Each hypersensitive site of the human β-globin locus control region confers a different developmental pattern of expression on the globin genes

Peter Fraser, Sara Pruzina, Michael Antoniou, and Frank Grosveld

Laboratory of Gene Structure and Expression, National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA, UK

We have tested the effect of the individual DNase I hypersensitive site (HS) regions of the globin locus control region (LCR) on the developmental expression pattern of the human γ and β genes in transgenic mice. The results show that HS3 is the most active site during the embryonic period. It is also the only site capable of high level expression of the γ genes during fetal hematopoiesis, in a population of cells that are capable of expressing both the γ and β genes. Region HS4 shows the highest activity during the adult stage and expresses the γ genes only at low levels during the embryonic period. HS2 drives equivalent levels of γ or β transgene expression throughout development. HS1 has a similar pattern to HS2, although the activity of HS1 is very low. From these results we conclude that the HS regions have distinct developmental specificities and suggest that in the complete LCR they interact with each other to form a larger complex which, in turn, interacts with the globin genes.

[Key Words: Human β-globin; locus control region; developmental regulation; transgenic mice]

Received August 27, 1992; revised version accepted November 2, 1992.

The human β-globin gene locus contains the embryonic-specific ε gene, the fetal-specific γ genes, and the adult-specific δ and β genes in the same order as they are expressed during development. The entire locus is controlled by the locus control region (LCR), which is located upstream of the ε gene [Fig. 1, below]. The LCR is characterized by a set of four developmentally stable DNase I hypersensitive regions [Tuan et al. 1985; Forrester et al. 1987; Grosveld et al. 1987], containing a number of sites for DNA-binding proteins [Philipsen et al. 1990; Talbot et al., 1990, Pruzina et al. 1991]. Transgene constructs containing the LCR show copy number-dependent expression of linked globin or heterologous genes in erythroid cells, independent of the site of integration of the transgene construct in the host genome [Grosveld et al. 1987; Blom van Asendel et al. 1989; Talbot et al. 1989]. When the developmental regulation of the individual globin genes in the absence of the LCR was studied in transgenic mice, it was found that the ε gene was inactive [Shih et al. 1990] but that the γ and β-globin genes were expressed in a developmental-specific manner, albeit at low levels and dependent on the position of integration in the host genome [Magram et al. 1985; Townes et al. 1985; Chada et al. 1986; Kollas et al. 1986]. When the genes were studied in the context of the complete LCR in transgenic mice, the ε gene was shown to be expressed at the embryonic stage only [P. Fraser, unpubl.] in agreement with the data published for the ε gene in combination with part of the LCR [Raich et al. 1990; Shih et al. 1990]. The γ-globin gene, like its murine structural homolog the βγ1 gene, is expressed in the embryonic yolk sac. However, in contrast to βγ1, the γ gene is expressed in the early fetal liver and is only silenced after day 16 of development. It is not expressed in the adult [Dillon and Grosveld 1991]. An individual β gene linked to the LCR is expressed prematurely at the embryonic stage [Behringer et al. 1990; Lindenbaum and Grosveld 1990], although not at maximal levels. Like the murine β gene, the human β gene is expressed at high levels in the fetal liver and adult [Grosveld et al. 1987; Behringer et al. 1990; Enver et al. 1990]. These data suggest that a large part of the developmental regulation of the globin genes is specified by the regions immediately flanking the genes. However, it also appears that the LCR is not developmentally neutral and that it influences the expression pattern of the genes.

When combinations of genes are used, the premature expression of the β gene is abrogated by competition for the LCR with another globin gene, for example, γ or α [Behringer et al. 1990; Enver et al. 1990; Hanscombe et al. 1991]. These results were confirmed by the expression of the complete human β-globin locus in transgenic mice [Strouboulis et al. 1992]. This competition appears to operate in a polar fashion, providing an advantage to a
gene that is proximal to the LCR and a disadvantage to a gene that is distal to the LCR (Hanscombe et al. 1991). The fact that competition between the genes is observed indicates that at least part of the developmental-specificity provided by the LCR is the result of differential interaction with the genes.

A key question, therefore, is whether the individual HS regions, which make up the complete LCR, interact differently with the genes or have distinct developmental-specificities. To answer this question, we introduced each HS region in combination with a set of human γ- and β-globin genes into transgenic mice and determined the expression pattern of the human globin genes during development. The results show that the individual HS regions confer distinct developmental patterns of expression on the γ or β-globin genes. These results have a number of implications for the mechanism of LCR function and show that developmental studies based on experiments with small parts of the LCR have to be treated with caution.

