Molecular Cell, Vol. 1, 131–139, December, 1997, Copyright ©1997 by Cell Press

The Effect of Distance on Long-Range Chromatin Interactions

Niall Dillon,*‡ Tolleiv Trimborn,† John Strouboulis,‡§ Peter Fraser,† and Frank Grosveld†‡‖
*MRC Clinical Sciences Centre
Imperial College School of Medicine
Hammersmith Hospital
London W12 ONN
United Kingdom
†MGC Department of Cell Biology and Genetics
Erasmus University
3000 DR, Rotterdam
The Netherlands
‡Gene Structure and Expression
National Institute for Medical Research
London NW7 1AA
United Kingdom

Summary

We have used gene competition to distinguish between possible mechanisms of transcriptional activation of the genes of the human β-globin locus. The insertion of a second β-globin gene at different points in the locus shows that the more proximal β gene competes more effectively for activation by the locus control region (LCR). Reducing the relative distance between the genes and the LCR reduces the competitive advantage of the proximal gene, a result that supports activation by direct interaction between the LCR and the genes. Visualization of the primary transcripts shows that the level of transcription is proportional to the frequency of transcriptional periods and that such periods last approximately 8 min in vivo. We also find that the position of the β-globin gene in the locus is important for correct developmental regulation.

Introduction

A variety of phenotypic effects in eukaryotes are known to be the result of long-range action by regions of DNA on one or more target sequences. Such effects include transcriptional activation by distal control elements, heterochromatinization of translocated genes to give position-effect variegation, and phenomena such as silencing and insulation. The phenomenon of transcriptional regulation by sequences that are located at long distances from the promoter has been known for many years, but there is still intense debate about how such regulation takes place. Looping models that bring distally located sequences into direct contact with promoters have been widely favored (Ptashne, 1988; Muller et al., 1989; Bickel and Pirotta, 1990). However, arguments have also been put forward in favor of a binary model that proposes that the sole function of distal elements is to generate a favorable chromatin structure that allows the promoter to function effectively (Weintraub, 1988; Martin et al., 1996; Walters et al., 1996).

The study of long-range effects presents formidable technical difficulties. Although it has been possible to observe looping between protein binding sites in vitro (Li et al., 1991; Mastrangelo et al., 1991; Su et al., 1991), such systems are necessarily extremely simplified compared with the in vivo situation. Direct visualization of contacts between transcriptional control elements in vivo has not yet been achieved, and most of our information about such contacts comes from genetic studies in Drosophila (e.g., Bickel and Pirotta, 1990). Methods for probing active and inactive chromatin configurations are limited and provide little direct information about the role played in transcriptional activation by changes in chromatin structure. An alternative approach that we have pursued is to use detailed functional analysis of a multigene locus to make inferences about chromatin dynamics in vivo.

The system that we use, the human β-globin locus, has long been a prototypic system for the study of transcription in vertebrates (reviewed in Grosveld et al., 1993). The locus consists of five developmentally regulated genes that are activated at different stages of erythroid development (Figure 1). The genes are arranged in the order in which they are activated during development with ε expressed first in the embryonic yolk sac. Between 6 and 10 weeks gestation, there is a gradual switch to expression of the γ genes, which predominate during the fetal liver stage. In the later fetal liver and neonatal stages, there is a further transition to expression of the β gene, and the γ genes are almost completely silenced during adult life.

The locus has been characterized in detail over a number of years, and this has resulted in the identification of the principal functional elements involved in its regulation. The entire locus is activated by the locus control region (LCR) contained within a 15 kb region located 5′ of the ε gene (Grosveld et al., 1987). It has long been known from genetic studies that expression of one gene in the locus can reduce that of the others (Giglioni et al., 1984). This effect has been studied extensively in transgenic mice (Enver et al., 1990; Hanscombe et al., 1991; Peterson and Stamatoyannopoulos, 1993), and there is clear evidence that it operates in a polar manner with the genes located proximally to the LCR having a stronger suppressive effect on the more distally located genes (Hanscombe et al., 1991). The simplest explanation for the observed down-regulation would be that activation is achieved through direct interaction between the genes and the LCR and that only one gene can interact with the LCR at any one time. In contrast, the binary model (Walters et al., 1996) excludes direct interaction and proposes transcriptional interference to explain the down-regulation of distal genes (Martin et al., 1996).

§Present Address: National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892.
‖To whom correspondence should be addressed.

