

Looping and Interaction between Hypersensitive Sites in the Active β -globin Locus

Bas Tolhuis,² Robert-Jan Palstra,² Erik Splinter,
Frank Grosveld, and Wouter de Laat¹
Department of Cell Biology and Genetics
Faculty of Medicine
Erasmus University, Rotterdam
P.O. Box 1738
3000DR Rotterdam
The Netherlands

Summary

Eukaryotic transcription can be regulated over tens or even hundreds of kilobases. We show that such long-range gene regulation *in vivo* involves spatial interactions between transcriptional elements, with intervening chromatin looping out. The spatial organization of a 200 kb region spanning the murine β -globin locus was analyzed in expressing erythroid and nonexpressing brain tissue. In brain, the globin cluster adopts a seemingly linear conformation. In erythroid cells the hypersensitive sites of the locus control region (LCR), located 40–60 kb away from the active genes, come in close spatial proximity with these genes. The intervening chromatin with inactive globin genes loops out. Moreover, two distant hypersensitive regions participate in these interactions. We propose that clustering of regulatory elements is key to creating and maintaining active chromatin domains and regulating transcription.

Introduction

Transcriptional activation in higher eukaryotes frequently involves the long-range action of a number of regulatory DNA elements. Although this has been recognized for more than 20 years, it is still not clear how enhancers (Banerji et al., 1981; Wasylyk et al., 1983), LCRs (Grosveld et al., 1987), or insulators/boundaries (Kellum and Schedl, 1991; Dorsett, 1999; Gerasimova and Corces 2001; West et al., 2002) exert their effect on the process of chromatin modification and transcription over distance (up to hundreds of kilobases). Many different models have been put forward to explain distant effects. The looping model states that enhancers and promoters communicate through direct interactions between proteins bound to the DNA elements, with the intervening DNA looping out (Ptashne, 1986; Mueller and Schaffner, 1990; Hanscombe et al., 1991). Other models imply a role for the DNA in between to support the transmission of some signal from enhancer to promoter. Direct support for the latter type of models comes from bacteria. Here, activation of the phage T4 late genes was found to involve loading on and sliding from the enhancer of trimeric gp45 along the DNA to the promoter to allow the forming of the transcription initiation complex (Herendeen et al., 1992). The looping model also

receives support from studies on transcriptional regulation of many different prokaryotic genes. In fact, the model was originally based on work on bacterial and phage repressor proteins, like the Gal, AraC, and λ repressor proteins, which were found to function only when homomultimerized and bound to two separate operator sites. Electron microscopy visually demonstrated the DNA in between to loop out (reviewed in Ptashne, 1986). Thus, both types of mechanisms appear to function in bacteria. Eukaryotes have more complex gene clusters with regulatory elements functioning over much greater distances. To date, there are no data that unambiguously demonstrate one (or more or combinations) of the models to be correct for the regulation of a given eukaryotic locus. Support for models has come from indirect and/or *in vitro* observations, and often the distinction between the activation and actual transcription of a locus is not made. However with respect to transcription, a number of observations can only easily be explained by the looping model. The first type of experiments involves studies on *trans*-activation, i.e., the ability of an enhancer to activate a promoter present on a physically separate DNA molecule. Most important in this respect is the naturally occurring phenomenon of transvection in *Drosophila* (Bickel and Pirrotta, 1990). In addition, Schaffner and coworkers demonstrated *in vitro* that enhancers can stimulate transcription in *trans*, by coupling an enhancer- to a promoter-containing plasmid via a biotin-streptavidin bridge (Mueller et al., 1989). Similarly, *trans*-activation of transcription was observed when enhancer- and promoter-containing plasmids were injected as intertwined catenates into frog oocytes (Dunaway and Droge, 1989). More recently, transient transfection assays with reporter plasmids and GAGA as a DNA-bridging factor also demonstrated transcriptional activation in *trans* in mammalian cells (Mahmoudi et al., 2002). All these studies on *trans*-activation demonstrate that a *cis* configuration of enhancer and promoter is not an absolute prerequisite for interaction, as predicted only by the looping model.

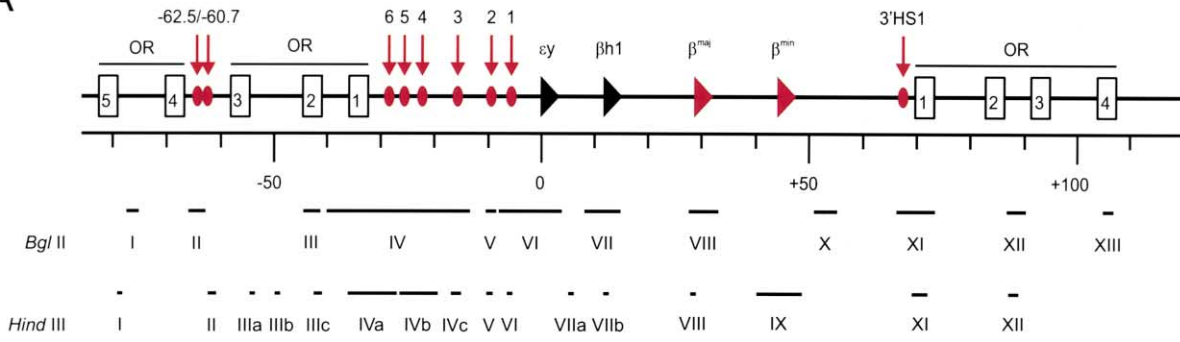
In addition, gene competition for a single regulator (Wasylyk et al., 1983; de Villiers et al., 1983; Hanscombe et al., 1991), leading to alternate transcription (Wijgerde et al., 1995; Gribnau et al., 1998; Trimborn et al., 1999), is also most easily explained by looping, particularly because the competitive advantage of the enhancer-proximal gene is lost when the genes are closely spaced at further distance from the regulator (Hanscombe et al., 1991; Heuchel et al., 1989; Dillon et al., 1997). Finally, in yeast, a downstream enhancer was recently demonstrated to activate gene expression from a distance by making use of loops induced by telomeres (de Bruin et al., 2001). However all these experiments were either done *in vitro* or are indirect in nature. None of them directly shows *in vivo* that two distal elements linked in *cis* interact by coming in close spatial proximity with intervening DNA looping out.

Here, we provide evidence that looping occurs during transcription *in vivo*. We demonstrate that the murine β -globin LCR is in physical proximity to the active globin

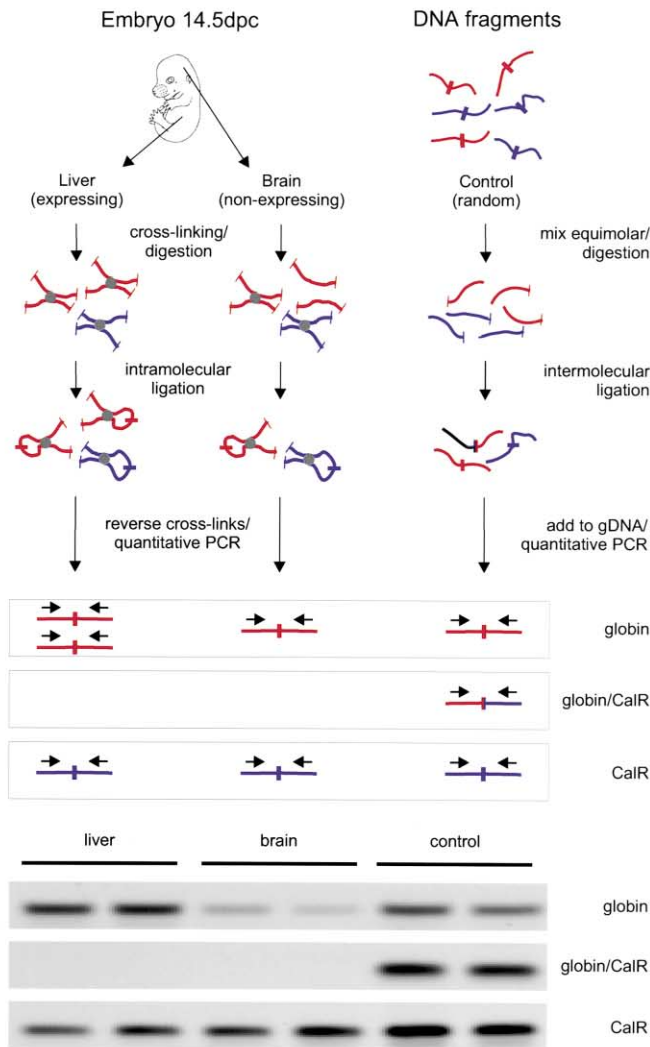
¹Correspondence: delaata@ch1.fgg.eur.nl

²These authors contributed equally to this work.

