The EMBO Journal vol.11 no.8 pp.2961 – 2969, 1992

Identification of a stage selector element in the human γ-globin gene promoter that fosters preferential interaction with the 5′ HS2 enhancer when in competition with the β-promoter

Stephen M.Jane, Paul A.Ney, Elio F.Vanin[1], Deborah L.Gumucio[2] and Arthur W.Nienhuis

National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, MD 20892, [1]Genetic Therapy Incorporated, Gaithersburg, MD 20878 and [2]Department of Anatomy and Cell Biology, University of Michigan, Ann Arbor, MI 48109, USA

Communicated by F.Grosveld

The erythroid-specific enhancer within hypersensitivity site 2 (HS2) of the human β-globin locus control region is required for high level globin gene expression. We investigated interaction between HS2 and the γ- and β-promoters using reporter constructs in transient assays in human erythroleukemia (K562) cells. The β-promoter, usually silent in K562 cells, was activated by HS2. This activity was abolished when a γ-promoter was linked *in cis*. Analysis of truncation mutants suggested that sequences conveying the competitive advantage of the γ-promoter for HS2 included those between positions −53 and −35 relative to the transcriptional start site. This sequence, when used to replace the corresponding region of the β-promoter, increased β-promoter activity 10-fold when linked to HS2. The modified β-promoter was also capable of competing with a γ-promoter modified internally in the −53 to −35 region, when the two promoters were linked to HS2 in a single plasmid. The corresponding sequences from the Galago γ-promoter, a species which lacks fetal γ-gene expression, were inactive in analogous assays. We have identified and partially purified a nuclear protein found in human (fetal stage) erythroleukemia cells, but present in much lower concentration in murine (adult stage) erythroleukemia cells, that binds the −53 to −35 sequence of the γ-promoter. We speculate that this region of the γ-promoter functions as a stage selector element in the regulation of hemoglobin switching in humans.

Key words: β-globin/enhancer/hypersensitivity site/promoter/*trans*-acting factor

Introduction

The human β-like globin genes (ε, ^{G}γ, ^{A}γ, δ and β) exist in a cluster on the short arm of chromosome 11. The genes are expressed in a temporal and tissue specific fashion. In the early stages of human development, embryonic hemoglobin is synthesized in the yolk sac. During the fifth week of gestation a switch in β-chain subtype occurs and fetal hemoglobin (HbF) synthesis begins in the liver. HbF remains the predominant form of hemoglobin until birth, when a second switch occurs with a sharp increase in β-chain production in the bone marrow resulting in a preponderance of adult hemoglobin (HbA) by the fifth week of life (reviewed in Stamatoyannopoulos and Nienhuis, 1987).

Developmental regulation of the β-globin gene is mediated by several mechanisms. Sequences within or immediately flanking the genes themselves can confer tissue and temporal specificity (Chada *et al.*, 1985, 1986; Kollias *et al.*, 1986; Rutherford and Nienhuis, 1987). However, the level of expression is low and sensitive to the site of integration into the host genome (Townes *et al.*, 1985). High levels of tissue specific expression are restored when the genes are linked to the Locus Control Region (LCR), a 20 kb segment of DNA 5′ to the β-like globin genes characterized by four DNase I hypersensitivity sites (Tuan *et al.*, 1985; Grosveld *et al.*, 1987; Talbot *et al.*, 1989). Despite its importance in globin gene expression, the LCR's role in determining development specificity of the individual globin genes is unclear. Initial reports suggested that the LCR was capable of overriding the appropriate temporal regulation of linked γ- or β-genes in transgenic mice (Enver *et al.*, 1989; Behringer *et al.*, 1990). Restoration of developmentally specific expression of these genes could be achieved if the LCR was linked to both genes in the same construct, suggesting a competitive interaction between the two genes for LCR sequences (Behringer *et al.*, 1990; Enver *et al.*, 1990). More recent work has indicated that, although a competitive mechanism is still applicable for γ-gene silencing of the β-gene, the γ-genes appear to be silenced independently of the other globin genes, possibly by stage-specific factors binding to sequences flanking the genes (Dillon and Grosveld, 1991). In addition, the relative position of the genes with respect to the LCR appears to polarize the competitive interactions, potentially allowing the proximal γ-genes to silence the distal β-genes competitively, but not vice versa (Hanscombe *et al.*, 1991).