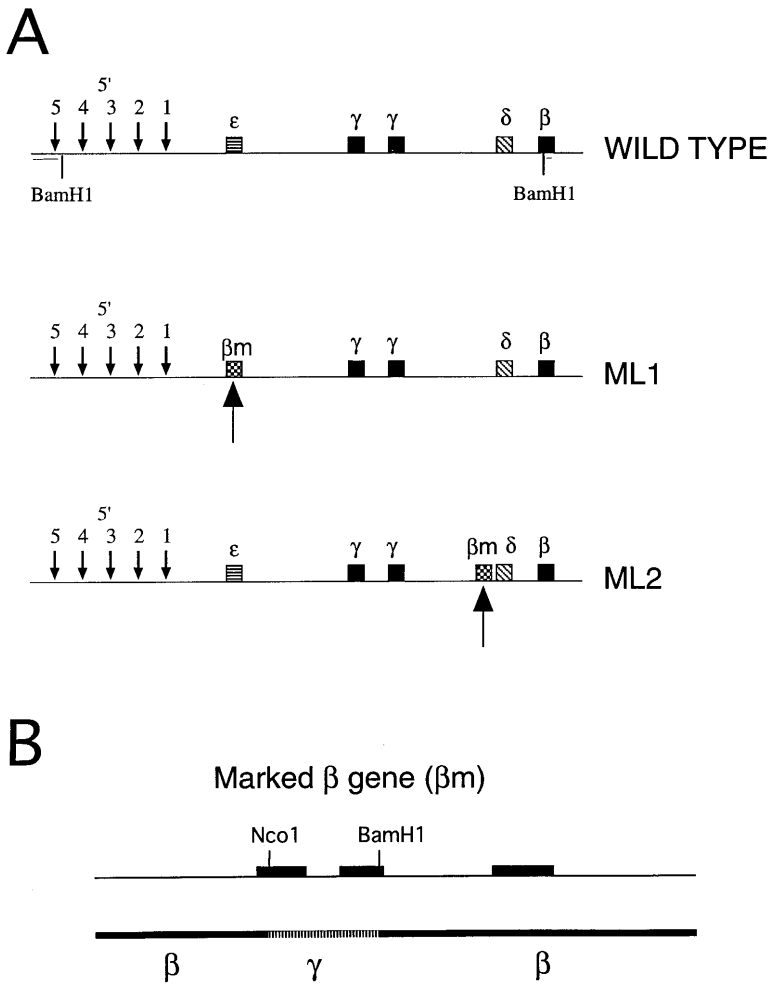

Figure 1. Structure of Normal and Modified β-Globin Loci

(A) Structure of the mutant β-globin loci used to generate transgenic mice. The LCR is indicated by the vertical arrows (1)–(5), and the genes as boxes. The thin lines and the BamHI sites indicate the positions of the probes that were used to analyze the ends of the transgenic loci (see Figure 2). Wild-type and mutant loci span a distance of 70 kb. In the ML1 locus, the β and βm genes are 50 kb and 6.8 kb, respectively, from the 3′ end (HS1) of the LCR (56.7 kb and 13.6 kb from the middle of the LCR), while the equivalent distance for the Gγ gene is 28 kb. In ML2, the distances of the β and βm genes from the 3′ end of the LCR are 53.7 kb and 42.8 kb, respectively (60.5 kb and 49.4 kb from the middle of the LCR).

(B) Structure of the marked β-globin gene (βm). A 410 bp fragment extending from the NcoI site at the ATG codon to a BamHI site located at the end of exon II was replaced with the equivalent fragment from the Aγ gene.

The study of interactions within the locus has recently been taken a step further by the use of in situ hybridization to directly visualize primary transcripts at individual alleles in single cells (Wijgerde et al., 1995). In transgenic mice homozygous for a human β-globin locus, cells from 12.5 day fetal liver were found to contain both γ and β primary transcripts, but almost all (87%) of the loci showed transcripts from only one gene. The result was interpreted as indicating that only one gene in the locus could be activated by the LCR at any one time. The presence of a minority of loci that gave double signals was interpreted as being due to the dynamics of the process. After a switch in gene activation, the newly synthesized RNA would colocalize with the decaying RNA of the previously transcribed gene. On the basis of these results and data obtained by manipulation of regulatory elements or transcription factors (Milot et al., 1996; Wijgerde et al., 1996), we postulated that transcriptional activation by distal elements is an all-or-nothing effect with transcription only taking place when the distal element is complexed with the gene. From this, it follows that the level of RNA produced would be determined by the frequency with which such transcriptional complexes are formed and the duration of the interactions.

However, it has never been possible to examine frequency and duration as independent variables because γ and β are different genes whose levels of expression change during the process of switching from γ to β. These factors introduce unknown variables into the analysis that limit the precision with which the kinetics of transcriptional competition can be measured (Wijgerde et al., 1995). In this study, we have eliminated these variables by generating mutant loci that contain a second functionally equivalent β gene at two different positions in the locus to allow the study of only the frequency of interaction.

We find that the steady-state RNA levels of the two β-globin genes are dependent on their relative distance from the LCR. The measurement of primary transcripts in individual cells shows that the number of transcriptional periods of each gene correlates with the levels of steady-state RNA. This provides independent evidence for direct interaction between the LCR and individual promoters and excludes transcriptional interference as a mechanism for the polarity of gene activation and silencing observed in the locus.

Results

Generation of Transgenic Mice Carrying Mutant β-Globin Loci

To test the effect of gene position on expression within the locus, we made use of a technique that involves

ligating two cosmid inserts to generate a 70 kb fragment containing the complete locus (Strouboulis et al., 1992b). Conventional cloning methodology carried out on the individual cosmids was used to generate two modified loci, each containing a second β-globin gene (βm) that was marked so that its transcript could be distinguished from that of the wild-type gene. In mutant locus 1 (ML1), the βm replaced the ε gene, while in mutant locus 2 (ML2), it was inserted close to the cap site of the δ gene (Figure 1A). Marking of the gene was achieved by replacing part of exon 1 and 2 and all of intron 1 with the equivalent sequences from the Aγ gene (Figure 1B). When marking the gene, it was important to reconstruct a fully functional globin gene to avoid generating a transcript that would be less stable in erythroid cells than the wild-type β-globin mRNA. Extensive functional testing has not detected any regulatory sequences within this region of the γ or β genes (Behringer et al., 1987; Bodine and Ley, 1987; Antoniou et al., 1988).

Transgenic mice were generated by microinjection of the mutant loci into oocytes. Since a multicopy tandem array would place an LCR close to the 3′ end of the wild-type β-globin gene, it was necessary to analyze animals that carried the modified loci at single copy. To do this, we generated a large number of transgenic founder animals and then mapped the transgenes extensively with particular emphasis on the ends of the injected fragment, to identify mice that carried a complete single copy of the locus.

Analysis of Transgene Structure

A total of 57 founder transgenics were generated (23 for ML1 and 34 for ML2). Tail DNA from these founders was digested with BamHI and probed with fragments from either end of the locus (Figure 2). Founders that showed a joining fragment indicative of a multicopy tandem repeat were discarded, while putative single-copy animals were bred to generate transgenic lines. This approach resulted in the generation of 2 single-copy lines for ML1 and 3 for ML2 (see below and Figure 2 with accompanying legend). The integrity of the locus in each of these lines was tested by probing blots of EcoRI-digested DNA with the complete cosmids used to generate the locus (Strouboulis et al., 1992a). Five of the lines had the locus fully intact, while a small rearrangement was detected in the middle of the locus in one of the ML1 lines (data not shown). This line was not analyzed further. Line 610 had 2 copies of the locus integrated as a tandem head-to-tail repeat, while line 217 had 2 copies of the locus linked together but separated by an unknown amount of mouse DNA (Figure 2). Such events appear to occur frequently during transgene integration (Singh et al., 1991; Strouboulis et al., 1992a). The structure of the locus in line 335 is not clear. The EcoRI blot failed to show a 3′-end fragment although end fragments were observed with other digests (data not shown). This indicates that there is at least one intact copy of the locus in this line. However, it is also clear that the intensity of the 5′-end fragment (Figure 2, left panel) is greater than that observed for the other lines when compared to the weakly hybridizing mouse band observed in all lanes. This suggests that line 335 contains a second truncated copy of the locus (with a segment of unknown size missing from the 3′ end of the second copy).