A



B



C

$$X(\text{gl}) = \frac{[A(\text{gl}) / A(\text{CaIR})]_{\text{tissue}}}{[A(\text{gl}) / A(\text{CaIR})]_{\text{control}}}$$

genes in vivo in expressing tissue with the intervening DNA looping out. Interaction and looping are not observed in nonexpressing tissue. In addition, DNase I hypersensitive sites at both ends of the locus participate in these interactions, again by looping out intervening DNA. Thus, multiple hypersensitive sites spread over 130 kilobases interact to form a cluster in the nuclear space. On the basis of these data we propose that direct interactions between distal DNase I hypersensitive sites and looping out of chromatin is crucial in establishing an open chromatin domain and activating transcription.

Results

Applying 3C Technology to the Murine β -globin Locus

We applied methodology recently developed by Dekker et al. (2002) to gain insight into long-range interactions between the LCR and the genes in the murine β -globin locus. The principle of this technique, chromosome conformation capture (3C), is that cells are treated with formaldehyde to crosslink proteins to other proteins nearby and DNA (see also Figure 1B). The resulting DNA-protein network is then subjected to cleavage by a restriction enzyme, which is followed by ligation at low DNA concentration. Under such conditions, ligations between crosslinked DNA fragments, which are intramolecular, is strongly favored over ligations between random fragments, which are intermolecular (Dekker et al., 2002). After ligation, the crosslinks are reversed and ligation products are detected and quantified by polymerase chain reaction (PCR). The crosslinking frequency of two specific restriction fragments, as measured by the amount of corresponding ligation product, is proportional to the frequency with which these two genomic sites interact (Dekker et al., 2002). Thus, 3C analysis provides information about the spatial organization of chromosomal regions in vivo.

A schematic presentation of the murine β -globin locus is given in Figure 1A. In brief, the locus contains an LCR, comprising six HS (5' HS1-6), two embryonic genes, ϵ and β h1 (expressed in the yolk sac), and two adult genes, β -major and β -minor (expressed in fetal liver and adult spleen/bone marrow). The LCR is required for high levels of expression of all β -globin genes. Similar to what is observed in the human β globin locus, the murine

β -globin locus is flanked by olfactory receptor (OR) genes, which are inactive in globin-expressing erythroid tissue (Bulger et al., 1999; 2000). Also similar to the human locus are the strong erythroid-specific DNase I HS at the 3' side (3' HS1) between β -minor and the OR genes and two closely spaced HS (HS-60.7 and HS-62.5) at the far 5' side located between 5' OR3 and 5' OR4 (Farrell et al., 2000).

Two independent sets of restriction fragments (BglII and HindIII fragments, respectively) were used for 3C analysis of the β -globin locus. Each set covers the 200 kb region depicted in Figure 1A, with intervals between analyzed DNA fragments of approximately 20 kilobases or smaller. Analysis was performed on 14.5 dpc mouse fetal livers, which express the most distal globin genes, β -major and β -minor. Brain from the same 14.5 dpc embryos was simultaneously analyzed as a nonexpressing control tissue.

A number of experimental controls were included. First, we checked the efficiency of restriction enzyme digestion. Southern blotting and PCR analysis showed that the restriction sites analyzed were cleaved without any preference for any particular region(s) after overnight incubation with an excess of enzyme (data not shown). Second, we determined the range of amount of template that shows linear PCR product formation. Similar ranges were found with both liver and brain template (data not shown), and roughly equal amounts (\sim 300 ng DNA template per reaction) were used in all subsequent experiments. Third, to correctly interpret signal intensities obtained with a given primer set by quantitative PCR, one needs to correct for the PCR amplification efficiency of that set. Thus, a control template is required in which all possible ligation products are present in equimolar amounts. In yeast, this was done by digesting and randomly ligating noncrosslinked genomic DNA (Dekker et al., 2002). For mammalian cells, with a genome one hundred times the size of the yeast genome, we found that random ligation of two specific loci is too rare an event to be detected by PCR. We therefore enriched for ligation products of interest by mixing equimolar amounts of DNA fragments that span each of the restriction sites analyzed (see Figure 1B). After digestion and ligation, this mix was added to genomic DNA to serve as a control template (see also Experimental Procedures). As a result, the crosslinking fre-

Figure 1. 3C Technology in the Murine β -globin Locus

(A) Schematic presentation of the murine β -globin locus. Red arrows and ellipses depict the individual HS. The globin genes are indicated by triangles, with active genes (β^{maj} and β^{min}) in red and inactive genes (ϵ and β h1) in black. The white boxes indicate the olfactory receptor (OR) genes (5' OR1-5 and 3' OR1-4). The two sets of restriction fragments (BglII and HindIII) that were used for 3C analysis are shown below the locus. The individual fragments are indicated by Roman numerals. Identical numbering between BglII and HindIII indicates that two fragments colocalize. Distances are in kb counting from the site of initiation of the ϵ gene.

(B) Schematic outline of the 3C analysis. Globin fragments (red), CaIR fragments (blue), restriction sites (perpendicular bars on fragments), crosslinks, and PCR primers are indicated. Examples of PCR results (always done in duplo) show products obtained with HindIII globin fragments VIII and IV-b (top), globin fragment VIII and one of the HindIII CaIR fragments (middle), and the two HindIII CaIR fragments (bottom). Tissue lanes in the middle panel were always empty, with every globin fragment tested. The CaIR products (bottom) were used for normalizing signals.

(C) Equation used to calculate relative crosslinking frequency between two given globin fragments $[X(\text{gl})]$. $A(\text{gl})$: Peak area (determined with ImageQuant 5.2) of PCR signal obtained with a given globin-globin ligation product. $A(\text{CaIR})$: Peak area of PCR signal obtained with the CaIR-CaIR ligation product. $A(\text{gl})$ and $A(\text{CaIR})$ are determined for both the tissue (fetal liver or brain) and the control (random ligation; see Results). The calculation gives a relative ligation crosslinking frequency for each tissue since it corrects for differences in PCR amplification efficiencies, crosslinking and ligation efficiencies, amounts of template, and size of PCR fragments (see text). These values are plotted on the y axis in Figures 2-6 for the various globin fragments.

quency between two loci can be expressed as the ratio of signal obtained by quantitative PCR on crosslinked template versus that obtained on control template. Fourth, we measured the crosslinking and ligation efficiencies in both tissues to be able to compare crosslinking frequencies. This was done by comparing the crosslinking frequency between two restriction fragments present on an unrelated locus situated on another chromosome. Two neighboring fragments were used, with the restriction sites analyzed ~ 1.5 kilobases apart, in the transcribed part of the calreticulin locus (CalR) on chromosome 8 (the β -globin locus is on mouse chromosome 7). The CalR locus, embedded in an area of ubiquitously expressed genes, is expressed at similar levels in 14.5 dpc brain and liver (W.d.L., unpublished data). It is therefore reasonable to assume that it adopts a similar spatial conformation in both tissues. Thus, by normalizing each crosslinking frequency to the crosslinking frequency observed between the CalR fragments within a tissue, we could correct for differences in the amount and quality of template. Similarly, by normalizing the observed random ligation efficiency of two given fragments to that observed of the CalR fragments, we corrected for differences in the amount of control template between experiments. The equation used to calculate the relative crosslinking frequency is given in Figure 1C. As a result of this normalization, the crosslinking frequency value 1 arbitrarily corresponds to the crosslinking frequency between our control CalR fragments. Finally, the crosslinking frequencies between globin fragments and CalR fragments were always measured as an additional control. As expected for the interaction between two unrelated loci, globin-CalR crosslinking frequencies were always found to be zero (no PCR signals observed in tissues; see Figure 1B).

The Globin Locus Adopts a Linear Conformation in Nonexpressing Brain Cells

We performed 3C analysis on expressing and nonexpressing tissue from 14.5 dpc embryos to be able to relate the spatial conformation of the β -globin locus to its transcriptional status. Figure 2 shows results obtained in a nonexpressing tissue, the brain. Depicted are locus-wide crosslinking frequencies for two different BglIII fragments (fixed fragments), one in the middle of the 200 kb region (fragment V) and one at the 5' end (fragment II). The central fragment V, a relatively small fragment containing HS2 of the LCR, showed the highest crosslinking frequency with the closest fragments IV and VI. Crosslinking frequency gradually decreased with fragments located further away on the linear DNA template (Figure 2A). No significant peaks of interactions were observed between fragment V and more distal DNA fragments. Similar results were obtained for the DNA fragment at the 5' end of the region (II) (Figure 2B). Thus, in brain, we observed a direct correlation between spatial proximity and distance along the linear β -globin DNA template. This holds for any fixed fragment in this region, independent of the restriction enzyme used (see data below). Such a correlation between distance in space and distance in kilobases would be expected of a linear structure (Rippe, 2001). Hence, we conclude that the 200 kb region encompassing the β -globin locus adopts an essentially linear conformation in the nucleus of the nonexpressing brain cell.

