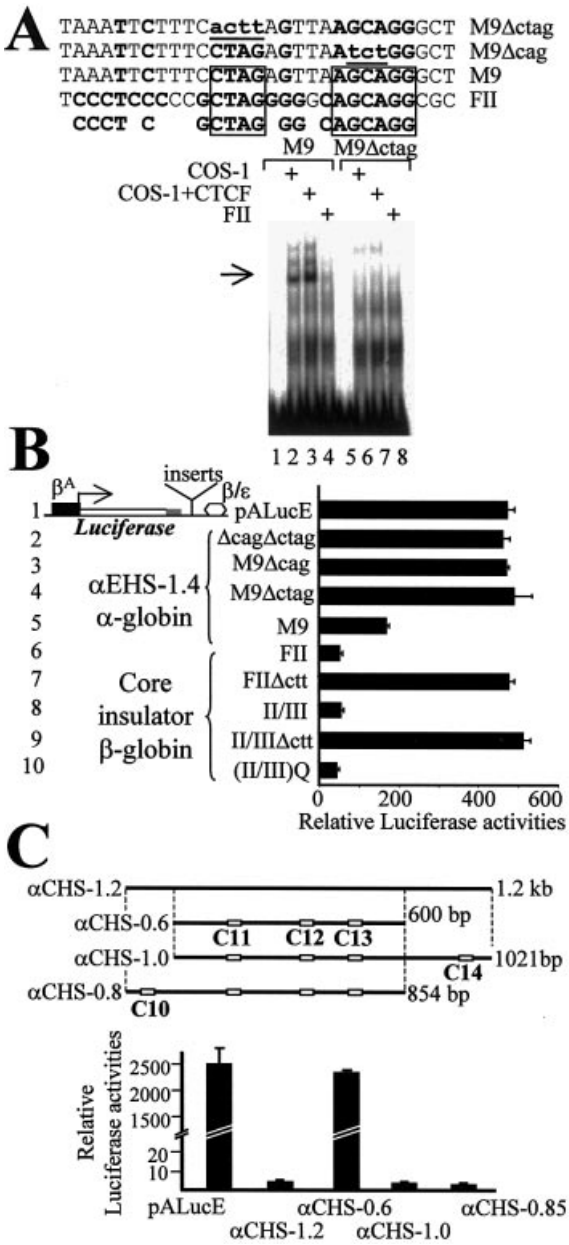

Figure 5. Detailed analysis of the enhancer blocking activities of the M9 site in the αEHS-1.4 and αCHS-1.2 enhancer-blocking elements using linearized pALucE luciferase reporter plasmids transiently transfected in chicken RBCs. (**A**) CTCF binding to the M9 and mutant M9Δctag sites. Each experiment was carried out with extracts from normal COS-1 cells and CTCF-overexpressing COS-1 cells. Lanes 4 and 8 show competition with 100-fold molar excess of the FII motif. (**B**) Enhancer blocking activity of the M9 site of the αEHS-1.4 element. Results are presented as relative luciferase activities and error bars show the SEM. Data from constructs 1–5 are the average of seven independent experiments and constructs 6–10 of four. (**C**) Sub-division of the αCHS-1.2 fragment harboring the second enhancer blocking element. In the scheme above the histogram, putative CTCF-binding elements are shown with open boxes indicated as C10–C14. Data in the histogram represent an average from six independent experiments.