Effect of Position in the Locus on Transcriptional Function

Three different analytic approaches were used to obtain a comprehensive picture of the effect of position on the regulation of the two β genes in the mutant loci. Steady-state levels of transcripts from the two genes were measured by S1 analysis of adult blood RNA from each transgenic line. RNA FISH using intronic probes (Wijgerde et al., 1995) was used to detect the presence of primary transcripts at the site of transcription of the individual genes in 16.5 day fetal liver cells. This gives a measure of the relative frequency of transcription of the two β genes in erythroblasts. S1 analysis was also used to measure transcription of the genes at different stages of mouse development and to assess whether position in the locus affects developmental regulation. The results of these analyses are summarized in Figures 3 and 4 and described below.

Effect of Placing βm in a Distal Position in the Locus

The mutant locus construct ML2 places the βm gene in a position 10 kb upstream from the wild-type β gene. Placing the marked gene relatively close to the wild-type gene should show whether gene order affects transcription, either through competition for the LCR or transcriptional interference. Analysis of steady-state levels of mRNA from both genes in adult blood from four transgenic lines carrying the construct showed that the combined output of the two β genes is the same as that of endogenous mouse β (per copy of the mouse β locus). The data in Figure 3 show that expression of the two human β-globin genes is not equal. The presence of the βm gene at this position results in an expression level of approximately 75% (lines 301, 604, and 623) of the total and a reduction of expression of the wild-type human β gene to around 25% as measured by S1 protection analysis of steady-state RNA (Figure 3).

FISH analysis of primary transcripts was used to determine the number of actively transcribing genes in erythroblasts from 16.5 day fetal liver (when only the adult β genes are transcribed). βm and β primary transcription signals were detected by in situ hybridization with probes specific for the first intron of the γ genes (which detect primary transcripts from the βm gene) and the first intron of the β gene (which detects wild-type β transcripts). For each of the ML2 lines (301 and 623), approximately 500 expressing loci were counted for primary transcription signals. 50% (51% and 49% for lines 301 and 623, respectively) of the alleles show a single transcription signal for βm (Figure 3, green), whereas 13% (in 301 and 623) have a single β signal (Figure 3, red). A combined signal (red + green = yellow) is the result of simultaneous transcription of one gene and decay of primary transcripts of the other gene (Wijgerde et al., 1995; Gribnau and de Boer, personal communication). This is observed in 37% of the cells (36% and 38% in lines 301 and 623, respectively). Since the βm and β