Spatial Interaction and Looping between the LCR and the Active Genes in the Expressing Fetal Liver

Next, we analyzed the spatial organization of the β -globin locus in the expressing 14.5 dpc fetal liver cells. The active globin genes, β -major and β -minor, are 34 and 49 kb away from the 3' side of the LCR, respectively. We first focused on a BglIII fragment (fragment VIII) containing the active β -major gene with all the known local regulatory elements, including the promoter and the enhancer ~ 1 kb downstream of the transcribed sequence. In agreement with the findings presented above, the curve for brain was indicative of a linear conformation (Figure 3A). In fetal liver, crosslinking frequencies identical to those in brain were observed for fragments closest to fragment VIII. However, when DNA elements more toward the 5' side of the region were analyzed, up to 3-fold elevated crosslinking frequencies were found in liver as compared to brain with fragments IV, V, and VI. Most interestingly, these are the three BglIII fragments that together cover all six hypersensitive sites of the LCR. Beyond the LCR, even further 5' from the β -major gene, crosslinking frequencies dropped again to the levels observed in brain (with the exception of fragment II, discussed below). These data indicate that in the nucleus of the expressing fetal liver cell, the active β -major gene comes in close vicinity to the LCR.

This is confirmed when the reciprocal experiment is carried out using an LCR fragment as the fixed fragment. BglIII restriction sites flank HS2 of the LCR, resulting in fragment V. When this fragment was tested versus the others in fetal liver, fragment VIII (β -major), but also fragment X, containing DNA sequences just 3' of the active β -minor gene, showed highly elevated crosslinking frequencies in fetal liver compared to brain (Figure 3B). In fact, in fetal liver but not in brain, the crosslinking frequency between HS2 and the active adult genes is much higher than that between HS2 and the inactive, embryonic genes ($\epsilon\gamma$ and $\beta\text{h}1$, present on fragments VI and VII, respectively). Thus, these data show that in expressing cells, the β -globin LCR and the distal active genes come in physical proximity, whereas the inactive genes appear to be located further away from the LCR fragment V.

In order to determine whether in fetal liver the inactive genes indeed do not come in close proximity to other sequences in the locus, we looked at locus-wide crosslinking frequencies of the $\beta\text{h}1$ gene (fragment VII). Almost identical crosslinking frequencies between $\beta\text{h}1$ and the rest of the locus were observed in liver and in brain for both a BglIII (Figure 3C) and HindIII digest (data not shown). Similar results were obtained for a HindIII fragment close to $\epsilon\gamma$ (VII-a, data not shown and see Figures 4–6). This suggests that the inactive genes are not interacting with the LCR. We conclude that the LCR interacts specifically with the active distal β -globin genes with intervening DNA containing the inactive genes looping out.

All Hypersensitive Sites of the LCR Participate in the Long-Range Interactions

Whereas BglIII cuts relatively infrequently in the murine β -globin locus, resulting in the large fragments analyzed and described above, digestion by HindIII yields smaller DNA fragments, which may allow fine-mapping of the

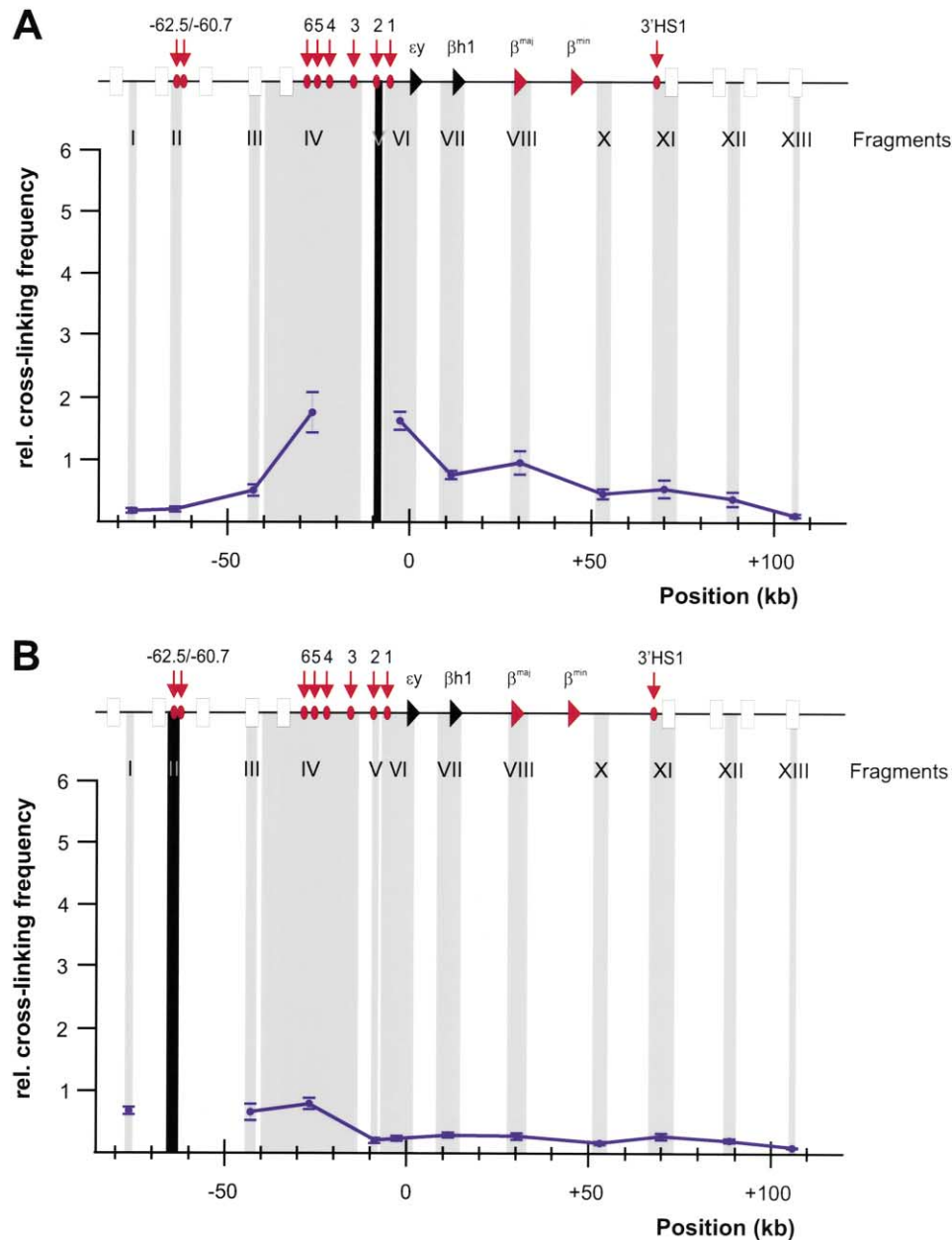


Figure 2. Linear Conformation of the β -globin Locus in Nonexpressing Brain Cells

The murine β -globin locus is depicted on top of each graph (for explanation of symbols, see Figure 1A). The x axis shows the position in the locus. The black shading shows the position and size of the fixed fragment. The gray shading indicates the position and size of other fragments. Standard error of the mean is indicated. Crosslinking frequency with a value of 1 arbitrarily corresponds to the crosslinking frequency between two neighboring *CaI*R control fragments (with restriction sites analyzed being 1.5 kb apart). Scaling on the y axis (from 0 to 6) allows direct comparison with Figures 3–6.

(A) Relative crosslinking frequencies between fixed *Bgl*III fragment V (5' HS2 in LCR) and the rest of the locus.