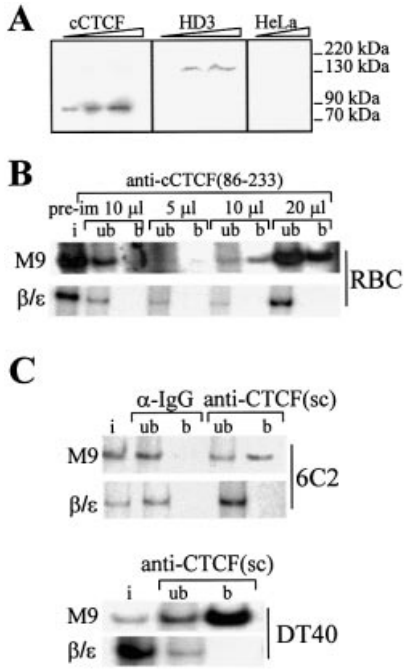

Figure 6. *In vivo* interaction of CTCF with the M9 binding site in erythroid and non-erythroid cells. To further confirm the interaction of CTCF with the M9 site, a ChIP experiment was carried out. (**A**) Characterization of the chicken cCTCF(86–233) antibody by western blot. Bacterially produced peptide [cCTCF(86–233)] and full-length peptide from total HD3 and HeLa cell extracts. (**B**) ChIP assay using chicken RBCs. PCR amplification of the test fragments (M9 and β^A/ε enhancer) using as template DNA input fraction (i) and DNA recovered from immunoprecipitated fractions bound (b) versus unbound (ub) by the chicken CTCF-specific antibody. In control experiments a preimune serum was used (pre-im, 10 µl). Note the increase of PCR-amplified M9 test fragment in bound fractions obtained with increased amount of CTCF-specific antibody. (**C**) ChIP assay in the same conditions except that the ChIP was done with a commercial CTCF antibody [anti-CTCF(sc)]. CTCF binds to the M9 site in chicken pre-erythroblast 6C2 cells and in non-erythroid DT40 cells. An irrelevant antibody was also used for ChIP showing no amplification product (α-IgG). These data are representative of three independent experiments.

activity equivalent to the activity of the entire αCHS-1.2 element (Fig. 5C, constructs αCHS-1.0, αCHS-0.85 and αCHS-1.2). These results suggest that the C10 and C14 CTCF binding sites, possibly in combination with other sites, are relevant for the overall αEHS-1.2 enhancer blocking activity.

CTCF binds *in vivo* to the M9 site

To find out if CTCF is bound *in vivo* to the M9 sequence, a ChIP assay was employed. To carry out these experiments, we first produced chicken-specific polyclonal antibodies against a non-conserved region of chicken CTCF, adjacent to the zinc-fingers [cCTCF(86–233); Fig. 6A]. Antibody specificity was confirmed by western blot (Fig. 6A). These antibodies were used to precipitate the chromatin fragments bearing bound CTCF. The concentration of the M9 motif in the precipitated fraction was compared with the concentration of a chicken reference sequence, the chicken β^A/ε enhancer element which does not interact with CTCF *in vivo* (21,26). The relative representations of both fragments in the immunoprecipitated DNA fraction were assayed using PCR with two sets of primers permitting the amplification of a 120 bp DNA fragment bearing the M9 motif and a 301 bp DNA fragment bearing the β^A/ε enhancer (Fig. 6B and C). The ChIP assay

was first carried out using chicken RBCs. The fragment bearing the M9 motif was clearly enriched (as compared to the fragment bearing the β^A/ε enhancer) in the chromatin fraction precipitated by our anti-cCTCF(86–233) antibody (Fig. 6B). Importantly, when 10 μl of pre-immune fraction was used, no enrichment was observed. The same results were obtained when commercial anti-CTCF antibodies were used (Fig. 6C). ChIP was repeated with the chicken pre-erythroblast 6C2 transformed cell line and with non-erythroid DT40 cells (pre-B cell line). These experiments demonstrated that CTCF is bound *in vivo* to the M9 motif in all the cell types tested (Fig. 6B and C).

DISCUSSION

This study describes the identification and characterization of two new enhancer blocking elements located in the upstream area of the chicken α-globin domain with a possible function to separate control elements within the same open chromatin domain or sub-domains. It should be underlined that the present knowledge about vertebrate insulators is based on the results of detailed analysis of only a few examples (7,10,11). Characterization of any new insulators may significantly extend our knowledge about these important genomic elements. This seems to be the case as far as the results of the present study are concerned. The chicken α-globin domain HSs are composed of two separable enhancer blocking activities. This is a distinguishing feature of the chicken α-globin domain upstream insulator, as compared to the insulators present at upstream and downstream flanks of the vertebrate β-globin domains (7,8). The second unexpected, and thus distinguishing, feature of one of the two enhancer blocking elements described in the present work is that it co-localizes with an initially defined erythroid-specific HS. There are previously described insulators which co-localize with tissue-specific HSs such as the mouse 3′HS1 (25). We thus decided to characterize more precisely the enhancer blocking element localized in the erythroid-specific HS. This element (referred to throughout this paper as αEHS-1.4, including the M9 motif) turned out to be a typical vertebrate enhancer blocker whose activity depends on the CTCF recognition sequence, which was found associated *in vivo* with the M9 motif in both immature erythroid and non-erythroid cells (DT40 cell line). It is also worth mentioning that this element co-localizes with the previously defined MAR sequence (15). Co-localization with MAR elements has been noted for some vertebrate insulators studied (2,17,19). Moreover it has recently been described that CTCF is able to bind to the nuclear matrix; the fact that CTCF is found *in vivo* in the DT40 cell line could mean that it plays a relevant structural role in the MAR function in addition to its enhancer blocking activity (27).