Results

The individual HS regions of the human globin LCR were tested by cloning each HS region in a unique SmaI site located upstream of the γy globin gene in the cosmid HG28 (Taramelli et al. 1986). This cosmid contains a genomic fragment encompassing the γy, δ-, and β-globin genes and is ~36 kb in length, with 3.4 kb of flanking DNA 5’ to the γy gene and 3.8 kb 3’ of the β-globin gene (Fig. 1). The restriction fragments containing HS regions 1–4 were the same (each ~1–2 kb) as used in previous studies (Talbot et al. 1987, 1990; Collis et al. 1990; Fraser et al. 1990; Philippsen et al. 1990; Pruzina et al. 1991) and cloned in the same orientation relative to the genes as they occur in the human genome. The vector sequences were removed from each of the four new cosmids [γγδβHS1–γγδβHS4] and the control cosmid HG28 (CosHG28) by restriction enzyme digestion and salt gradient centrifugation before injection into fertilized mouse eggs. Tail DNA was analyzed to identify transgenic animals, and different restriction digests were used to identify those animals that contained only intact copies of the construct (data not shown). Two transgenic founder animals for each construct were bred to establish transgenic lines to study the developmental pattern of expression of the human globin genes. RNA was prepared from 10.5-day embryos, 16.5-day fetal liver, and adult blood from animals >6 weeks of age. The level of human γ, β, and mouse βH1 and β major (Bmaj) gene expression was determined by S1 nuclease protection assays. Analysis of the two HG28 [i.e., no LCR sequences] lines, HG28A and HG28B, carrying 8 and 30 copies of the transgene, respectively, shows that the expression of the transgenes is detectable by in situ RNA. The highest level of transgene expression occurs in line HG28A, which has approximately fourfold fewer copies than line HG28B, clearly displaying a lack of copy number-dependent expression. In addition, the level of expression per

Figure 1. Schematic diagram of human β-globin locus on chromosome 11 and constructs used for microinjection. The top line represents the region of human chromosome 11, which contains the β-globin locus. The position of the ε, γ, and β-globin genes are indicated, as are the SmaI (S) and ClaI (C) restriction sites. The vertical arrows above the line indicate the positions of the HS sites, which make up the β-globin LCR (5’HS1–HS4 and 3’HS1). The lower lines (HG28, γγδβHS1–γγδβHS4) show the five constructs used for microinjection, which contain the 36-kb SmaI–ClaI fragment and the various LCR HSs.

Figure 2. S1 nuclease protection assay of HG28A and HG28B transgenic lines, which contain 8 and 30 copies, respectively, of HG28. [Lanes 1, 2] 12 µg of total 10-day embryonic RNA; [lanes 3, 4] 4 µg of total 16-day fetal liver RNA. [Lane 5] 2 µg of total 6-week adult blood RNA. Identification of the protected fragments for mouse βH1, βmaj, and human γ and β-globin is shown at left.
copy is, at best, 1/150th of an endogenous globin gene. These results not only confirm that high level expression of the transgenes is dependent on the presence of the LCR but also demonstrate convincingly that residual expression is extremely dependent on position effects.

Transgenic lines carrying HSI linked to HG28 [lines HS1A and HS1B; 34 and 8 copies of γδβH51, respectively] express low levels of γ and β [Fig. 3]. However, this level is 3- to 50-fold higher per copy than the HG28 lines [cf. Fig. 2], and the level of expression increases proportionally with copy number (see Fig. 7, below). Expression of γ-globin is observed only in the embryonic yolk sac [Fig. 3, lanes 1,2] and not in the fetal liver or adult blood [Fig. 3, lanes 3–5]. The apparent variation in the γ signal [relative to the mouse βH1 gene] within the same line is due to the fact that the human and mouse genes do not show an identical pattern of expression in the yolk sac [Strouboulis et al. 1992]. As a consequence, a small difference in gestational age of an individual embryo can alter the γ/βH1 ratio. The β-globin gene is expressed at low levels in the fetal liver and adult blood [Fig. 3, lanes 3–5] with a constant human β/mouse βmaj ratio. Such a constant ratio is also observed with the entire locus [Strouboulis et al. 1992]. Quantitation of the transgene expression relative to the endogenous murine βH1 (embryonic) or βmaj (fetal/adult) globin gene output at each stage shows that the γ gene is expressed at ~2% per copy, and the β-globin gene at ~4% per copy [Fig. 7, below]. These levels are quite low when compared to the complete locus in transgenic mice [Strouboulis et al. 1992]. We therefore conclude that HSI has a very weak activity and is at its lowest in the embryonic yolk sac when the γ gene is expressed.