(B) Relative crosslinking frequencies between fixed *Bgl*III fragment II (5' HS –62.5/60.7) and the rest of the locus.

interactions. Most relevant to our studies, *Hind*III cuts in between most of the hypersensitive sites of the LCR (with the exception of HS4 and HS5, which are present on one *Hind*III fragment). Analysis of crosslinking frequencies with a fixed *Hind*III fragment VIII, containing 300 base pairs of the β -major promoter plus one-third of the coding part of this gene, confirmed the fetal liver-specific interaction with the LCR (Figure 4A). In fact, elevated crosslinking frequencies with the β -major frag-

ment were observed for all fragments containing a hypersensitive site of the LCR (fragments IV-a, -b, and -c, and fragments V and VI). As seen in the *Bgl*III experiments, crosslinking frequencies with β -major dropped for fragments flanking the LCR (again with the exception of fragment II, discussed below). Thus, the *Hind*III data indicate that all individual hypersensitive sites of the LCR (HS1–6) participate in long-range interaction. The same results were obtained with fragment IX, encom-

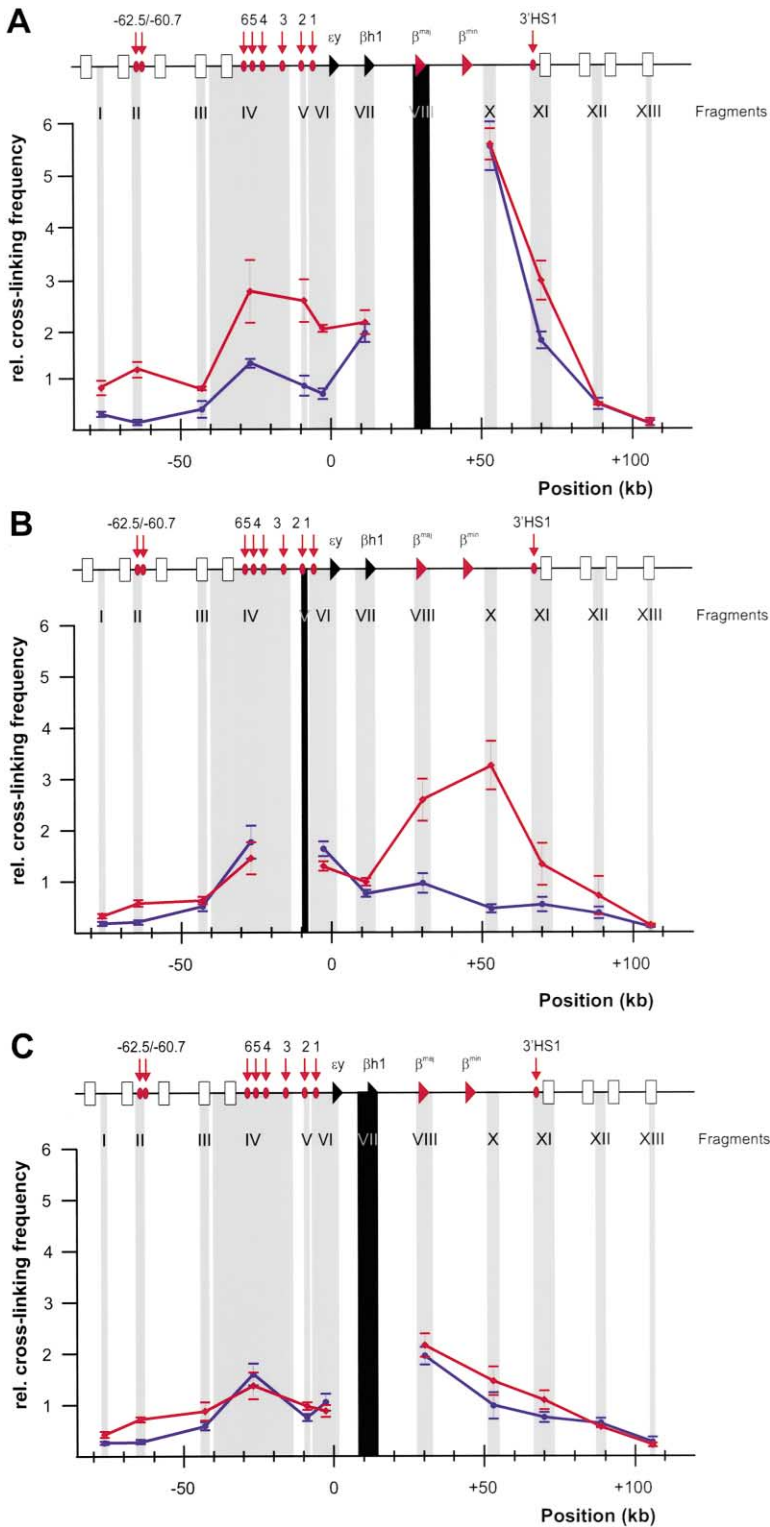


Figure 3. Erythroid-Specific Interaction and Looping between the LCR and an Active β -globin Gene

Relative crosslinking frequencies observed in fetal liver are shown in red. For comparison, data obtained in brain are depicted in blue. Standard error of the mean is indicated. Crosslinking frequency with a value of 1 arbitrarily corresponds to the crosslinking frequency between two neighboring CaIR control fragments (with restriction sites analyzed being 1.5 kb apart). Scaling on the y axis (from 0 to 6) allows direct comparison with Figures 2 and 4–6.

- (A) Fixed BglII fragment VIII (β^{ma}) versus the rest of the locus.
- (B) Fixed BglII fragment V (5'HS2) versus the rest of the locus.
- (C) Fixed BglII fragment VII ($\beta H1$) versus the rest of the locus.

passing the active β -minor gene (Figure 4B), although here the data suggest that HS2 (fragment V) and HS3 (fragment IV-c) do not participate as actively in the interaction as the other hypersensitive sites do. This may indeed be the case, but it may equally well reflect a technical problem (see Discussion). Nevertheless, these data strongly support the hypothesis that the individual hypersensitive sites of the LCR act together to contact

distal genes in the fetal liver (Ellis et al., 1993; Wijgerde et al., 1995).

If indeed the LCR forms one spatial entity in expressing cells, tissue-specific high crosslinking frequencies among the individual hypersensitive sites of the LCR would be expected. This is indeed what we observe. For example, taking HS2 (fragment V) as the fixed fragment, we found fetal liver-specific high crosslinking fre-

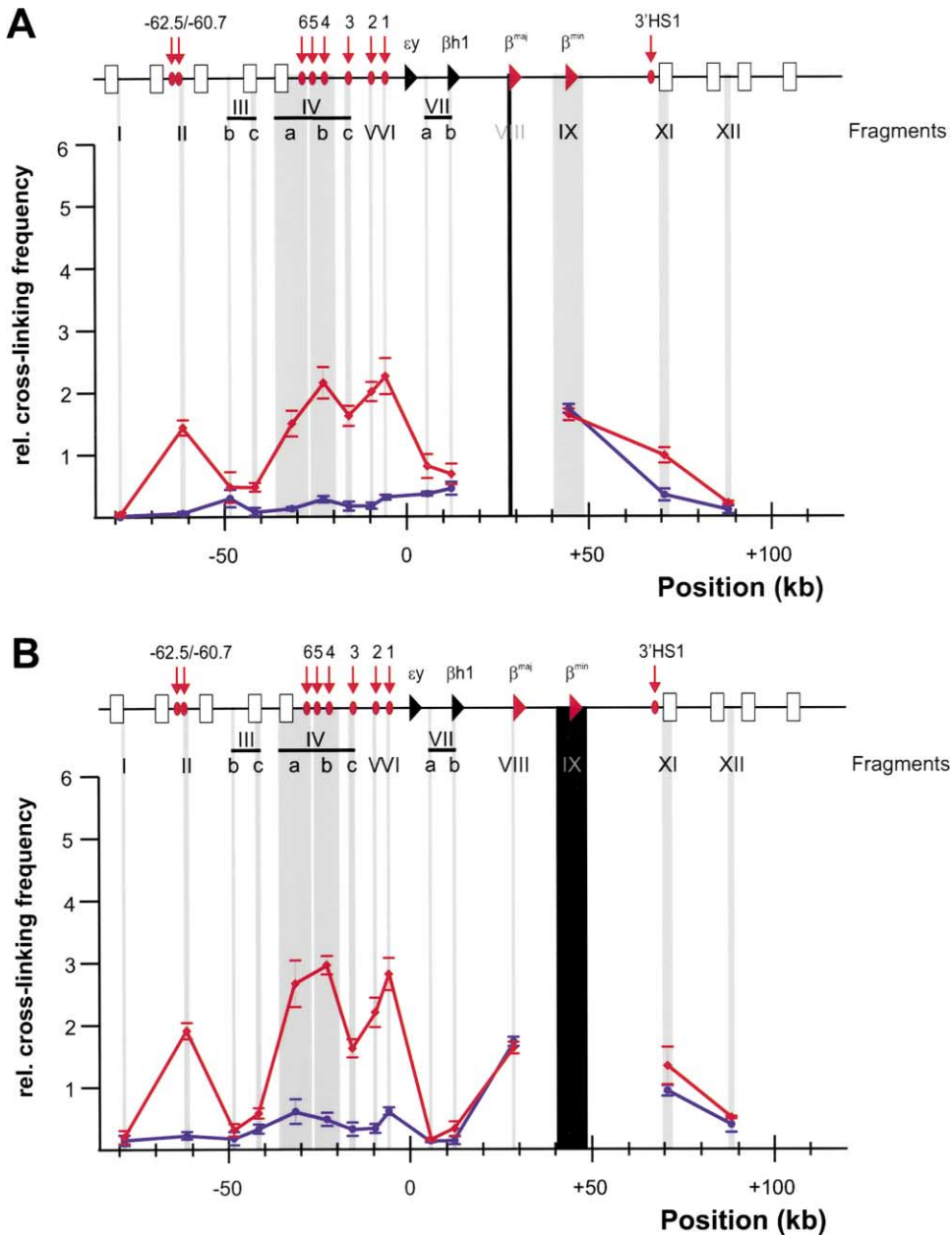


Figure 4. Erythroid-Specific Interactions between the Active β -globin Genes and Individual Hypersensitive Sites in the LCR

Relative crosslinking frequencies observed in fetal liver (red) and brain (blue) are shown. Standard error of the mean is indicated. Crosslinking frequency with a value of 1 arbitrarily corresponds to the crosslinking frequency between two neighboring *CaIR* control fragments (with restriction sites analyzed being 1.5 kb apart). Scaling on the y axis (from 0 to 6) allows direct comparison with other figures.