The ability of the γ-gene to silence a linked β-gene competitively, may be mediated by sequences analogous to the stage selector element (SSE) identified in the chicken β-globin promoter and centered at nucleotide position −50 (Choi and Engel, 1988). Temporal regulation in this system may be mediated by a developmentally specific *trans*-acting factor, NF-E4, which binds to the SSE (Gallarda *et al.*, 1989). Despite the recognition and characterization of several human erythroid nuclear regulatory factors (Mignotte *et al.*, 1989; Trainor *et al.*, 1990; Zon *et al.*, 1990) no stage-specific proteins comparable to the chicken system have been identified. In this paper we document the presence of a putative SSE in the proximal γ-promoter, capable of silencing linked β-promoter mediated expression. In addition, we have identified a protein with erythroid and developmental specificity that binds to this region of the promoter.

Results

Competition between the γ- and β-promoters for hypersensitivity site 2

DNase I hypersensitivity site 2 (HS2) confers high level tissue-specific expression on a linked β-globin gene in transgenic mice (Ryan *et al.*, 1989). Within HS2 is a powerful erythroid-specific enhancer required for the activity of HS2 in transgenic mice and active in transient transfection assays utilizing reporter genes (Tuan *et al.*, 1989; Ney *et al.*, 1990). We sought to determine whether HS2 was sufficient to override the developmental specificity of β-promoter mediated expression in a fetal environment, analogous to the effect seen with a μ-LCR linked to a β-globin gene in transgenic mice (Enver *et al.*, 1989; Behringer *et al.*, 1990). The human erythroleukemia cell line K562 was utilized as a model of fetal erythropoiesis, as it constitutively expresses the ε- and γ-genes but not the β-gene (Lozzio and Lozzio, 1975; Rutherford *et al.*, 1979; Benz *et al.*, 1980).

When transfected into K562 cells, a chimeric −376 β-promoter−CAT reporter gene construct (βCAT) was not expressed. In contrast, when the βCAT chimera was linked to a 1.45 kb fragment of the LCR containing the HS2 enhancer (HS2βCAT) (P.A.Ney, unpublished observation) β-promoter activity was observed (Figure 1). This expression was unchanged by co-transfection of a 6-fold molar excess of a construct containing a chimeric −260 γ-promoter−firefly luciferase reporter gene (γLUC) (Figure 2, constructs 1 and 2). Co-transfection of the γLUC chimera linked to the HS2 enhancer fragment (HS2γLUC) also had no effect (Figure 2, construct 3 and 4). In contrast, when the β and γ-promoters with reporter genes were linked *in cis* in the presence of a single HS2 fragment, β-promoter activity was reduced 10-fold, whereas γ-promoter activity remained unchanged (Figure 2, construct 5). This effect occurred irrespective of the orientation of the two promoters with respect to the enhancer (data not shown).

The ability of the γ-promoter to silence the β-promoter *in cis* but not *in trans* could be explained by two models. Firstly, the γ-promoter could exert a direct negative influence on the β-promoter only when the two are in close proximity in a single construct. Secondly, the γ-promoter could contain sequences which convey a competitive advantage for this promoter over the β-promoter in the interaction with the HS2 enhancer. To determine which of these models was applicable, a construct in which each promoter−reporter was linked 3′ to an HS2 fragment was transfected into K562 cells.

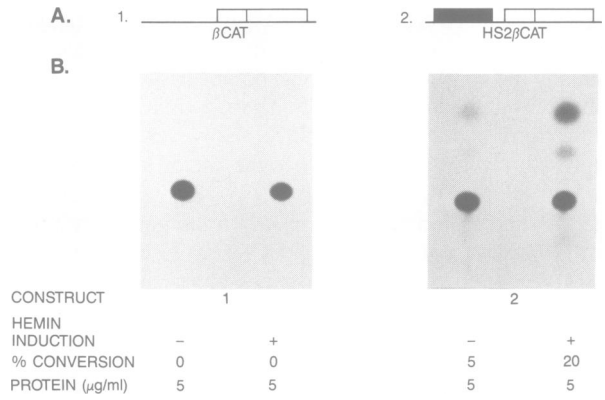

Fig. 1. HS2 activation of the β-globin promoter in K562 cells. (**A**) Diagrammatic representation of the constructs used for transfection into K562 cells. The solid box represents the 1.45 kb *Kpn*I−*Bgl*II HS2 fragment and the open box the β-promoter−CAT gene chimera. The vertical line in the chimeric gene represents the 3′ end of the −376 to +45 β-promoter fragment utilized. All numbers refer to positions relative to transcriptional start sites. (**B**) Reporter genes activity of the constructs described in (A). Equal volumes of extracts (10 μl) with protein concentrations as shown, were used for the assays.