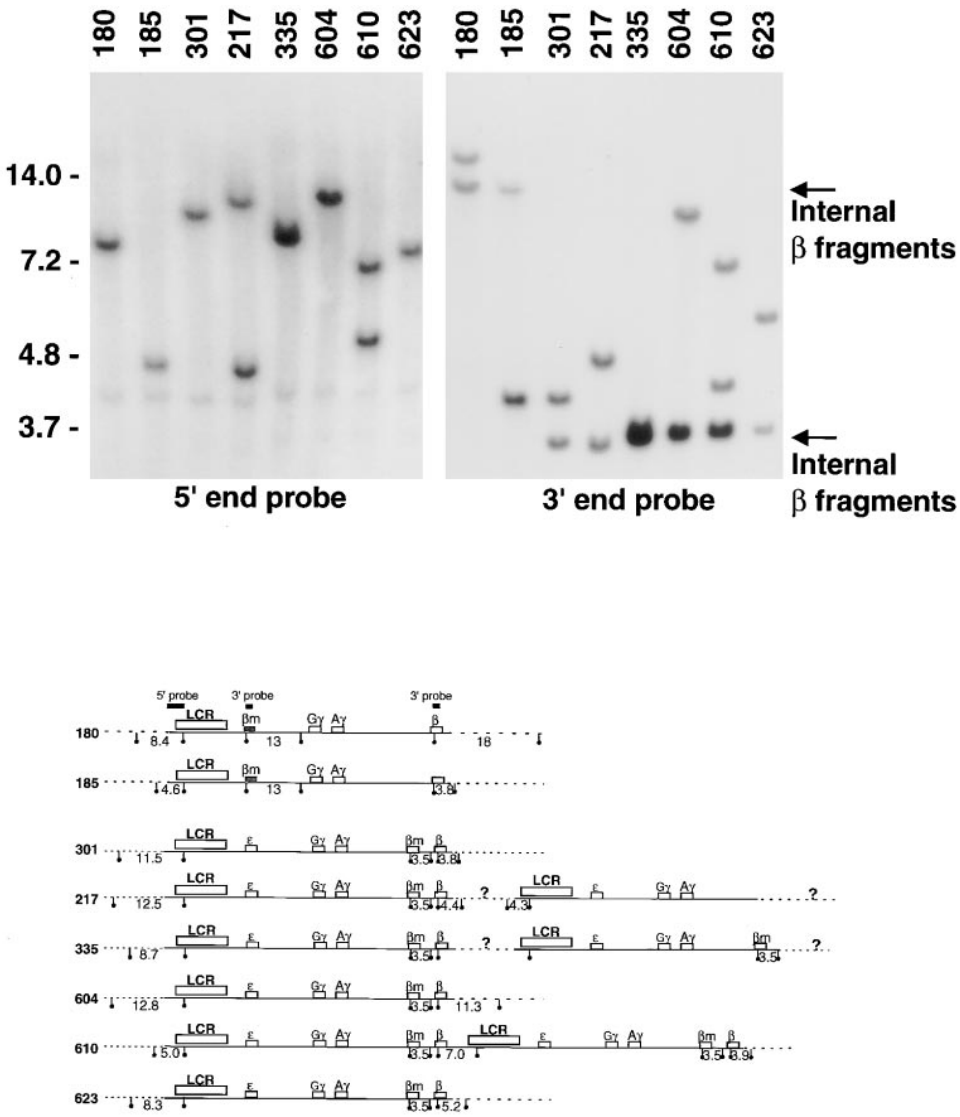

Figure 2. Analysis of Transgene Structure

End-blot of five single-copy lines together with two additional two-copy lines for ML2. The right-hand panel shows a BamHI digest probed with a 3.3 kb EcoRI fragment from the 5′ end of the locus (see Figure 1). The left-hand panel shows the same blot probed with a fragment from the second intron of the human β-globin gene. The size of the internal fragment differs in the ML1 and ML2 lines because of the different location of the gene in the locus. Each of the five single-copy lines gives a differently sized end fragment when hybridized with probes specific for the 5′ and 3′ ends of the locus (the additional internal fragment observed with the 3′ probe comes from the βm gene). Line 335 shows a higher signal than expected for a single-copy integration, but we have not detected a joining fragment as would be expected for a multicopy integration. It also fails to show a 3′-end fragment when digested with BamHI; possibly the BamHI end fragment is larger than the average size of the DNA in these preparations. Further blots did not clarify the situation, and hence, the exact structure of line 335 is not clear. Line 610 gave two end fragments and a fragment that corresponds to the size expected for a head-to-tail joining fragment, indicating that the line carries two copies of the transgene arranged as a head-to-tail repeat. Line 217 also gives two different end fragments when probed with the 5′-end probe. Although one end fragment is visible for the 3′-end probe, NcoI and HindIII digests each give two 3′-end fragments (not shown), indicating that this line contains two copies of the locus integrated in the same region but separated by an unknown amount of mouse DNA. The end fragments segregated together when the line was bred, indicating that two integration events have occurred close together.

genes are the same, the number of double spots (red + green = yellow) is only dependent on their frequency of transcription and signal decay. Using the average numbers (50% green, 13% red, and 37% yellow), it can be calculated (see Experimental Procedures and Discussion) that the frequency of β transcription is 31% and that of βm is 69%. This is in good agreement with the S1 protection analysis.

We conclude that the presence of the additional β gene down-regulates the distal β gene. Since the two β genes have identical promoters and flanking sequences, we also conclude that the down-regulation is caused

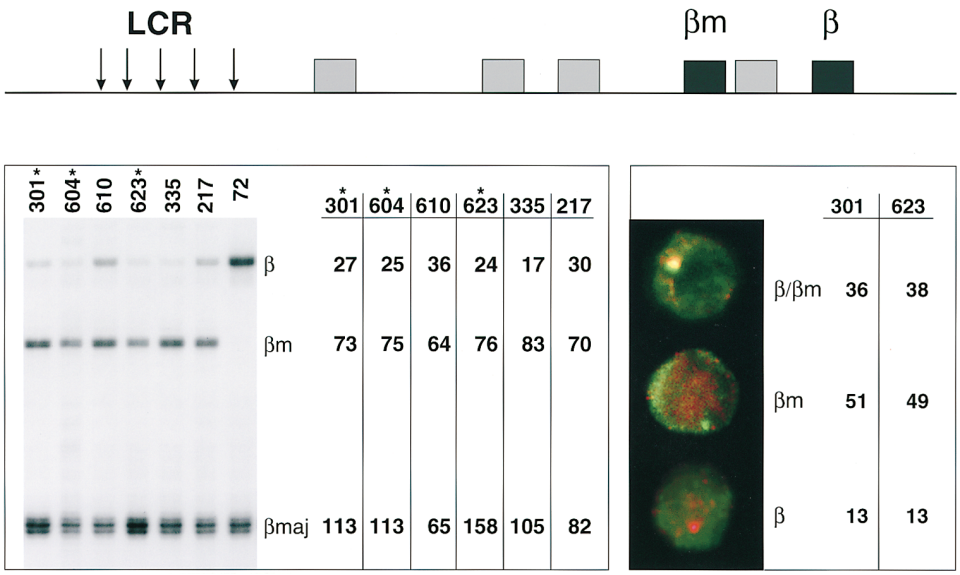

Figure 3. Transcriptional Analysis of the ML2 Locus

(Top Line) The ML2 construct with the βm and β genes indicated in black.

(Left Bottom) S1 analysis of adult blood RNA from lines transgenic for ML2 and the wild-type β-globin locus (Strouboulis et al., 1992a). The protected fragments are indicated on the right. The bands were quantitated by phosphorimage analysis and corrected for copy number. The numbers were normalized to 100% total human β and are shown on the right.

(Right Bottom) In situ hybridization of two of the ML2 lines with the percentages of cells showing βm/β double, βm single, and β single signals.

by a decrease in the frequency of transcription of the distal β gene. Possible explanations for this decrease would be competition for the LCR (Enver et al., 1990; Hanscombe et al., 1991) or transcriptional interference (Proudfoot, 1986; Martin et al., 1996).

Effect of Placing βm in a More Proximal Position

If transcriptional interference plays a role in down-regulating distal genes, then the effect would be expected to be most pronounced when the βm gene is closest to the β gene and to be reduced when the βm gene is placed in a more proximal position (i.e., closer to the LCR). If relative distance from the LCR is the main determinant, then the opposite effect would be expected. These predictions were directly tested using the construct ML1 in which βm is in the position normally occupied by the ε gene (Figure 1). The results of S1 analysis of adult blood from the two transgenic lines (180 and 185) obtained for ML1 are shown in Figure 4. Again, the total output of the locus is very similar to that of the endogenous mouse loci. In line 180, expression of the level of β RNA was 12% of the total human β RNA, while in line 185, expression was reduced to less than 1% (Figure 4). These results are confirmed by in situ analysis. Line 185 shows only 0.4% of the expressing loci with a single β (red) and 0.4% with a double βm/β (yellow) signal, while 99.2% express βm only (Figure 4). Line 180 shows 1.2% single β, 14.4% double βm/β, and 84.4% single βm signals. From these data, it can be calculated (see Discussion) that the frequency of transcription of βm in line 185 is 99% of the total, and in line 180, 91% of the total. As in the ML2 mice, these in situ data are in agreement with the S1 protection analysis. However, the variation between the ML1 mice indicates that one of these lines is subject to a position effect (see Discussion).