(A) Fixed HindIII fragment VIII (β^{maj}) versus the rest of the locus.

(B) Fixed HindIII fragment IX (β^{min}) versus the rest of the locus.

quencies with all other hypersensitive sites of the LCR (Figure 5A). Similar results were obtained with fixed fragment IV-b (HS4-5, Figure 5B), IV-a, IV-c, and VI (HS6, HS3, and HS1, respectively; data not shown). Together, these data provide strong support for the LCR acting as a holocomplex in erythroid cells to activate the globin genes.

HS at Both Ends of the Locus Participate in the Interactions between the LCR and the Active Genes
Two other erythroid-specific interactions stand out. In Figure 5B, for example, high crosslinking frequencies

were observed between HS4/5 and the fragments II and XI, at the far 5' and 3' end of the region, respectively. Interestingly, fragment II contains (part of) the recently identified hypersensitive sites -62.5 and -60.7 (Farrell et al., 2000), and fragment XI is located just 3' of another erythroid-specific hypersensitive site, 3'HS1 (Tuan et al., 1985; Grosveld et al., 1987). Interaction with both of the distal hypersensitive sites was seen with all other hypersensitive sites of the LCR, both in the HindIII experiments (see Figure 5 and data not shown) and in the BglIII experiments (see Figure 3B and data not shown). Moreover, the active β -major and β -minor genes also

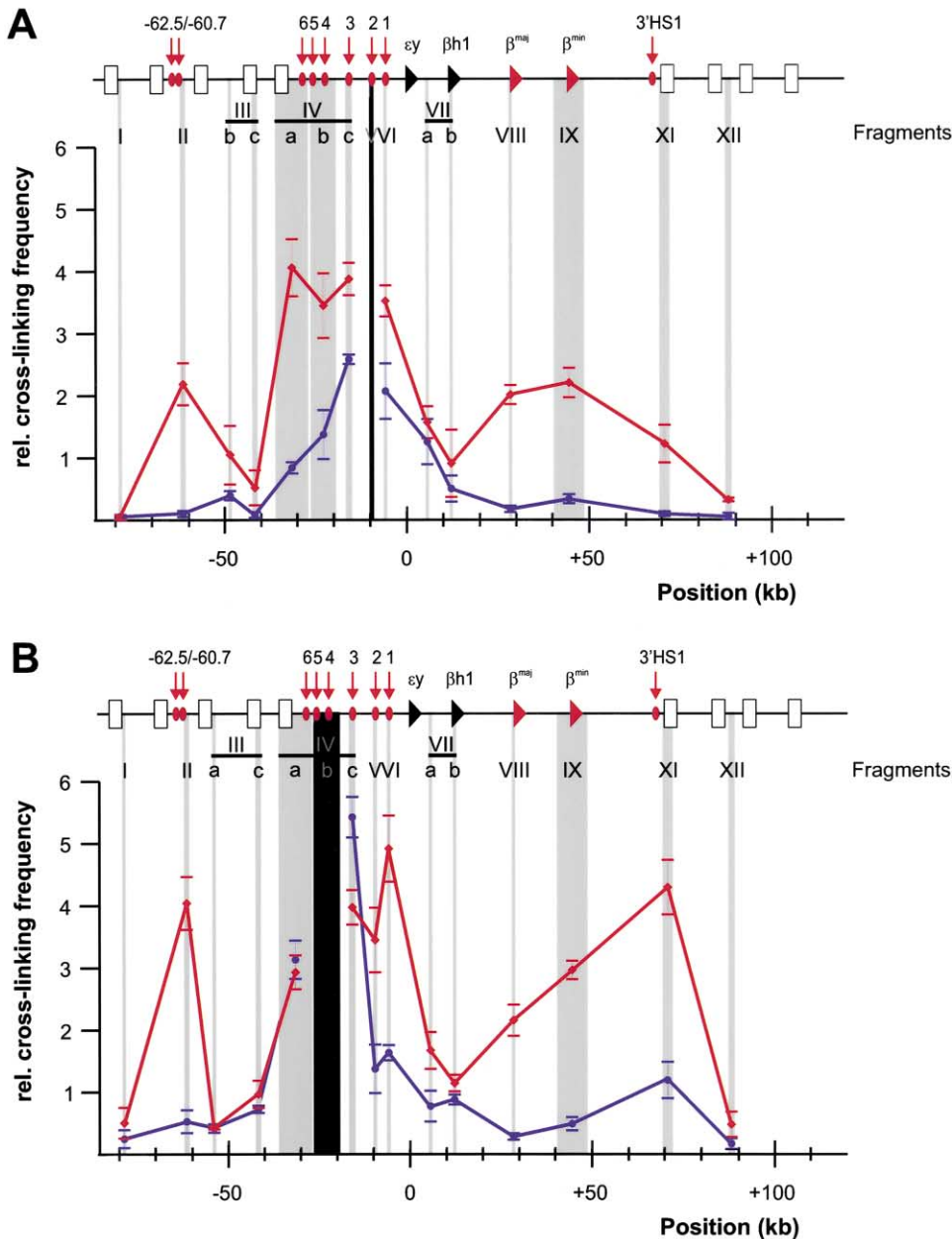


Figure 5. Erythroid-Specific High Crosslinking Frequencies among the Individual Hypersensitive Sites of the LCR and Two Distal Hypersensitive Sites

Relative crosslinking frequencies observed in fetal liver (red) and brain (blue) are shown. Standard error of the mean is indicated. Crosslinking frequency with a value of 1 arbitrarily corresponds to the crosslinking frequency between two neighboring CaIR control fragments (with restriction sites analyzed being 1.5 kb apart). Scaling on the y axis (from 0 to 6) allows direct comparison with other figures.

(A) Fixed HindIII fragment V (5'HS2 of the LCR) versus the rest of the locus.

(B) Fixed HindIII fragment IV-b (5'HS4-5 of the LCR) versus the rest of the locus.

showed erythroid-specific interactions with 5'HS62.5/–60.7 (Figures 3A, 4A, and 4B), despite being approximately 100 kb away. These data suggest a complex series of interactions between hypersensitive sites in the β -globin locus in expressing tissue.

To further investigate this, we analyzed locus-wide interactions with the distal hypersensitive sites. Figure 6A shows the results for fragment II, which confirm the interaction between 5'HS–62.5/–60.7 and LCR elements in the fetal liver. Fragments I and III, flanking these

5'HS, do not participate in this interaction (both in the BglIII and HindIII digestions), suggesting that the intervening DNA loops out. High crosslinking frequencies were also found between 5'HS–62.5/–60.7 and 3'HS1, which is remarkable considering the two sites are 130 kb apart on the linear chromatin template. Comparable interactions were observed using 3'HS1 as the fixed fragment (Figure 6B). However, it should be noted that the data for 3'HS1 are similar to those found for β -major and β -minor and that this region appears to act as one