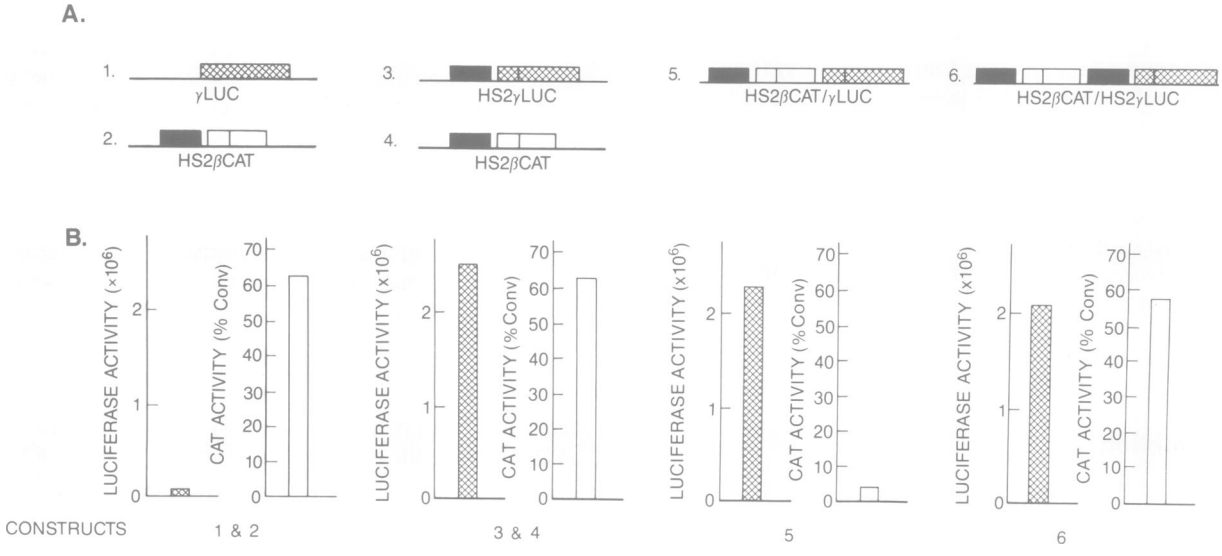

Fig. 2. Competition between the γ- and β-promoters for the HS2 enhancer. (**A**) Diagrammatic representation of the constructs used for transfection into K562 cells. The solid box represents the 1.45 kb *Kpn*I−BglIII HS2 fragment, the hatched box the −260 γ-promoter−luciferase gene chimera and the open box the −376 β-promoter−CAT gene chimera. The vertical lines in the chimeric genes represent the 3′ ends of the promoters. The γ-promoter fragment used was from −260 to +35 relative to the transcriptional start site. Constructs 1 and 2 (γLUC and HS2βCAT) were co-transfected. Increasing concentrations of γLUC (up to a 6-fold molar excess) were used in separate experiments. The result depicted was obtained with the highest concentration used and is representative of the other experiments. Constructs 3 and 4 (HS2γLUC and HS2βCAT) were co-transfected. Constructs 5 and 6 (HS2βCAT/γLUC and HS2βCAT/HS2γLUC were transfected separately. (**B**) Reporter gene activity of the constructs as transfected in experiments described in (A). The hatched bars represent luciferase activity, the open bars, CAT conversion. Equal volumes of extracts (10 μl) with equivalent protein concentrations (15 μg/ml) were used in each experiment.

As seen in Figure 2 (construct 6), β-promoter activity is restored with the provision of its own enhancer fragment, supporting the competition model proposed above.

Localization of sequences within the γ-promoter allowing preferential interaction with the HS2 enhancer

Truncation mutants from the 5′ end of the −260 γ-promoter were linked to the luciferase reporter gene and tested for their ability to suppress the β-promoter in the presence of HS2 in K562 cells. We chose to position the β-promoter adjacent to the HS2 fragment and 5′ to the γ-promoter. Although this orientation is opposite to that seen naturally in the β-globin cluster, we felt that any positional advantage would be conveyed to the β-promoter, thereby increasing the significance of any suppression. As seen in Figure 3, two regions of the γ-promoter appear to be important for its competitive advantage. Removal of sequences between −260 and −167 allows a small increase in β-promoter activity, which is reduced with further truncation to −137. Subsequent truncation of the γ-promoter to −53 maintained suppression of the β-promoter, despite removal of canonical promoter elements, including the CACCC and CCAAT boxes. Only with removal of sequence to the −35 position (6 bp 5′ to the TATA box) was β-promoter activity restored to the levels observed in the absence of the γ-promoter. Examination of the γ-promoter activity of these constructs revealed that the truncations were associated with a progressive reduction in γ-promoter strength. The −167 to −137 and −53 to −35 truncations induced the most profound reductions, the −35 promoter having barely detectable levels of activity. To ensure that the increase in β-promoter activity and loss of γ-promoter activity observed in the −53 to −35 γ-promoter truncation was due to reversal of the competitive advantage of the respective promoters and not merely a reflection of crippling of the γ-promoter, the γ-truncation mutants were linked to HS2 and transfected into K562 cells in the absence of a linked β-promoter. As seen in Table I, the truncation from −53 to −35 did not significantly reduce γ-promoter activity in the absence of the β-promoter, confirming that this sequence, although not necessary for γ-promoter activity is important for the fetal advantage of the γ-promoter when in competition with the β-promoter for the HS2 enhancer.