These data clearly show that moving the proximal gene closer to the LCR and much farther away from the distal gene results in a more effective suppression of the distal gene. In addition, since the proximal βm gene is now in a position where transcriptional interference on the distal β gene would be expected to be reduced (and would certainly not be increased), we conclude that down-regulation of the distal gene is the result of competition between the genes for direct interaction with the LCR.

Developmental Regulation

It has been suggested that the β-globin gene is normally at least in part suppressed in the early stages of embryonic development due to its position in the locus and, as a consequence, the competition from the ε and the γ genes (Enver et al ., 1990; Hanscombe et al., 1991; Dillon and Grosveld, 1993; Peterson and Stammatoyannopoulos, 1993). Since the ML1 mice carry a β-globin gene in the position of the ε gene, they provide an excellent opportunity to test the competitive silencing early in development. Timed matings were carried out for four of the transgenic lines (two for each construct), and the RNA was analyzed from 10.5 day yolk sac, 12.5 day fetal liver, and adult blood for the presence of transcripts from the human β, βm, and γ genes and the mouse embryonic βh1 and adult βmaj genes (Figure 5). Embryonic yolk sac from the transgenic line 72 carrying the wild-type locus shows the expected pattern of human

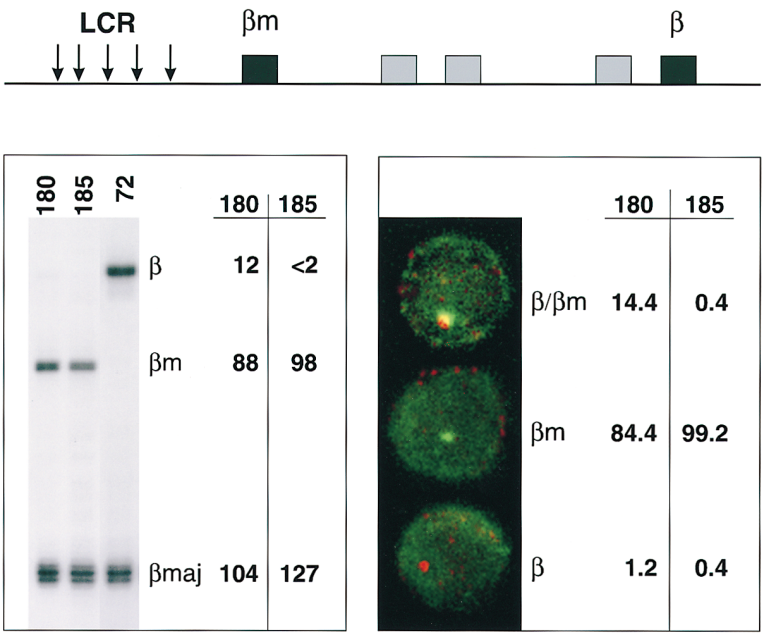

Figure 4. Transcriptional Analysis of the ML1 Locus

(Top Line) The ML1 construct with the βm and β genes indicated in black.

(Bottom Left) S1 analysis of adult blood RNA from lines transgenic for ML1 and the wild-type β-globin locus (Strouboulis et al., 1992a). The protected fragments are indicated on the right. The bands were quantitated by phosphorimage analysis, and the normalized numbers are shown on the right.

(Bottom Right) In situ hybridization on day 16.5 of development of two of the ML1 lines with the percentages of cells showing βm/β double, βm single, and β single signals.

globin expression with the γ genes expressed at high levels and the β gene completely suppressed. A similar pattern is observed for the ML2 lines (the βm gene in the distal position) with both βm and β completely suppressed. In contrast, in the two ML1 lines (the βm in the proximal position), βm is expressed at levels that are approximately equivalent to those of the γ genes. This result shows that the β-globin gene does not contain any flanking sequences (2 kb in either the 5′ or 3′ direction) that actively suppress its transcription in early development.

In fetal liver, the control line 72 shows the expected levels of human γ and β expression (Strouboulis et al., 1992a; Peterson et al., 1993, Figure 5). In contrast to both ML1 lines (βm located upstream from the γ genes), expression of γ in 12.5 day fetal liver is completely suppressed. In the ML2 lines (βm located downstream from γ), expression is observed in the early fetal liver. We therefore conclude that the position of the genes relative to the LCR profoundly affects their developmental regulation and that this is the effect of transcriptional competition. Our results indicate that the β-globin gene is normally supressed in embryonic and fetal tissues because its distal location results in a competitive disadvantage compared with the ε and γ genes.

Discussion

In this study, we have used a combination of technologies to examine chromatin interactions over large DNA regions. The LCR approach (i.e., using all of the elements required for physiological expression) allows quantitative conclusions to be drawn from measurements of steady-state levels of mRNA. The in situ analysis of primary transcripts provides qualitative information about transcription in each cell in a population. Techniques for manipulating large fragments and introducing them into mice permits the application of reverse genetics at the level of a complete locus. Using these approaches, we have been able to address a number of different questions.

Competition versus Transcriptional Interference

It has long been known that the transcription of one gene can affect the expression of other genes in the β-globin locus. Two different models have been suggested as explanations for this phenomenon. The competition model proposes that transcription depends on direct interactions between the genes and the LCR (Hanscombe et al., 1991; Wijgerde et al., 1995) and that such interactions are monogenic. The binary model proposes that the effect of the LCR on transcription is due to an alteration in chromatin structure and that the level of expression is determined entirely by the promoter without direct interaction with distal elements. As this model does not explain the fact that up-regulation of one gene results in a reduction in the expression of the others, it was necessary to invoke transcriptional interference as an additional mechanism to explain this effect (Martin et al., 1996). We distinguished between the competition and binary model by placing an extra β-globin gene (βm) at two different positions in the β locus. It would be expected that if down-regulation is caused by transcriptional interference, then the effect would decrease when the extra β gene is moved away from the gene with which it would interfere (ML1 versus ML2, Figure 1). In fact, suppression is substantially increased in ML1, and we conclude that interference cannot explain our result. By inference, transcriptional interference plays at most a minor role in the regulation of the β-globin locus. Of course, our results do not exclude the possibility that transcriptional interference plays a role in other situations.