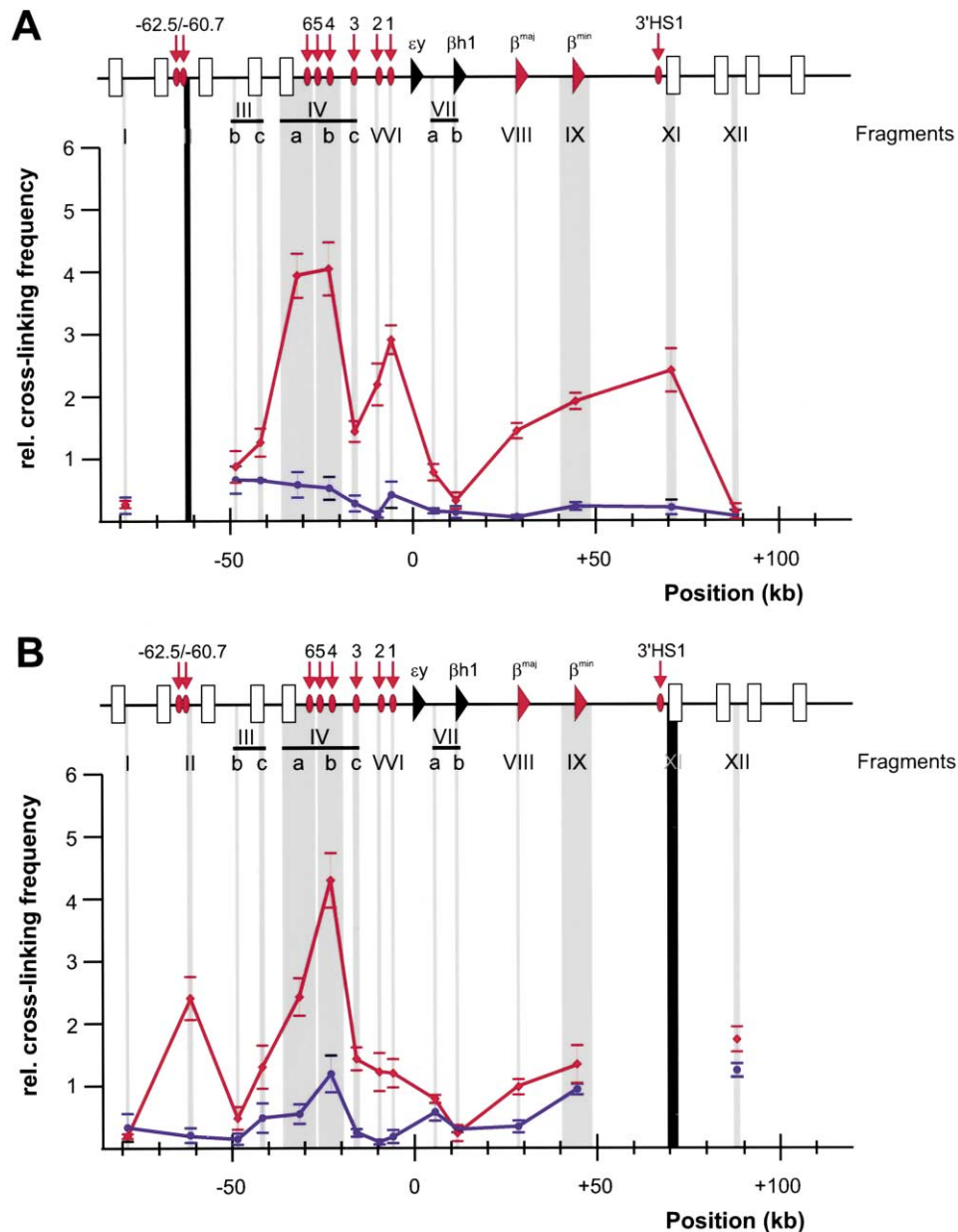


Figure 6. Two Distal Hypersensitive Sites at Each Side of the Locus Cluster with the LCR and the Genes

Relative crosslinking frequencies observed in fetal liver (red) and brain (blue) are shown. Standard error of the mean is indicated. Crosslinking frequency with a value of 1 arbitrarily corresponds to the crosslinking frequency between two neighboring CaIR control fragments (with restriction sites analyzed being 1.5 kb apart). Scaling on the y axis (from 0 to 6) allows direct comparison with other figures.

(A) Fixed HindIII fragment II (5'HS-62.5/-60.7) versus the rest of the locus.

(B) Fixed HindIII fragment XI (3'HS1) versus the rest of the locus.

block. The latter may point at some compaction, perhaps caused by the large amount of repetitive DNA present in this region (Bulger et al., 1999). Nevertheless, our data demonstrate that all the hypersensitive sites and the active genes of the β -globin locus cluster together in space in the erythroid nucleus.

Discussion

The looping model postulates that regulatory elements and genes/promoters communicate through direct in-

teractions between proteins bound to the DNA, with intervening chromatin looping out. In this paper we have demonstrated that the distal regulatory elements and the active genes, which are linked in *cis* in the murine β -globin locus, interact *in vivo* while the intervening DNA loops out. This looping is only seen in expressing cells and provides direct *in vivo* evidence for the looping model. Previous support for this model has come from several types of studies. *Trans*-activation, i.e., the ability of an enhancer to activate a promoter located on a physically separate DNA molecule, is most easily ex-

plained by direct contact between the enhancer and the gene. This has been observed in transvection in *Drosophila* (Bickel and Pirrotta, 1990) and in a number of in vitro experiments with artificial DNA constructs (Mahmoudi et al., 2002; Dunaway and Droge, 1989; Mueller et al., 1989). Competition between genes for a single regulator (Wasylyk et al., 1983; de Villiers et al., 1983; Hanscombe et al., 1991) leading to alternate transcription (Wijgerde et al., 1995) is also most easily explained by looping, particularly because the competitive advantage of the enhancer-proximal gene is lost when the genes are closely spaced at distances further from the enhancer (Heuchel et al., 1989; Dillon et al., 1997). However, all this evidence is indirect, and each can also be explained by other mechanisms. The findings presented here show direct evidence for looping in the active β -globin locus, whereas a linear type of structure is found for the nonexpressing locus. In particular, the observation that two hypersensitive sites at the far ends of the region cluster with the LCR and the active genes (i.e., all hypersensitive sites) provides new insights into long-range interactions (see below). However, the limitations of the 3C technique should also be noted in order to avoid overinterpretation of the results.

Interpreting 3C Analysis of the β -globin Locus

Some technical and biological aspects of the results by 3C analysis should be considered. As pointed out originally by Dekker et al. (2002), measuring crosslinking efficiency by the formation of ligation products largely depends on the frequency with which two genomic sites interact. They showed that contributions of other parameters, such as local protein concentrations or a favorable geometry of the crosslinked intermediate, are minor. Our results support this notion. However, we further believe that additional parameters, e.g., the fragment size, notably affect the crosslinking efficiency. Comparison of crosslinking frequencies observed with the large (26 kb) BglIII fragment IV (covering HS3-6 of the LCR and 12 kb upstream), to those observed with the much smaller HindIII fragments IV-a, -b, and -c (containing HS6, 4-5 and 3 as separate entities) reveals an increased background in brain for the large fragment. This can be explained by assuming that the chance of being crosslinked per se increases with fragment size. Also, an increase in ligation to irrelevant fragments will compete with ligation to specifically interacting fragments, causing underestimation of specific interactions in the fetal liver. Thus, to determine whether a specific interaction occurs between two given DNA sequences, it is best to study smaller fragments containing isolated entities.

The accuracy of signals obtained with the control template is crucial for our analysis. Since crosslinking values in brain and in liver are both normalized to the same control value, we were concerned about the fact that HindIII fragment IV-c showed a dip in relative crosslinking frequency with every fragment tested, both in brain and in liver. This result was due to high PCR signals in the control rather than low signals in the tissue samples (data not shown). Designing new primers did not solve this problem. Thus, although the observed crosslinking frequencies with HindIII fragment IV-c may be real, it is more likely that it reflects an as yet unresolved technical issue.

Purely biological parameters also play a role. For example, in 14.5 dpc fetal liver about 15%–20% of the cells are not expressing globin (judged by many RNA FISH experiments). These are likely to adopt a conformation similar to that observed in brain and to contribute to the total amount of substrate in the ligation reaction but not to the specific ligation frequency. Thus, the real value of erythroid-specific interactions will be underestimated, which increases the significance of finding these interactions, particularly the ones over large distances. Perhaps most importantly, interactions between distal DNA elements are thought to be dynamic (Wijgerde et al., 1995), while these measurements represent steady-state average levels. For example, a very important but short-lived interaction for transcription initiation (Wijgerde et al., 1995) may score much lower than a more long-lived interaction that would only stabilize the complex.

Given these limitations and the unknown dimensions of the chromatin fiber in the globin locus in vivo, the results presented here do not allow a strictly quantitative interpretation or conclusions as to what HS is responsible for in a given interaction and/or function. Predictions about the dynamics of the interactions or real nuclear distances are therefore not possible at this stage of development of the technique.

The Hypersensitive Sites, Looping, and an Open Chromatin Domain

Despite the limitations of the 3C technique, we can conclude that the six hypersensitive sites of the LCR, HS1-6, interact with the active genes, β -major and β -minor, in the 14.5 dpc fetal liver, with the inactive $\epsilon\gamma$ and βH1 genes on the intervening DNA fiber looping out. The upstream 5'HS–62.5/–60.7 participate in this interaction, again with the intervening DNA looping out. At the other end of the locus the 3'HS1 is also involved in the contacts, but we have no evidence for DNA looping out between the genes and 3'HS1. This region contains a large amount of repetitive DNA and may adopt a compacted structure as it appears to act in concert. The data also show a subdivision of the interactions, because we consistently observe the extreme 5' and 3' HS (5'HS–62.5/–60.7 and 3'HS1, respectively) to be closer to the 5' half of the LCR (HS4-6), which is not observed for the expressed genes.