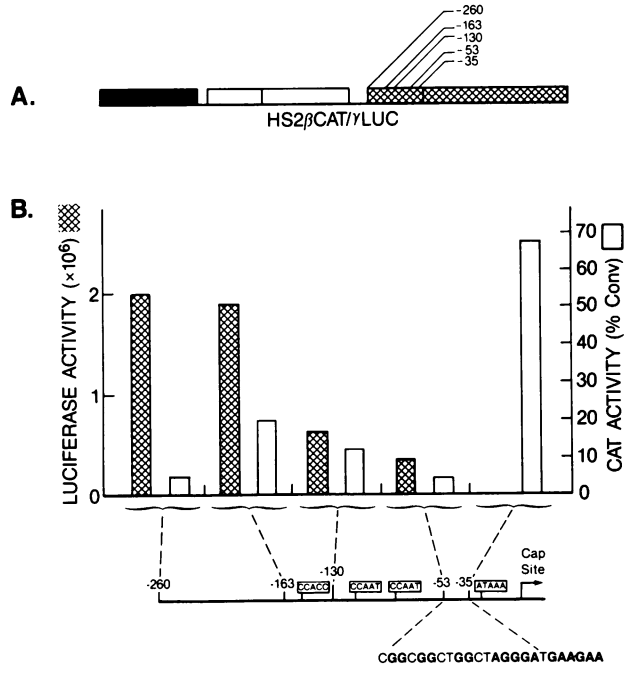

Fig. 3. Effects of truncation of the γ-promoter on its competition with the β-promoter for the HS2 enhancer. (**A**) Diagrammatic representation of the γ-promoter truncation mutant constructs transfected into K562 cells. The solid box represents the *KpnI*−*BglII* HS2 fragment and the open box the β-promoter−CAT gene chimera. The hatched box represents the γ-promoter truncations−luciferase gene chimeras. The truncation numbers refer to the position relative to the transcriptional start site of the γ-promoter. The plasmids are otherwise identical. (**B**) Reporter gene activity of the corresponding truncation constructs depicted in (A). The hatched bars represent luciferase activity, the open bars, CAT conversion. Equal volumes (10 μl) of extract with protein concentrations of 15 μg/ml were employed in all assays. The −260 γ-promoter showing the position of the truncations relative to the canonical promoter elements is shown in the lower portion of the figure. The sequence between −53 and −35 is expanded.

Further evaluation of the −53 to −35 γ-promoter region

To assess the activity of the region between −53 and −35 in the context of the −260 γ-promoter, the 18 bp region was mutated, and this modified promoter tested for its ability to compete with a wild-type β-promoter. As shown in Figure 4 (construct 2), mutation of this region resulted in a 2-fold increment in β-promoter activity. This increase was less than we predicted based on the truncation data, suggesting that there may be other sequences in the γ-promoter, presumably in the region between −260 and −167 (see Figure 3) that favor its interaction with the HS2 region. The −53 to −35 γ-sequence was inserted into the corresponding position in the β-promoter (replacing the β-sequence) and this hybrid promoter was examined for its ability to compete with either the modified γ-promoter (with the −53 to −35 mutated) or a wild-type β-promoter for HS2. As seen in Figure 4 (construct 3), addition of the 18 bp of γ-sequence to the β-promoter allows it to compete effectively with the previously dominant γ-promoter. In contrast, the wild-type β-promoter is incapable of providing any significant competition for the hybrid (construct 4) suggesting that this is 18 bp γ-sequence is sufficient to provide a competitive advantage. Despite the increase in activity of the hybrid β-promoter, no alteration in γ-promoter strength was observed in any of these constructs, suggesting that interaction of HS2 with other regions of the γ-promoter suffices to maintain promoter activity.