To understand a possible significance of the location of an enhancer blocker element in an erythroid-specific HS, one should consider special features of α-globin domains as compared to β-globin domains (2,12,28). In contrast to the latter, the α-globin domains are located in permanently open chromatin areas which contain several non-globin genes, including a housekeeping gene located very close to the α-globin gene cluster and overlapping the α-globin gene domain (14). To this end the original prediction is that the α-globin domain constitutes a typical example of weak genomic domains that do not have defined structural boundaries (2). Nevertheless, the transcriptional status of the domain as a whole is regulated by special, although as yet poorly understood mechanisms. Operation of these domain-level regulatory systems became especially clear when it was demonstrated that the domain as a whole (with flanking sequences) is characterized by increased levels of site-specific histone acetylation in erythroid cells but not in cells of other lineages (13,29,30). The LCR-like regulatory element of mammalian α-globin domains was found 40 kb upstream of the first α-globin gene in one of the so-called '–14' gene introns (14,31). Analysis of the distribution of conserved regions in the upstream areas of vertebrate α-globin domains has permitted the prediction of the positions of similar regulatory elements in the chicken genome [about 20 kb upstream to π gene (32)]. The intriguing fact is that this regulatory region is located upstream of the enhancer blockers characterized in the present study. If indeed this regulatory region is involved in control of α-globin gene expression, the enhancer blockers should be somehow inactivated in cells expressing α-globin genes. Conditional insulators have been described in the literature (10,11,33,34). In the majority of the known cases, their activity is regulated by DNA methylation, which interfere with the CTCF binding. It is, however, easy to imagine a situation when CTCF binding would be prevented by binding of other sequence-specific factors (in our case erythroid-specific transcription factors or ubiquitous factors) close to the CTCF recognition site. To this end it might be of importance that the genomic region studied in the present work contains multiple recognition motifs for different erythroid-specific factors (15). However, our results indicate that CTCF is capable of interacting with the M9 recognition site by itself according to the *in vitro* transcription/translation results, and also in the presence of other proteins according to the gel-shift experiments performed with erythroid nuclear extracts.

Although the second enhancer blocking element (αCHS-1.2) identified in our study co-localizes roughly with the constitutive HS, the short fragment (αCHS-0.6) containing this constitutive HS does not possess enhancer blocking activity by itself, and binding of CTCF to the recognition motifs located at the flanks of the αCHS-1.2 DNA fragment might be affected by binding of erythroid-specific transcription factors to the recognition sequences present within the two erythroid-specific HSs. If indeed erythroid-specific transcription factors can compete with the CTCF binding, a simple increase in the concentration of these factors occurring in the course of differentiation of erythroblast precursors would inactivate the enhancer blocker.

An alternative model has to do with the location of the newly defined enhancer blockers, and the bi-directional presence of intergenic transcripts over the studied area (Fig. 1) (14,16). One possible speculation in terms of the location and presence of two enhancer blockers in the 5′ upstream area of the chicken α-globin domain is related to those intergenic transcripts. Each opposing transcript should have regulatory elements that drive their initiation and elongation. Based on such supposition, the newly defined αEHS-1.4 and αCHS-1.2 elements could block non-specific regulatory signals affecting those transcripts. Such an idea

seems attractive since one insulator could act on the 5′ strand and the other on the opposite strand based on the transcript directions and we have demonstrated that the enhancer blocking activity of these elements is independent of their orientation (see Fig. 2) (14,16). There are experiments currently in progress in our laboratory to test this hypothesis. In addition, we should recall that there is recent evidence supporting the presence of intergenic transcripts in chromatin domains and sub-domains that contribute to their formation and developmental regulation (35,36).

Interestingly, detailed comparative analysis of histone acetylation levels on the 5′ side of the human, mouse, pufferfish and chicken α-globin gene clusters have revealed histone acetylation transitions (29). For the chicken α-globin domain this group has found the highest levels of H4 Lys-5 and H3 Lys-14 hyperacetylation overlapping the αEHS-1.4 and αCHS-1.2 enhancer blocking elements. They have defined an area with a high level of histone acetylation, which extends from 6 kb upstream of the αEHS-1.4 [primer C6 (29)] through 200 bp downstream of the αCHS-1.2 element [primer C7 (29)]. This area overlaps the three HSs under study but the transition on histone hyperacetylation is located 6 kb upstream of the αEHS-1.4 (primer C6). Taking all this data into account we can conclude that the newly defined enhancer blockers are located within a zone of high levels of histone acetylation and 6 kb downstream of a sharp histone acetylation transition which could correspond to the boundary of the domain [see figures 6 and 7 from Anguita *et al*. (29)].