The clustering of all hypersensitive sites in the β -globin locus is intriguing. Interactions are not confined to the outermost HS (we cannot exclude the presence of even more distal erythroid-specific hypersensitive sites), as proposed in boundary models (for review see Gerasimova and Corces, 2001), nor to sequences that have been proposed to act as insulators (Farrell et al., 2002) but include all HS and the promoters/enhancers of the genes. Thus, rather than being a particular type of transcription element, hypersensitivity appears to be the determining criterion for a DNA element to participate in clustering. We anticipate that this clustering is not confined to the β -globin locus only. We propose to name a 3D clustering of hypersensitive sites an active chromatin hub or ACH (Figure 7). Its formation is required to initiate transcription in repressive chromatin surroundings. The affinity between distal DNA hypersensitive

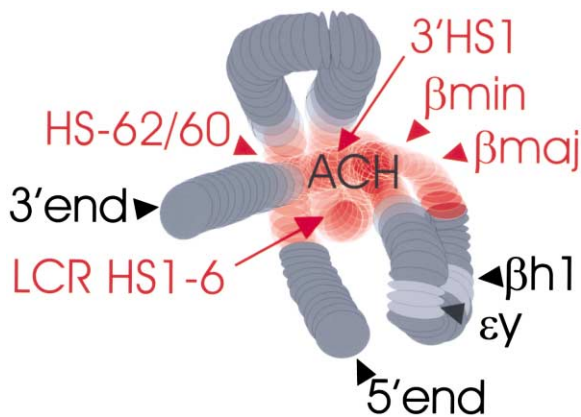


Figure 7. A 3D Model of the ACH

A hypothetical model of the active chromatin hub (ACH) is shown to illustrate the 3D nature of the ACH (not to scale), not the actual position of the elements relative to each other in vivo. Red indicates the active regions (hypersensitive sites and active genes) of the locus forming a hub of hyperaccessible chromatin (ACH). The inactive regions of the locus, having a more compact chromatin structure, are indicated in gray, with the inactive β h1 and $\epsilon\gamma$ genes in lighter gray. The olfactory genes are not shown. The interactions in the ACH would be dynamic in nature, in particular with the active genes (β maj and β min), which are alternately transcribed.

sites determines whether an ACH is productively formed or not. Affinity depends on the transcription factors bound to these DNA elements and can therefore be modulated (Wijgerde et al., 1996; Milot et al., 1996; Lundgren et al., 2000). Entry of new HS may stabilize or destabilize existing interactions, which in turn can alter expression levels of genes present in the ACH. The model does not predict how DNA sequences become hypersensitive in the first place (e.g., by mass action [Locke et al., 1988]), but stabilization/maintenance of hypersensitivity is proposed to depend on ACH formation. Surrounded by less active chromatin, the ACH would create a biphasic system, ensuring and stabilizing a high local concentration of transcription factors and associated chromatin modifying proteins to allow efficient transcription. The hypersensitive regions and promoters of the genes would have very high levels of, for example, histone acetylation (Burgess-Beusse et al., 2002; Bulger et al., 2002), whereas the chromatin outside the ACH would be less acetylated. An ACH need not occupy a fixed position in the nucleus but can be a dynamic fluid entity, possibly inside the interchromatin domain (ICD) compartment (Cremer et al., 2000). We propose that stable formation of an ACH underlies position-independent expression in transgenic experiments, which indeed can be accomplished by various combinations of HS. Such a scenario would explain why multicopy inserts may give position-independent expression (Ellis et al., 1993; Sabbattini et al., 1999).

Although formation of HS in the LCR precedes transcription (Blom van Assendelft et al., 1989; Groudine and Weintraub, 1982), we presently do not know whether the same holds for ACH formation. However, it is tempting to speculate that the ACH would take shape first, creating the appropriate environment, by modification of the locus, to recruit the actual transcription machinery. The observation that the globin genes are alternately

transcribed (Wijgerde et al., 1995; Trimborn et al., 1999) shows that only one of the genes is transcribed at any given moment. This implies that there is only one position of interaction within the ACH that allows initiation of globin gene transcription. In other gene clusters such a productive interaction may become stabilized and explain, for example, single gene expression (olfactory receptor genes).

We presently do not know how looping in the β -globin locus is accomplished. Although we like to think that initial contact occurs through random collision between distal elements, we cannot exclude other mechanisms from being involved in loop formation. Also, we do not know whether sequences other than HS (and cognate factors) participate directly in the ACH or perhaps stabilize its structure. Evidence from both *Drosophila* (Morcillo et al., 1997; Rollins et al., 1999; Zhou et al., 1999; Zhou and Levine, 1999; Sipos et al., 1998) and mammalian systems (Calzolari et al., 1999; Liu et al., 1997; Kmita et al., 2000) strongly suggests that there are elements and protein factors that stabilize long-range interactions. It will be interesting to determine whether such sequences are indeed part of the ACH.

Experimental Procedures

Chromosome Conformation Capture (3C)

We used the procedure recently developed by Dekker and coworkers (Dekker et al., 2002) with small adaptations to determine the spatial organization of the murine β -globin locus in 14.5 dpc embryos. Per experiment, 10–12 fetal livers or fetal brains were resuspended in DMEM supplemented with 10% FCS. The equivalent of two fetal livers or four fetal brains (approximately 4×10^7 cells) was diluted to 50 ml with DMEM (10% FCS). Formaldehyde was added to 2%, and the samples were crosslinked for 10 min at room temperature. The reaction was quenched by the addition of glycine to 0.125 M. Nuclei were harvested by lysis of the cells in ice-cold lysis buffer (10 mM Tris, 10 mM NaCl, 0.2% NP-40 [pH 8.0]) containing protease inhibitors. Nuclei were resuspended in the appropriate restriction buffer containing 0.3% SDS and incubated for 1 hr at 37°C while shaking. Triton X-100 was added to 1.8%, and the nuclei were further incubated for 1 hr at 37°C to sequester the SDS. The crosslinked DNA was digested overnight with the restriction enzyme (BglII or HindIII). Overnight incubation at 37°C did not result in any specific loss of hypersensitive sites due to the action of endogenous nuclease activity (data not shown). The restriction enzyme was inactivated by the addition of SDS to 1.6% and incubation at 65°C for 20 min. The reaction was diluted (to 2.5 ng/ μ l of genomic DNA) with ligation buffer (30 mM Tris-HCl, 10 mM MgCl₂, 1 mM DTT, 1 mM ATP [pH 7.8]), and Triton X-100 was added to 1% and incubated for 1 hr at 37°C. The DNA was ligated using T4 ligase for 4.5 hr at 16°C followed by 30 min at room temperature. Proteinase K was added, and samples were incubated overnight at 65°C to reverse the crosslinks. The following day, samples were incubated for 30 min at 37°C with RNase, and the DNA was purified by phenol extraction and ethanol precipitation.

To prepare a control template with detectable amounts of randomly ligated DNA fragments, we had to enrich for ligation products of interest (see also Results). PCR fragments spanning the restriction sites of interest were gel purified, and the DNA concentration was carefully determined using a Cary 100 Bio spectrophotometer (Varian). Equimolar amounts of the different PCR fragments were mixed and digested with the appropriate restriction enzyme followed by ligation. The mix was purified by phenol extraction and ethanol precipitation. The ligated fragments were diluted to the appropriate concentration (see below) and mixed with 300 ng digested and ligated genomic DNA.

PCR Analysis of the Ligation Products

The linear range of amplification was determined for the fetal liver samples and fetal brain samples by serial dilution. An appropriate

amount of DNA within the linear range (typically 300 ng of DNA for both liver and brain) was subsequently used for the experiments. The linear range of the control template was determined with a serial dilution of the random ligation mix made in the same amount (300 ng) of digested and ligated genomic DNA. Standardly, the 5' side of each restriction fragment was used to design primers unless this coincided with repetitive DNA sequences. Primer sequences are available on request. PCR products were run on 2% agarose gels and quantified on a Typhoon 9200 imager (Molecular Dynamics). All data points were generated from an average of five (with a minimum of three) different experiments performed in duplo. PCR products of the ligated fragments were run on agarose gels and quantitated. Crosslinking frequencies were calculated using the equation shown in Figure 1C. All probes (I–XIII) were tested against all other probes. A selection of the results is presented, and data not shown are in agreement.

As shown before (Dekker et al., 2002), formation of ligation products was strictly dependent on both ligation and crosslinking, i.e., lowering the amount of formaldehyde resulted in the loss of PCR product, as did the omission of T4 ligase (data not shown).

Acknowledgments

We would like to thank John Strouboulis, Niels Galjart, and Sjaak Philipson for comments on the manuscript. Also, we thank our colleagues in the lab for help with the logistics. This work is supported by NWO (The Netherlands Organization for Scientific Research) to W.d.L. as part of the Innovational Research Incentives Scheme and by NWO and EC grants to F.G.