The presence of HS2 [transgenic lines HS2A and HS2B, 12 and 8 copies of γδβH52, Fig. 4] results in a quantitatively different, but qualitatively similar, result as that observed with γδβH51. The γ- and β-globin genes are expressed much more efficiently at ~15% per copy [Fig. 7] and, as observed with HSI, the γ genes are silenced at the end of the embryonic period [Fig. 4, lanes 1,2] and are not expressed in the fetal liver [Fig. 4, lanes 3,4]. In contrast to HSI, the relative increase in fetal β expression versus embryonic γ expression (4% vs. 2%) is not observed with HS2 (15% vs. 15%).

Addition of HS3 [transgenic lines HS3A and HS3B, two and six copies of γδβH53] leads to an interesting change in the expression pattern of the genes (Fig. 5). This is the most active site in the yolk sac and fetal liver. The highest level of expression occurs in the embryo, where the γ genes are expressed at ~40% of the level of mouse βH1 genes per copy [Fig. 5, lanes 1,2; Fig. 7, below]. In the fetal liver, the β gene is expressed at ~15%; however, γ expression persists [Fig. 5, lanes 3,4] at a level similar to that observed for the full locus [Strouboulis et al. 1992]. The β gene is expressed in adult blood at ~20% per copy [Fig. 5, lane 5; Fig. 7, below], compared with 90–100% for the full locus [Strouboulis et al. 1992]. This suggests that HS3 is the major component driving γ gene expression in the embryo and early fetal liver and is the principal determinant involved in the competitive interaction between the γ and β genes in the fetal liver of transgenic mice [Berry et al. 1992].

Transgenic lines HS4A and HS4B, which contain 14 and 7 copies of γδβH54, respectively, show yet another pattern of expression. The γ genes are poorly expressed in the embryonic yolk sac [Fig. 6, lanes 1,2] at ~2% per copy, and their expression is suppressed in the fetal liver. The β gene is expressed at 8% per copy in the fetal liver.
but is increased to 25% in adult blood [Fig. 6, lanes 3, 4, and 5, respectively; Fig. 7], providing the most efficient expression of any of the sites in the adult stage. This suggests that HS4 plays a minor role in embryonic expression in the yolk sac but has a major role in the adult expression of the β gene.

The γβββHS3-bearing mice provide a good opportunity to determine whether the fetal coexpression of γ and β (Fig. 4, lanes 3, 4) occurs in the same cell or in separate cells. That is, the coexpression of γ and β during the fetal period the result of the existence of two populations of cells, one only capable of expressing γ and one capable of expressing β, or are mouse fetal liver cells capable of expressing both γ and β as found in the case of human fetal cells and adult F cells [Stamatoyannopoulos and Nienhuis 1987]. Fetal liver cells (16.5 day) from transgenic line HS3B were therefore stained with a monoclonal antibody specific for human β chain and a rabbit polyclonal antibody specific for human γ [Fig. 8, D–I]. The antibodies were visualized by indirect immunofluorescent confocal microscopy. Figure 8A shows cytoplasmic staining with Texas red of γ-expressing cells. Figure 8B shows the same field of fetal liver cells viewed under the FITC channel [green], which identifies β-expressing cells. Computer-generated superimposition of the images clearly shows that the majority of cells express both the γ and β genes [Fig. 8C; note orange/yellow cells]. We therefore conclude that murine fetal liver cells resemble human fetal liver and adult F cells in that they are capable of expressing both γ and β genes.

Discussion

The results described in this paper have a number of implications for our understanding of the developmental regulation of the globin genes and the role of the LCR. The results show that the transgenic expression of cosHG28, without LCR sequences, is up to several hundred-fold below the level of an endogenous murine globin gene or the entire human β-globin locus in transgenic mice. In addition, the level of expression was not dependent on transgene copy number, as our 8-copy line (HG28A, Fig. 2) expressed higher levels of the transgenes than our 30-copy line (HG28B). This clearly demonstrates the position-dependent nature of transgenes that lack an LCR. Addition of LCR HS1, HS2, HS3, or HS4 to cosHG28 not only provides large increases in the level of expression of the transgenes but also results in expression levels directly proportional to gene copy number. This result is in agreement with and confirms our previous data using individual HSs driving the expression of a β-globin gene in 14-day fetal liver of transgenic mice [Fraser et al. 1990; Philipsen et al. 1990; Talbot et al. 1990; Prauzina et al. 1991; Ellis et al. 1992; Philipsen et al. 1993].