At this point we can only speculate how these enhancer blocking elements work in the chicken α-globin domain. Nevertheless, the proposed models may constitute a starting point to design further experiments aimed at better understanding the regulation of α-globin gene expression.

ACKNOWLEDGEMENTS

We are indebted to the support of Alicia González Manjarrez, Jaime Mas-Oliva and members of their laboratories, Gary Felsenfeld for the chicken cHS4 β-globin insulators plasmids, and Amy Weinmann and Peggy Farnham for ChIP advice. We are also grateful to Catherine M. Farrell, Paul Delgado-Olguín and Héctor Rincón-Arano for critical reading of the manuscript, useful suggestions and constant support. We thank Georgina Guerrero Avendaño for excellent technical assistance. This work was supported by the Dirección General de Asuntos del Personal Académico-UNAM (IN203200), the Consejo Nacional de Ciencia y Tecnología, CONACyT (33863-N), the Third World Academy of Sciences (TWAS, Grant 01-055 RG/BIO/LA) and Fundación Miguel Alemán. V.V.-G. was supported by a CONACyT Ph.D. fellowship. The work of S.V.R. was supported by a Grant on Physico-Chemical Biology given by the Presidium of the Russian Academy of Sciences.

REFERENCES

1. Recillas-Targa,F. and Razin,S.V. (2001) Chromatin domains and regulation of gene expression: familiar and enigmatic clusters of chicken globin genes. *Crit. Rev. Eukaryot. Gene Exp.*, **11**, 227–242.
2. Razin,S.V., Farrell,C.M. and Recillas-Targa,F. (2003) Genomic domains and regulatory elements operating at the domain level. *Int. Rev. Cytol.*, **226**, 63–125.
3. Gerasimova,T.I., Byrd,K. and Corces,V.G. (2000) A chromatin insulator determines the nuclear localization of DNA. *Mol. Cell*, **6**, 1025–1035.
4. Cremer,T. and Cremer,C. (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.*, **2**, 292–301.
5. Ishii,K., Arib,G., Lin,C., Van Houwe,G. and Laemmli,U.K. (2002) Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell*, **109**, 552–562.
6. West,AG., Gaszner,M. and Felsenfeld,G. (2002) Insulators: many functions, many mechanisms. *Genes Dev.*, **16**, 271–288.
7. Burgess-Beusse,B., Farrell,C., Gaszner,M., Litt,M., Mutskov,V., Recillas-Targa,F., Simpson,M., West,A. and Felsenfeld,G. (2002) The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl Acad. Sci. USA*, **99** Suppl 4, 16433–16437.
8. Bell,A.C., West,A.G. and Felsenfeld,G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, **98**, 387–396.
9. Ohlsson,R., Renkawitz,R. and Lobanenkov,V. (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.*, **17**, 520–527.
10. Hark,A.T., Schoenherr,C.J., Katz,D.J., Ingram,R.S., Levorse,J.M. and Tilghman,S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature*, **405**, 486–489.
11. Bell,A.C. and Felsenfeld,G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature*, **405**, 482–485.
12. Recillas-Targa,F. (2000) The chicken α- and β-globin gene domains and their chromatin organization. *Cell. Mol. Biol. Lett.*, **5**, 451–467.
13. Litt,M.D., Simpson,M., Recillas-Targa,F., Prioleau,M.-N. and Felsenfeld,G. (2001) Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J.*, **20**, 2224–2235.
14. Vyas,P., Vickers,M.A., Picketts,D.J. and Higgs,D.R. (1995) Conservation of position and sequence of a novel, widely expressed gene containing the major human α-globin regulatory element. *Genomics*, **29**, 679–689.
15. Razin,S.V., Shen,K., Ioudinkova,E. and Scherrer,K. (1999) Functional analysis of DNA sequences located within a cluster of DNase I hypersensitive sites colocalizing with a MAR element at the upstream border of the chicken α-globin gene domain. *J. Cell. Biochem.*, **74**, 38–49.
16. Sjakste,N., Iarovaia,O.V., Razin,S.V., Linares-Cruz,G., Sjakste,T., Le Gac,V., Zhao,Z. and Scherrer,K. (2000) A novel gene is transcribed in the chicken α-globin gene domain in the direction opposite to the globin genes. *Mol. Gen. Genet.*, **262**, 1012–1021.
17. Stief,A., Winter,D.M., Strätling,W.H. and Sippel,A.R. (1989) A nuclear DNA attachment element mediates elevated and position-independent gene activity. *Nature*, **341**, 343–345.
18. Recillas Targa,F., Razin,S.V., De Moura Gallo,C.V. and Scherrer,K. (1994) Excision close to matrix attachment regions of the entire chicken α-globin gene domain by nuclease S1 and characterization of the framing structures. *Proc. Natl Acad. Sci. USA*, **91**, 4422–4426.
19. Antes,T.J., Namciu,S.J., Fournier,R.E.K. and Levy-Wilson,B. (2001) The 5′ boundary of the human apolipoprotein B chromatin domain in intestinal cells. *Biochemistry*, **40**, 6731–6742.
20. Recillas-Targa,F., Bell,A.C. and Felsenfeld,G. (1999) Positional enhancer-blocking activity of the chicken β-globin insulator in transiently transfected cells. *Proc. Natl Acad. Sci. USA*, **96**, 14354–14359.
21. Reitman,M. and Felsenfeld,G. (1988) Mutational analysis of the chicken β-globin enhancer reveals two positive-acting domains. *Proc. Natl Acad. Sci. USA*, **85**, 6267–6271.
22. Prioleau,M.-N., Nony,P., Simpson,M. and Felsenfeld,G. (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.
23. Beug,H., Palmieri,S., Freudenstein,C., Zentgraf,H. and Graf,T. (1982) Hormone-dependent terminal differentiation in vitro of chicken erythroleukemia cells transformed by ts mutants of avian erythroblastosis virus. *Cell*, **28**, 907–919.
24. Weinmann,A.S., Yan,P.S., Oberley,M.J., Huang,T.H. and Farnham,P.J. (2002) Isolating human transcription factor targets by coupling chromatin