To study further the role of the −53 to −35 region of the γ-promoter as a developmentally-specific element, the hybrid β-promoter containing the 18 bp γ-sequence was linked to HS2 in the absence of a competing γ-promoter, and transfected into K562 cells. As shown in Figure 5, the

Table I. Comparison of −53 and −35 γ-promoters

		CAT activity (% conversion)	Luciferase activity
HS2 γ-luciferase	−53	−	8.0×10^5
	−35	−	4.9×10^5
HS2 β-CAT−γ-luciferase	−53	5	3.4×10^5
	−35	68	3.3×10^2

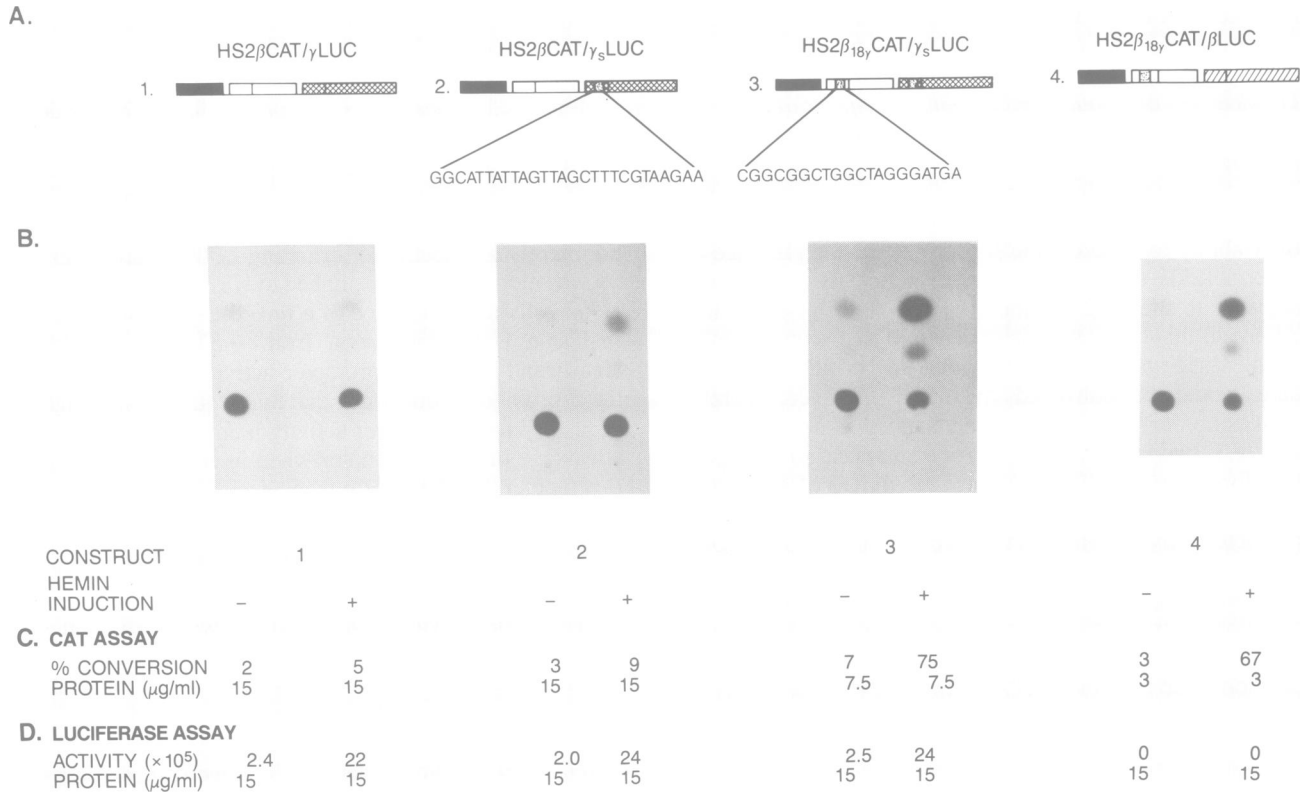

Fig. 4. Effects of mutation of the −53 to −35 γ-promoter sequence and replacement of the −53 to −35 β-promoter sequence with the corresponding γ-sequence on competition for the HS2 enhancer. (**A**) Diagrammatic representation of the constructs transfected into K562 cells. The solid box represents the *Kpn*I−*Bgl*II HS2 fragment, the hatched box the γ-promoter−luciferase gene chimera, the open box the β-promoter−CAT gene chimera and the lined box the β-promoter−luciferase gene chimera. Mutations and insertions in all constructs are represented by a stippled box. Construct 1 (HS2βCAT/γLUC) contains wild-type γ- and β-promoter sequence. Construct 2 (HS2βCAT/γsLUC) contains a mutation of the region between −53 and −35 in the γ-promoter, but is otherwise identical to construct 1. The mutated sequence is expanded. Construct 3 (HS2β18γCAT/γsLUC) contains a replacement of the −53 to −35 β-sequence with the corresponding γ-sequence, in addition to the −53 to −35 mutation of the γ-promoter depicted in construct 2. The location of the γ-sequence within the β-promoter is expanded. Construct 4 contains a hybrid β-promoter identical to construct 3 and a wild-type β-promoter. (**B**) CAT assays of the constructs outlined in (A). In each assay 10 μl of extract at the stated protein concentration was used.