The most interesting aspect of this work is the fact that the precise developmental expression pattern of the genes is dependent on the HS used in the construct. Clearly, the HSs are functionally distinct in a developmental context. This has important implications for results that have been obtained by using only part of the LCR, particularly, those where HS3 and HS4 have been left out of the constructs. Conclusions about the developmental expression pattern of genes obtained with such constructs must be treated with caution when they are related back to the normal in vivo situation.

Although HS1 expression levels are low, both HS1 and
HS2 express either the γ or the β gene throughout development. Importantly, neither site is capable of driving γ-globin gene expression in the fetal liver. Although we do not as yet know the effect of each of these sites on the expression of the e-globin gene, it would appear from these data that HS1 plays a minor role throughout development. HS2 contributes substantially to the overall expression of the locus at all developmental stages, albeit at lower levels than HS4 in the adult and HS3 at all stages.

The results obtained with HS3 are the most interesting in that it is the only site capable of driving γ-globin expression in the fetal liver. We have not as yet found any stage-specific factors that bind to HS3, which consists of a core of repeated GATA1-binding sites, a G-rich motif that preferentially binds Sp1, and an NF-E2-binding site outside the core [Philipsen et al. 1990]. All of these cis-acting elements are bound by factors in vivo [Strauss and Orkin 1992], but none is stage specific for the fetal liver [S. Philipsen, unpubl.]. We therefore tentatively conclude that this particular arrangement of binding sites as compared to HS1, HS2 and HS4 leads to a preferential interaction with factors bound at the γ-globin gene promoter in the fetal liver. HS3 is also the most active site in the embryo and is still active in the adult. HS4 has low transcriptional activity during the embryonic and fetal stages but is the most active site in the adult stage in conjunction with β gene expression.

When the results described in this paper are compared with those obtained for the expression pattern of different combinations of genes or of the entire locus in transgenic mice, an interesting possibility emerges as to how the complete LCR may function in vivo. The expression levels achieved by individual HSs are lower than those observed with the complete LCR and locus [Strouboulis et al. 1992], suggesting that HS elements must interact with each other to realize full expression. This interaction does not result in a synergistic amplification of the expression level but appears to be more of an additive effect. A possible mechanism for such an effect is that the HS regions interact to form a larger complex that interacts with any gene or set of genes at any particular stage. This would agree well with the data obtained with a core HS2 fragment [Ellis et al. 1993, and in prep.], which is only active when more than one copy of the construct is present in transgenic mice, again suggesting that HS elements interact with each other to achieve activation. The formation of a large LCR complex by interaction between the HS elements would also explain why the γ-globin gene can effectively compete the β-globin gene [Behringer et al. 1990; Enver et al. 1990; Hanscombe et al. 1991; Berry et al. 1992]. Competition between the genes implies that the interaction with the LCR is the limiting factor. This is difficult to explain if the HSs are operating independently, as each HS can direct the transcription of the β-globin gene [Fraser et al. 1990, and this paper] and fetal liver cells are capable of expressing both γ and β genes [this paper]. However, if the individual elements interact to form a single large LCR complex, this would result in a competitive arrangement where only one regulator, the large LCR complex, is available for multiple genes. Experiments to test such a model in vivo have been made possible by our recent demonstration that the entire locus can be introduced as a single copy into transgenic mice [Strouboulis et al. 1992] and are now in progress.

Materials and methods

DNA constructs

Cloning of the individual DNase I HS fragments of the human β-globin LCR into synthetic polylinker plasmids has been de-
scribed previously in detail by Collis et al. (1990). The fragments were ligated into a unique Smal site located ~3.5 kb upstream of the Gγ gene of cosH28, which contains 36 kb of contiguous DNA encompassing the Gγ, Aγ, δ and β genes (Taramelli et al. 1986). The HS fragments were H54, Scal–PvuII; HS3, Scal–BstEII; HS2, Nael–Nael, HS1, XmnI–PvuII.

Purification of cosmids inserts for microinjection

The cosmid/LCR constructs were digested with SacII/Clai for the γγδβHS4 and γγδβHS3 cosmids; Asp718/Clai for γγδβHS2; NotI–HaeII for γγδβHS1, and Smal–Clai for cosH28, which contains no LCR HSs. Cosmid inserts were then purified by centrifugation on salt gradients. Briefly, 50–100 μg of digested cosmid DNA in 200 μl was layered on top of a continuous 5–25% NaCl gradient in 3 mM EDTA (pH 8.0) and centrifuged at 37 K rpm for 5.5 hr in a Beckman SW40 rotor at room temperature. Fractions of 500 μl were collected, and aliquots were run off a 0.6% agarose gel. Fractions containing only cosm id insert fragments were pooled, diluted with two volumes of H2O2, and precipitated with 2 volumes of ethanol. DNA fragments were redissolved in 10 mM Tris (pH 7.5) and 0.1 mM EDTA at 1–2 ng/μl for microinjection.