The LCR contains five hypersensitive sites spread over a distance of 15 kb. Functional analysis in transgenic mice has shown that four of the sites are required

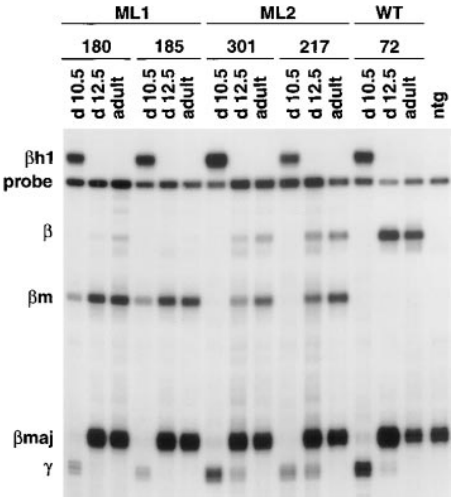

Figure 5. Developmental Regulation of the ML1 and ML2 Loci

S1 analysis of RNA from 10.5 day yolk sac, 12.5 day fetal liver, and adult blood from lines transgenic for ML1, ML2, and the wild-type β-globin locus. The protected fragments of the various probes are indicated on the left.

for full activation of a linked β-globin gene (Bungert et al., 1996; Milot et al., 1996; Peterson et al., 1996). The fact that the genes compete for the LCR and that a proximal gene can almost completely suppress transcription of a distal gene implies that the LCR functions as a single holocomplex as proposed by Ellis et al. (1996) rather than individual sites contacting the genes independently (Engel, 1993).

Kinetics of Complex Formation between the LCR and the Genes

The finding that there is competition between the genes in the locus and the observation of a predominance of single signals by in situ analysis demonstrates that the LCR can activate only one gene at a time (Wijgerde et al., 1995). The proportion of single and double signals depends on the frequency of switching, the duration of the interaction between the LCR and each of the genes, and the lifetime of the double signal after a switch. The two identical competing genes, βm and β, will have the same duration of interaction with the LCR. Since the lifetime of the signal (7 min, Gribnau and de Boer, personal communication) and the number of single and double signals are known, the frequency and duration of a transcriptional period can be estimated. The data obtained with the ML2 mice show that the number of single βm- and β-gene transcription signals for every 100 cells is 50 and 13, respectively, whereas 37 cells have a double signal (Figure 3). From these numbers, it is calculated (see Experimental Procedures) that there is a 69% probability of transcribing the βm gene versus a 31% probability of transcribing the β gene when both are in a distal position, as in the ML2 mice (301 and 623). Thus, the βm gene is transcribed 2.3 times more frequently than the β gene. It can also be calculated that the time of a transcriptional period lasts 8.2 min on average (see Experimental Procedures).

When the same calculation is used on the two ML1 lines (180 and 185) that were analyzed by in situ analysis, it results in two different values, namely a βm-to-β frequency of 160-fold and a transcriptional period of 20.5 min for line 185 (the very low β-globin expressor) and a βm-to-β frequency of 9-fold and a period of 7.5 min for line 180. The latter is very similar to the 8.2 min observed for the ML2 lines, and we conclude that the distal gene in line 185 is influenced by a negative position effect. In general, distal genes in single-copy integrations show some sensitivity to the position of integration in the mouse genome (Strouboulis et al., 1992a), in particular when the LCR interacts predominantly with a proximal gene. In that situation, the distal gene is not engaged by the LCR for most of the time and would therefore be available for an interaction with neighboring mouse sequences.

Effect of Position in the Locus on Competition

Our results show that placing a gene at different positions in the locus affects the frequency of its interaction with the LCR. A more distal gene is always at a disadvantage, but the effect is accentuated when the proximal gene is much closer to the LCR than the distal gene. What is the mechanism for this effect of relative distance from the LCR on competition? The more efficient competition by a proximal gene could be explained by a tracking model, where the LCR traverses along the DNA until it finds a gene and forms a complex with it. However, a simple tracking model predicts that the effect of the proximal gene would be the same whatever its position relative to the distal gene and does not explain the effect of relative distance. An alternative model (Dillon and Grosveld, 1993) proposed that LCR and genes can move freely in solution and that random contact between the genes and the LCR results in the formation of stable complexes. A gene that is closer to the LCR would come into contact with it more frequently and would therefore compete more effectively. As the proximal gene is moved away from the LCR and toward the distal gene, the difference in frequency of interaction would become less and the competitive advantage of the proximal gene would be reduced. The kinetics of free movement of two tethered elements predict that frequency of contact will be a function of relative distance to the power of 3/2 (Rippe et al., 1995). Calculation of the distance from the genes to the LCR is complicated by the fact that there are four hypersensitive sites spread out over a distance of 15 kb. Since the structure of the LCR in vivo is not known, only a range of 15 kb can be used by measuring the distance from the genes to either the 5′ end (HS4) or the 3′ end (HS1) of the LCR. This would result in a distance of β-LCR/βm-LCR to the power 3/2 of 6–20 for ML1 and 1.4–1.5 for ML2. Thus, if the genes were moving freely in solution, the predicted expression ratios of βm/β would be in the case of ML1 of 6- to 20-fold and ML2 of 1.4- to 1.5-fold.

However, the measured effect of competition by the proximal gene on the distal gene does not quite fit these predicted numbers (9-fold for ML1 line 180 [no position effect] and 2.3-fold for ML2). In particular, the deviation of the more reproducible numbers for ML2 implies that movement of two points on a fragment of DNA in the

nucleus is subject to additional constraints over and above those imposed by the fact that they are tethered together. The nature of these constraints is unclear, but the most likely constraint is probably the fact that the locus is part of a bigger loop.