immunoprecipitation and CpG island microarray analysis. *Genes Dev.*, **16**, 235–244.
25. Farrell,C.M., West,A.G. and Felsenfeld,G. (2002) Conserved CTCF insulator elements flank the mouse and human β-globin loci. *Mol. Cell. Biol.*, **22**, 3820–3831.
26. Recillas-Targa,F., Pikaart,M.J., Burgess-Beusse,B., Bell,A.C., Litt,M.D., West,A.G., Gaszner,M. and Felsenfeld, G. (2002) Position-effect protection and enhancer blocking by the chicken β-globin insulator are separable activities. *Proc. Natl Acad. Sci. USA*, **99**, 6883–6888.
27. Dunn,K.L., Zhao,H. and Davie,J.R. (2003) The insulator binding protein CTCF associates with the nuclear matrix. *Exp. Cell Res.*, **288**, 218–223.
28. Knezetic,J.A. and Felsenfeld,G. (1989) Identification and characterization of a chicken α-globin enhancer. *Mol. Cell. Biol.*, **9**, 893–901.
29. Anguita,E., Johnson,C.A., Wood,W.G., Turner,B.M. and Higgs,D.R. (2001) Identification of a conserved erythroid specific domain of histone acetylation across the α-globin gene cluster. *Proc. Natl Acad. Sci. USA*, **98**, 12114–12119.
30. Noma,K.-I., Allis,D.C. and Grewal,S.I.S. (2001) Transition in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science*, **293**, 1150–1155.
31. Higgs,D.R., Wood,W.G., Jarman,A.P., Sharpe,J., Lida,J., Pretorius,I.-M. and Ayyub,H. (1990) A major positive regulatory region located far upstream of the human α-globin gene locus. *Genes Dev.*, **4**, 1588–1601.
32. Flint,J., Tufarelli,C., Peden,J., Clark,K., Daniels,R.J., Hardison,R., Miller,W., Philipsen,S., Tan-Un,K.C., McMorrow,T. *et al*. (2001) Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. *Hum. Mol. Genet.*, **15**, 371–382.
33. Chao,W., Huynh,K.D., Spencer,R.J., Davidow,L.S. and Lee,J.T. (2002) CTCF, a candidate *trans*-acting factor for X-inactivation choice. *Science*, **295**, 345–347.
34. Lutz,M., Burke,L.J., LeFevre,P., Myers,F.A., Thorne,A.W., Crane-Robinson,C., Bonifer,C., Filippova,G.N., Lobanenkov,V. and Renkawitz,R. (2003) Thyroid hormone-regulated enhancer blocking cooperation of CTCF and thyroid hormone receptor. *EMBO J.*, **22**, 1579–1587.
35. Ashe,H.L., Monks,J., Wijgerde,M., Fraser,P. and Proudfoot,N.J. (1997) Intergenic transcription and transinduction of the human β-globin locus. *Genes Dev.*, **11**, 2494–2509.
36. Gribnau,J., Diderich,K., Pruzina,S., Calozolari,R. and Fraser,P. (2000) Intergenic transcription and developmental remodeling of chromatin subdomains in the human β-globin locus. *Mol. Cell*, **5**, 377–386.