Transgenic mice

Purified fragments were microinjected into the pronuclei of fertilized mouse eggs and transferred into the oviducts of pseudopregnant [CBA × C57Bl] F1 female mice. Tail DNA was prepared from 10-day-old offspring and screened for the presence of the β-globin transgene by slot blotting. On the basis of Southern blots, two transgenics that contained only intact copies of a particular construct were bred to produce transgenic lines. Second- or third-generation transgenics were bred to produce transgenic embryos and fetuses for developmental analysis. Placental DNA was prepared from 10-day embryos and 16-day fetuses and screened for the β-globin transgene by slot blots. Copy numbers were determined by quantitation of Southern blots via Phosphorimage analysis.

Preparation of RNA and S1 nuclease protection assays

RNA was prepared from frozen transgenic 10-day embryos, 16-day fetal livers, and adult blood (>6 weeks) and analyzed by S1 nuclease protection assays as described previously (Fraser et al. 1990). The βmaj S1 probe is a 700-bp NcoI–HindIII fragment with a protected size of 100 bp; the human β probe is a 525-bp Accl fragment, which protects 155 bp; the mouse βH1 probe is a 255-bp HindII fragment, which protects 180 bp; and the human γ probe is a BstNI fragment, which protects 136 bp. Equimolar amounts of each end-labeled probe (βmaj, 15 ng; human β, 11 ng; βH1, 5.5 ng; human γ, 4.3 ng per sample) were used with the amounts of RNA indicated in the figure legends.
Immunofluorescent staining

Indirect immunofluorescence was used to detect human β- and γ-globin in the cytoplasm of hematopoietic fetal liver cells from transgenic mice. The antibodies employed were a mouse monoclonal antibody against β-globin [Immuno-rx, Augusta, GA], detected via a FITC-conjugated, goat anti-mouse IgG [Sigma Chemical Co., St. Louis, MO], a polyclonal rabbit anti-γ-globin antibody detected by a Texas red conjugated goat anti-rabbit IgG [both from Calbiochem, La Jolla, CA]. Fetal livers were disrupted by repeated pipetting with a Gilson p-1000 pipetman in PBS, washed three times in PBS, and spread on 10-mm polylysine-coated cover slips. The cells were allowed to attach (1-2 min) and were then fixed with 3.7% paraformaldehyde in CSK [100 mM NaCl, 300 mM sucrose, 10 mM Pipes at pH 6.8, 3 mM MgCl₂, 1 mM EGTA] for 10 min with periodic swirling, followed by three washes with 5 min with PBS. The fixed cells were permeabilized with 0.5% Triton X-100 in CSK for 15 min followed by three washes with 5 min with PBS. Tissue culture cell lines were immobilized and fixed by the same procedure. Cover slips were then overlaid with 5 μl of a block solution consisting of 0.8% bovine serum albumin and 0.1% gelatin in PBS for 15 min. The antibodies diluted in the same block solution were then applied (10 μl) in the following sequence: [1] anti-β-globin (1:200); [2] FITC-anti-mouse IgG (1:100); [3] anti-γ-rabbit (1:1000); and [4] Texas red anti-rabbit IgG (1:200). The binding time for each antibody was 45 min at room temperature in a humidified chamber with three washes for 5 min with 0.05% Tween 20 in PBS between each application. The cover slips were then mounted with a drop of Univar (BDH, Poole, UK) containing 100 mg/ml of 1,4-diazabicyclo[2,2,2]octane (DABCO, Sigma) as an antifading agent. Samples were analyzed with the EMBL laser confocal microscope (Stelzer et al. 1992), as described by Carmo-Fonseca et al. (1993).

Acknowledgments

We thank Dr. Ernst Stelzer for guidance and use of the laser confocal microscope, developed and constructed at EMBL (Heidelberg, Germany). This work was supported by the Medical Research Council (UK).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References


Fraser, J., J. Hurst, P. Collis, and F. Grosveld. 1990. DNase I hypersensitive sites 1, 2 and 3 of the human beta-globin dominant control region direct position-independent expression. Nucleic Acids Res. 18: 3503–3508.


Philipsen, S., S. Puziina, and F. Grosveld. 1993. The minimal requirements for activity in transgenic mice of Hypersensitive Site 3 of the β globin Locus Control Region. EMBO J. [in press].