hybrid β-promoter (construct 2) was 8- to 10-fold more active in K562 cells than the wild-type β-promoter (construct 1). To test the specificity of this γ-promoter sequence, and potentially to localize the activity further, we made two other hybrids, one in which the 18 bp γ-sequence was replaced with an unrelated 18 bp sequence, and one in which we inserted the −53 to −35 region of the Galago monkey γ-promoter, a species lacking a fetal stage of expression (Tagle *et al.*, 1988). As seen in Figure 5, an irrelevant mutation in the 18 bp γ-region or insertion of the Galago region (constructs 3 and 4) did not result in activation of the β-promoter in K562 cells.

Analysis of proteins binding to the −53 to −35 region of the γ-promoter

Four probes were designed to evaluate protein binding to the −53 to −35 γ-promoter region. The 'wild-type' probe (Figure 6A, line 1) is the sequence between −71 and −37 relative to the transcriptional start site. The Galago probe is the corresponding sequence in the Galago γ-promoter (Figure 6A, line 2). The mutant probe (Figure 6A, line 3) differed from the wild-type only in mutation of the 8 bp (GGCTGGCT) between −49 and −41 and the Sp1 probe is a consensus sequence from the SV40 enhancer (Figure 6A, line 4). The Sp1 probe was selected because of an Sp1 consensus binding sequence in this region of the γ-promoter coupled with prior evidence that Sp1 bound to this region (Gumucio *et al.*, 1991). Gel mobility shift analysis with crude K562 extract and the wild-type probe revealed three major retarded bands (Figure 6B, lane 1). The upper and lower bands co-migrated with the complexes observed with the Sp1 probe (lane 3) and were effectively competed by excess of non-radiolabelled Sp1 (lane 4). The middle band (designated −50γ) did not bind to the Sp1 probe (lane 3) and was not competed off with excess cold Sp1 probe (lane 4), suggesting that this complex is unrelated to Sp1. Competition with excess non-radiolabelled wild-type probe effectively competed both −50γ and Sp1, confirming the specificity of these interactions. Assays using the mutant probe with K562 extract failed to demonstrate the −50γ protein complex (lane 2), indicating that the GGCTGGCT motif was integral for its DNA binding.

To evaluate the distribution of the −50γ protein complex, wild-type and mutant probes were used to assay crude non-erythroid (HeLa) and adult erythroid [mouse erythroleukemia (MEL) extracts (Figure 6B, lanes 6−9)]. These extracts contained only trace amounts of the −50γ binding activity compared with that observed in K562 cells.