Effect of Position on Developmental Regulation

Our results also demonstrate that the effect of position on competition between the genes of the β-globin locus for the activating function of the LCR has a substantial effect on the developmental regulation of the locus. Specifically, we have shown that embryonic silencing of the β-globin gene is dependent on its location in a distal position in the locus. We cannot say for certain which sequences in the proximal region of the locus are responsible for this embryonic activation, but the 2.7 kb fragment containing the ε gene that was removed and replaced with the β gene contains sufficient information to direct high-level expression when linked to the LCR (Raich et al., 1990; Shih et al., 1990), suggesting that any ε-specific enhancers have been removed from ML1. Embryonic expression of the γ genes in ML1 is significantly reduced compared with the wild-type locus, suggesting that the more proximal β gene actually competes more effectively with γ than with the ε gene that normally occupies this position. However, a similar reduction is observed when an extra β gene is located downstream from the γ genes, even though this gene is completely silent. Taken together with our other results, this suggests that the locus has evolved as a functional system in which the number of genes and the spacing between the genes are important parameters that act in concert with all of the regulatory elements to give the final level of expression. This has important implications for our understanding of the regulation of other complex multigene loci such as the highly conserved *Hox* loci and the imprinted murine *H19* region. In the latter case, an enhancer competition model has been proposed to explain the fact that deletion of the *H19* gene eliminates imprinting of the *Ins2* and *Igf2* genes located 95 and 80 kb away, respectively (Leighton et al., 1995). The study described here provides the first direct evidence that competition for shared elements does occur over distances of this magnitude in chromatin.

Polar competition clearly modulates switching between the γ and β genes in the early fetal liver and gives the γ genes a strong advantage by virtue of their more proximal position. However, the γ genes are expressed at similar levels to that of β in the early fetal liver, which means that other parameters must be acting to counteract this positional advantage. For example, destabilization of the interaction between the LCR and γ genes may lead to a reduced frequency of productive interaction following contact between the LCR and the γ genes.

Our results demonstrate the power of this type of functional analysis for making inferences about transcriptional behavior in chromatin. A particular advantage of this approach of functional probing is the fact that it uses the modulation of gene function in the native context to draw inferences about the relationship between structure and transcriptional behavior. The most interesting implication of this analysis is that the initiation of transcription of a gene can only take place while the distant regulatory sequence (LCR) is in direct contact with the gene and that the process of initiation stops as soon as the LCR/gene complex dissociates. Further use of functional probing should allow direct testing of a variety of different models for gene function in chromatin in vivo.

Experimental Procedures

Construction of Mutant Loci

Modifications were carried out on the cosmids Cos-LCRε and Cos γγδβ containing the 5′ and 3′ segments of the locus, respectively (Strouboulis et al., 1992a). A 2.7 kb ClaI/KpnI fragment containing the human ε gene was excised from cos-LCRε and replaced with a 4.9 kb BglII fragment containing the marked β-globin gene (Figure 1B). The same BglII fragment was inserted into the SalI site located at the cap site of the δ gene in Cos γγδβ. The two cosmids were joined together using the oligo-tailing procedure described by Strouboulis et al. (1992b).

Generation and Analysis of Transgenic Mice

The 70 kb fragments containing the complete mutant loci were purified by preparative agarose gel electrophoresis and injected into mouse oocytes. Southern blotting of tail DNA was carried out using probes specific for the ends of the locus. The intactness of the locus was confirmed by probing with the complete Cos-LCRε and Cos γγδβ. All methods used have been previously described by Strouboulis et al. (1992a).

S1 Analysis of RNA

RNA was isolated from 10.5 day yolk sac, 12.5 day fetal liver, and adult blood from transgenic lines and subjected to S1 analysis using 5′-end-labeled probes. The S1 procedure and the probes used for the mapping were identical to those used by Strouboulis et al. (1992a). The probe for the βm gene was a 190 bp fragment extending from −67 to +123 in the fusion gene. Quantitation was carried out by scanning.

In Situ Primary Transcription Analysis

The in situ analysis was carried out essentially as described by Wijgerde et al. (1995), using oligonucleotide probes that detect the first intron of the γ gene (visualizing βm as green signals) and the first intron of the β gene (visualizing the β gene as red signals). The equilibrium between βm and β transcription is then divided up in transcriptional time intervals T, while the lifetime of the double signal is 7 min (Gribnau and deBoer, personal communication). At the start, there is a possibility of each time interval to transcribe βm or β with a probability P or 1-P, respectively. There are four possible states per interval: (1) switching from β to βm (with probability P) results in 7 min double signal followed by T-7 min single βm signal; (2) switching from βm to β (with probability 1-P) results in 7 min double signal followed by T-7 min single β signal; (3) going from βm back to βm (with probability P) results in T min βm signal; and (4) going from β back to β (with probability P) results in T min β signal. This gives a Markov chain that can be solved in three equations per time interval T: (1) expected time that βm is observed, $P(1 - P)(T - 7) + P^2 = (\%\beta m \text{ signals})T$; (2) expected time that β is observed, $P(1 - P)(T - 7) + (1 - P)^2 = (\%\beta \text{ signals})T$; and (3) expected time that βm/β double is observed, $2P(1 - P)7 = (\%\beta m/\beta \text{ signals})T$. For example, these equations can be solved to result for line ML2 in $P = 69\%$ and $T = 8.2$ min.

Acknowledgments

We are grateful to Dr. T. Stijnen and J. Gribnau for their advice on the statistical evaluation of the in situ transcription data, to M. Kuit for photography, and to L. Braam for animal care. This work was supported by the Medical Research Council (United Kingdom), the Howard Hughes Medical Foundation, and the NWO (The Netherlands).

Received June 2, 1997; revised September 25, 1997.

References

Antoniou, M., de Boer, E., Habets, G., and Grosveld, F. (1988). The human β-globin gene contains multiple regulatory regions: identification of one promoter and two downstream enhancers. EMBO J. *7*, 377–384.