Received: August 27, 2002

Revised: October 29, 2002

References

- Banerji, J., Rusconi, S., and Schaffner, W. (1981). Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299–308.
- Bickel, S., and Pirrotta, V. (1990). Self-association of the *Drosophila* zeste protein is responsible for transvection effects. *EMBO J.* 9, 2959–2967.
- Blom van Assendelft, G., Hanscombe, O., Grosveld, F., and Greaves, D.R. (1989). The β -globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 56, 969–977.
- Bulger, M., van Doorninck, J.H., Saitoh, N., Telling, A., Farrell, C., Bender, M.A., Felsenfeld, G., Axel, R., Groudine, M., and von Doorninck, J.H. (1999). Conservation of sequence and structure flanking the mouse and human beta-globin loci: the beta-globin genes are embedded within an array of odorant receptor genes. *Proc. Natl. Acad. Sci. USA* 96, 5129–5134.
- Bulger, M., Bender, M.A., van Doorninck, J.H., Wertman, B., Farrell, C.M., Felsenfeld, G., Groudine, M., and Hardison, R. (2000). Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. *Proc. Natl. Acad. Sci. USA* 97, 14560–14565.
- Bulger, M., Sawado, T., Schubeler, D., and Groudine, M. (2002). ChIPs of the beta-globin locus: unraveling gene regulation within an active domain. *Curr. Opin. Genet. Dev.* 12, 170–177.
- 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*, in press. Published online August 1, 2002. 10.1073/pnas.162342499.
- Calzolari, R., McMorrow, T., Yannoutsos, N., Langeveld, A., and Grosveld, F. (1999). Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and delta-beta-thalassemia affects beta- but not gamma-globin gene expression. *EMBO J.* 18, 949–958.
- Cremer, T., Kreth, G., Koester, H., Fink, R.H., Heintzmann, R., Cremer, M., Solovei, I., Zink, D., and Cremer, C. (2000). Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. *Crit. Rev. Eukaryot. Gene Expr.* 10, 179–212.
- de Bruin, D., Zaman, Z., Liberatore, R.A., and Ptashne, M. (2001). Telomere looping permits gene activation by a downstream UAS in yeast. *Nature* 409, 109–113.
- de Villiers, J., Olson, L., Banerji, J., and Schaffner, W. (1983). Analysis of the transcriptional enhancer effect. *Cold Spring Harb. Symp. Quant. Biol.* 47, 911–919.
- Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science* 295, 1306–1311.
- Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P., and Grosveld, F. (1997). The effect of distance on long-range chromatin interactions. *Mol. Cell* 1, 131–139.
- Dorsett, D. (1999). Distant liaisons: long-range enhancer-promoter interactions in *Drosophila*. *Curr. Opin. Genet. Dev.* 9, 505–514.
- Dunaway, M., and Droge, P. (1989). Transactivation of the *Xenopus* rRNA gene promoter by its enhancer. *Nature* 341, 657–659.
- Ellis, J., Talbot, D., Dillon, N., and Grosveld, F. (1993). Synthetic human beta-globin 5'HS2 constructs function as locus control regions only in multicopy transgene concatamers. *EMBO J.* 12, 127–134.
- Farrell, C.M., Grinberg, A., Huang, S.P., Chen, D., Pichel, J.G., Westphal, H., and Felsenfeld, G. (2000). A large upstream region is not necessary for gene expression or hypersensitive site formation at the mouse beta-globin locus. *Proc. Natl. Acad. Sci. USA* 97, 14554–14559.
- Farrell, C.M., West, A.G., and Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol.* 22, 3820–3831.
- Gerasimova, T.I., and Corces, V.G. (2001). Chromatin insulators and boundaries: effects on transcription and nuclear organization. *Annu. Rev. Genet.* 35, 193–208.
- Gribnau, J., de Boer, E., Trimborn, T., Wijgerde, M., Milot, E., Grosveld, F., and Fraser, P. (1998). Chromatin interaction mechanism of transcriptional control in vivo. *EMBO J.* 17, 6020–6027.
- Grosveld, F., van Assendelft, G.B., Greaves, D.R., and Kollias, G. (1987). Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell* 51, 975–985.
- Groudine, M., and Weintraub, H. (1982). Propagation of globin DNAase I-hypersensitive sites in absence of factors required for induction: a possible mechanism for determination. *Cell* 30, 131–139.
- Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and Grosveld, F. (1991). Importance of globin gene order for correct developmental expression. *Genes Dev.* 5, 1387–1394.
- Herendeen, D.R., Kassavetis, G.A., and Geiduschek, E.P. (1992). A transcriptional enhancer whose function imposes a requirement that proteins track along DNA. *Science* 256, 1298–1303.
- Heuchel, R., Matthias, P., and Schaffner, W. (1989). Two closely spaced promoters are equally activated by a remote enhancer: evidence against a scanning model for enhancer action. *Nucleic Acids Res.* 17, 8931–8947.
- Kellum, R., and Schedl, P. (1991). A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64, 941–950.
- Kmita, M., Kondo, T., and Duboule, D. (2000). Targeted inversion of a polar silencer within the HoxD complex re-allocates domains of enhancer sharing. *Nat. Genet.* 26, 451–454.
- Liu, Q., Bungert, J., and Engel, J.D. (1997). Mutation of gene-proximal regulatory elements disrupts human epsilon-, gamma-, and beta-globin expression in yeast artificial chromosome transgenic mice. *Proc. Natl. Acad. Sci. USA* 94, 169–174.
- Locke, J., Kotarski, M.A., and Tartof, K.D. (1988). Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* 120, 181–198.
- Lundgren, M., Chow, C.M., Sabbattini, P., Georgiou, A., Minaee, S., and Dillon, N. (2000). Transcription factor dosage affects changes

in higher order chromatin structure associated with activation of a heterochromatic gene. *Cell* 103, 733–743.

Mahmoudi, T., Katsani, K.R., and Verrijzer, C.P. (2002). GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *EMBO J.* 21, 1775–1781.

Milot, E., Strouboulis, J., Trimborn, T., Wijgerde, M., de Boer, E., Langeveld, A., Tan-Un, K., Vergeer, W., Yannoutsos, N., Grosveld, F., and Fraser, P. (1996). Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 87, 105–114.

Morcillo, P., Rosen, C., Baylies, M.K., and Dorsett, D. (1997). Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in *Drosophila*. *Genes Dev.* 11, 2729–2740.

Mueller, H.P., and Schaffner, W. (1990). Transcriptional enhancers can act in trans. *Trends Genet.* 6, 300–304.

Mueller, H.P., Sogo, J.M., and Schaffner, W. (1989). An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. *Cell* 58, 767–777.

Ptashne, M. (1986). Gene regulation by proteins acting nearby and at a distance. *Nature* 322, 697–701.

Rippe, K. (2001). Making contacts on a nucleic acid polymer. *Trends Biochem. Sci.* 26, 733–740.

Rollins, R.A., Morcillo, P., and Dorsett, D. (1999). Nipped-B, a *Drosophila* homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics* 152, 577–593.

Sabbattini, P., Georgiou, A., Sinclair, C., and Dillon, N. (1999). Analysis of mice with single and multiple copies of transgenes reveals a novel arrangement for the lambda5-VpreB1 locus control region. *Mol. Cell. Biol.* 19, 671–679.

Sipos, L., Mihaly, J., Karch, F., Schedl, P., Gausz, J., and Gyurkovics, H. (1998). Transvection in the *Drosophila* Abd-B domain: extensive upstream sequences are involved in anchoring distant cis-regulatory regions to the promoter. *Genetics* 149, 1031–1050.

Trimborn, T., Gribnau, J., Grosveld, F., and Fraser, P. (1999). Mechanisms of developmental control of transcription in the murine alpha and beta-globin loci. *Genes Dev.* 13, 112–124.

Tuan, D., Solomon, W., Li, Q., and London, I.M. (1985). The “beta-like-globin” gene domain in human erythroid cells. *Proc. Natl. Acad. Sci. USA* 82, 6384–6388.

Wasylyk, B., Wasylyk, C., Augereau, P., and Chambon, P. (1983). The SV40 72 bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. *Cell* 32, 503–514.

West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Insulators: many functions, many mechanisms. *Genes Dev.* 16, 271–288.

Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. *Nature* 377, 209–213.

Wijgerde, M., Gribnau, J., Trimborn, T., Nuez, B., Philipsen, S., Grosveld, F., and Fraser, P. (1996). The role of EKLF in human beta-globin gene competition. *Genes Dev.* 10, 2894–2902.

Zhou, J., and Levine, M. (1999). A novel cis-regulatory element, the PTS, mediates an anti-insulator activity in the *Drosophila* embryo. *Cell* 99, 567–575.

Zhou, J., Ashe, H., Burks, C., and Levine, M. (1999). Characterization of the transvection mediating region of the abdominal-B locus in *Drosophila*. *Development* 126, 3057–3065.