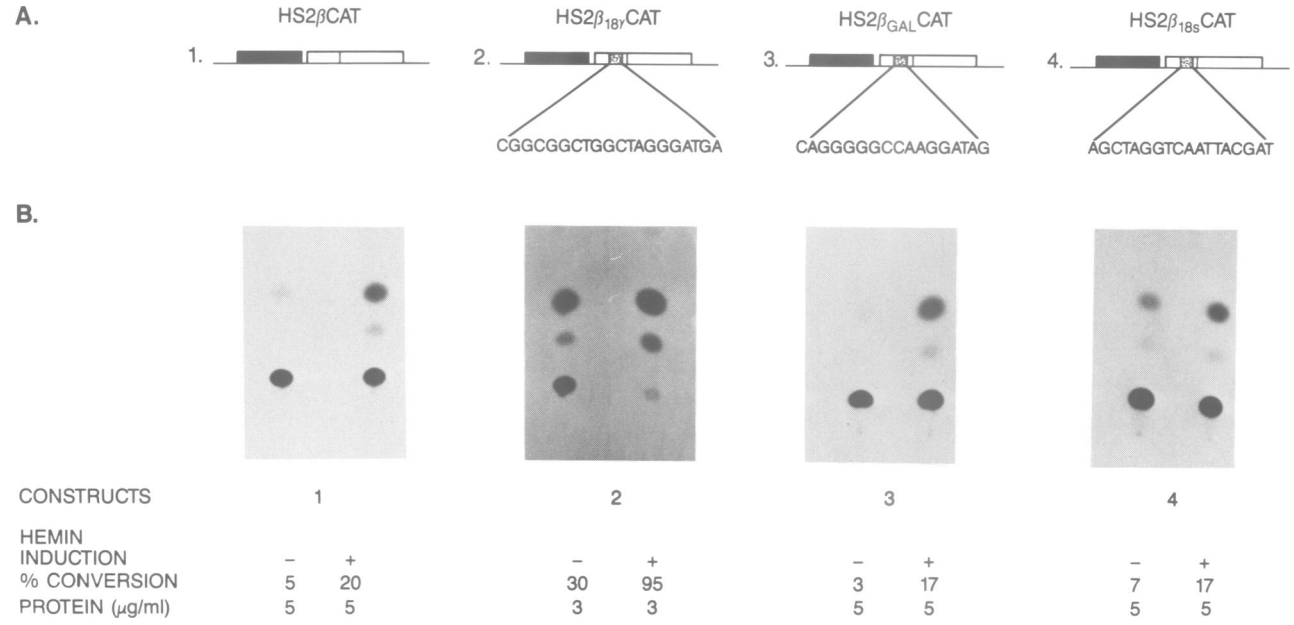

Fig. 5. Specificity of the −53 to −35 region of the γ-promoter in activating the β-promoter when linked to HS2 (**A**) Diagrammatic representation of the constructs transfected into K562 cells. The solid box represents the *KpnI*−*BglII* HS2 fragment and the open box the β-promoter−CAT gene chimera. Mutations and insertions in all constructs are represented by a stippled box. Construct 1 (HS2βCAT) contains wild-type β-promoter sequence. Construct 2 (HS2β18γCAT) contains an insertion of the −53 to −35 γ-sequence into the corresponding position in the β-promoter. Construct 3 (HS2βGALCAT) contains the analogous region of the Galago γ-promoter inserted into the corresponding position in the β-promoter. Construct 4 (HS2β18sCAT) contains a mutated 18 bp sequence unrelated to the globin genes transposed into the −53 to −35 region of the β-promoter. The sequence of all inserts has been expanded. (**B**) CAT assays of the constructs outlined in (A). 10 μl of extract at the stated protein concentration were used in all experiments.

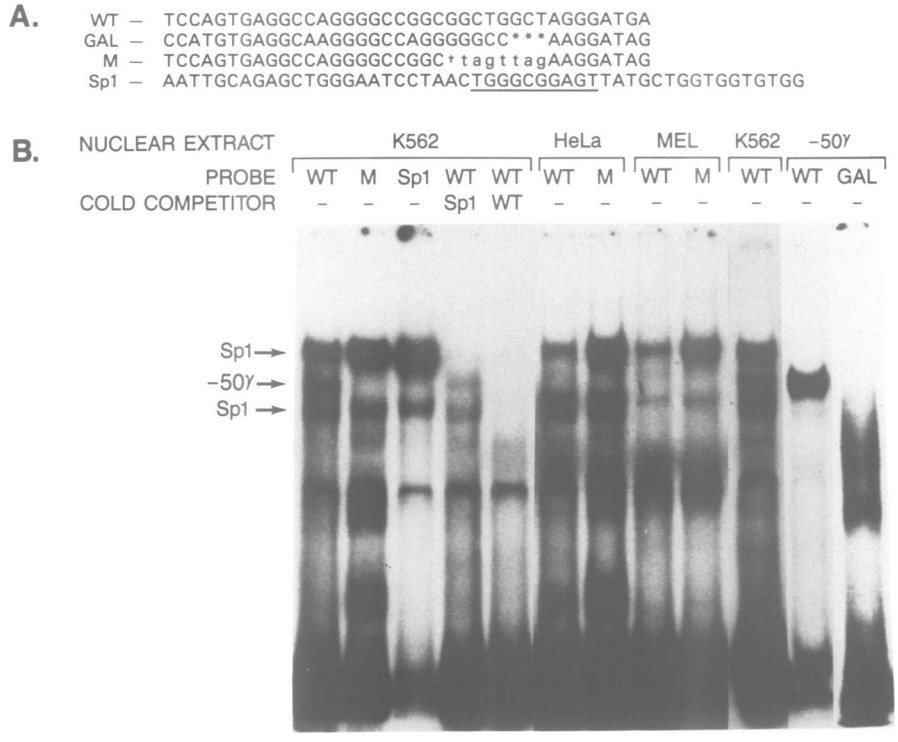

Fig. 6. Proteins binding to the −50 region of the γ-promoter. (**A**) Probes employed to study the −50 region: line 1 (WT), 35 bp oligonucleotide consisting of wild-type proximal γ-promoter sequence (−71 to −37 relative to the transcriptional start site); line 2 (GAL), a Galago proximal γ-promoter probe; line 3 (M), a mutant probe identical to the wild-type except for an 8 bp scramble between −49 and −43; line 4 (Sp1), an Sp1 probe derived from the SV40 enhancer (the SP1 consensus sequence is underlined). (**B**) Gel mobility shift assays were performed with a constant amount (5 μg) of nuclear extract as stated, in the presence of the specified probes. Cross competition experiments were performed with a 200 molar excess of unlabelled double-stranded oligonucleotides as specified. The partially purified −50γ protein was used as the extract source in lanes 11 and 12. Details of the purification are provided in the text. The Sp1 related complexes and the −50γ complex are labelled.