Behringer, R., Hammer, R., Brinster, R., and Palmiter, R.D. (1987). Two 3′ sequences direct adult erythroid specific expression of human β-globin genes in transgenic mice. Proc. Natl. Acad. Sci. USA *84*, 7056–7059.

Bickel, S., and Pirotta, V. (1991). Self association of the Drosophila zeste protein is responsible for transvection effects. EMBO J. *9*, 2959–2967.

Bodine, D., and Ley, T. (1987). An enhancer element lies 3′ to the human Aγ-globin gene. EMBO J. *6*, 2997–3004.

Bungert, J., Dave U., Lim, K., Lieuw, K., Shavit, J., Liu, Q., and Engel, J. (1996). Synergistic regulation of human β globin gene switching by LCR elements HS3 and HS4. Genes Dev. *9*, 3083–3096.

Dillon, N., and Grosveld F. (1993). Transcriptional regulation of multigene loci: multilevel control. Trends Genet. *9*, 134–137.

Ellis, J., Tan-Un, K.C., Harper, A., Michalovich, D., Yannoutsos, N., Philipsen, S., and Grosveld, F. (1996). A dominant chromatin-opening activity in 5′ hypersensitive site 3 of the human β-globin locus control region. EMBO J. *15*, 562–568.

Engel, J.D. (1993). Developmental regulation of human β-globin gene transcription: a switch of loyalties? Trends Genet. *9*, 304–309.

Enver, T., Raich, N., Ebens, A.J., Papayannopoulou, T., Costantini, F., and Stamatoyannopoulos G. (1990). Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. Nature *344*, 309–313.

Giglioni, B., Casini, C., Mantovani, R., Merli, S., Comi, P., Ottolenghi, S., Saglio, G., Camaschella, C., and Mazza, U. (1984). A molecular study of a family with Greek hereditary persistence of fetal hemoglobin and β-thalassemia. EMBO J. *3*, 2641–2645.

Grosveld, F., Blom Van Assendelft, G., Greaves, D., and Kollias, G. (1987). Position-independent, high-level expression of the human β-globin gene. Cell *51*, 975–985.

Grosveld, F., Dillon, N., and Higgs, D. (1993). The regulation of human globin gene expression. Baillieres Clin. Haematol. *6*, 31–55.

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.

Li, R., Knight, J.D., Jackson, S.P., Tjian, R., and Botchan, M.R. (1991). Direct interaction between Sp1 and the BPV enhancer E2 protein mediates synergistic activation of transcription. Cell *65*, 493–505.

Leighton, P., Ingram, R., Eggenschwiler, J., Efstratiadis, A., and Tilghman, S. (1995). Disruption of imprinting caused by deletion of the *H19* gene region in mice. Nature *375*, 34–39.

Martin, D.I., Fiering, S., and Groudine, M. (1996). Regulation of β-globin gene expression: straightening out the locus. Curr. Opin. Genet. Dev. *4*, 488–495.

Mastrangelo, I.A., Courey, A.J., Wall, J.S., Jackson, S.P., and Hough, P.V. (1991). DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. Proc. Natl. Acad. Sci. USA *88*, 5670–5674.

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.

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

Peterson, K., and Stamatoyannopoulos, G. (1993). Role of order in developmental control of human γ and β globin gene expression. Mol. Cell. Biol. *13*, 4836–4843.

Peterson, K.R., Clegg, C.H., Huxley, C., Josephson, B.M., Haugen, H.S., Furukawa, T., and Stamatoyannopoulos, G. (1993). Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human β-globin locus display proper developmental control of human globin genes. Proc. Natl. Acad. Sci. USA *90*, 7593–7597.

Peterson, K., Clegg, C., Navas, P., Norton, A., Kimbrough, T. and Stamatoyannopoulos, G. (1996). Effect of deletion of 5′HS3 or 5′HS2 of the human β globin LCR on the developmental regulation of globin gene expression in β globin locus YAC transgenic mice. Proc. Natl. Acad. Sci. USA, *93*, 6605–6609.

Proudfoot, N. (1986). Transcriptional interference and termination between duplicated α globin gene constructs suggests a novel mechanism for gene regulation. Nature *322*, 562–565.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature *335*, 683–689.

Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T., and Stamatoyannopoulos, G. (1990). Autonomous developmental control of human embryonic globin gene switching in transgenic. Science *250*, 1147–1149.

Rippe, K., von Hippel, P.H., and Langowski, J. (1995). Action at a distance: DNA-looping and initiation of transcription. Trends Biochem. Sci. *20*, 500–506.

Shih, D., Wall, R., and Shapiro, S. (1990). Developmentally regulated and erythroid-specific expression of the human embryonic β-globin gene in transgenic mice. Nucleic Acids Res. *18*, 5465–5472.

Singh, G., Supp, D., Schreiner, C., McNeish, J., Merker, H., Copeland, N., Jenkins, N., Potter, S., and Scott, W. (1991). Legless insertional mutation: morphological, molecular and genetic characterization. Genes Dev. *5*, 2245–2255.

Strouboulis, J., Dillon, N., and Grosveld, F. (1992a). Developmental regulation of a complete 70-kb human β-globin locus in transgenic mice. Genes Dev. *6*, 1857–1864.

Strouboulis, J., Dillon, N., and Grosveld, F. (1992b). Efficient joining of large fragments for transgenesis. Nucleic Acids Res. *20*, 6109–6110.

Su, W., Jackson, S., Tjian, R., and Echols, H. (1991). DNA looping between sites for transcriptional activation: self association of DNA-bound Sp1. Genes Dev. *5*, 820–826.

Walters, M.C., Magis, W., Fiering, S., Eidemiller, J., Scalzo, D., Groudine, M., and Martin, D.I. (1996). Transcriptional enhancers act in cis to suppress position-effect variegation. Genes Dev. *10*, 185–195.

Weintraub, H. (1988). Formation of stable transcription complexes as assayed by analysis of individual templates. Proc. Natl. Acad. Sci. USA *85*, 5819–5823.

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 β-globin gene competition. Genes Dev. *10*, 2894–2902.