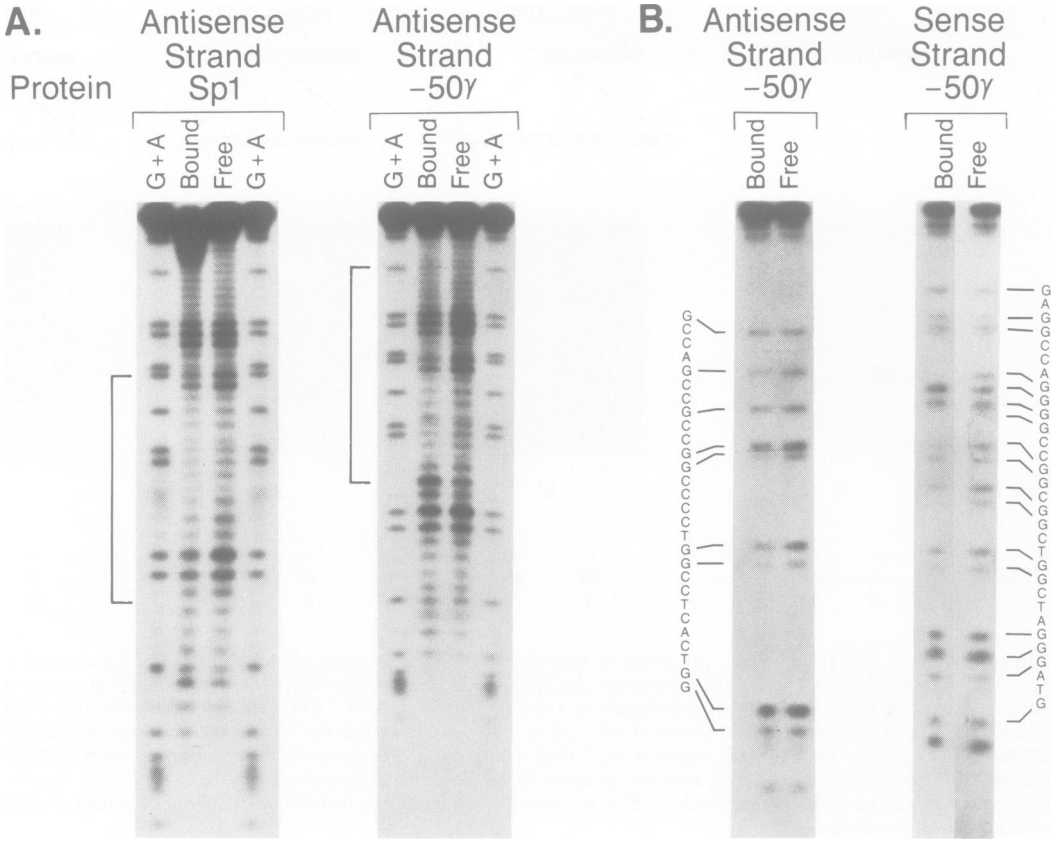

Fig. 7. Footprinting studies of purified Sp1 and the −50γ protein. (**A**) *In situ* copper−phenanthroline footprinting. A promoter fragment (−71 to −37) was labelled at the 3′ end of the antisense strand with Klenow polymerase, and incubated with purified Sp1 (1 μl) or partially purified −50γ (1.5 μl). Following gel mobility shift assay, the entire gel was subjected to copper−phenanthroline cleavage as previously described (Kuwabara and Sigman, 1987). DNA was then isolated from the retarded complexes and run on a sequencing gel. The respective footprint margins are indicated by square brackets. (**B**) Methylation interference analysis of −50γ. The probe used in (A) was uniquely end-labelled at the 3′ end of both strands, lightly methylated with dimethyl sulfate and complexed with partially purified −50γ (1.5 μl). Following gel mobility shift assay the DNA was isolated from the retarded complexes and run on a sequencing gel as previously described (Maxam and Gilbert, 1980). The probe sequence with guanine contact points is shown adjacent to each assay.

Comparison of Sp1 and −50γ binding sites in the proximal γ-promoter

Partial purification of the −50γ protein was performed to facilitate further characterization. Gel mobility shift analysis of the partially purified fractions showed a single retarded complex which co-migrated with −50γ from the crude extracts (Figure 6, lanes 10 and 11). The purified protein failed to bind the Galago promoter probe (Figure 6, lane 12). Utilizing this material and a commercially purified preparation of Sp1 (Promega), footprint comparison using the copper orthophenanthroline method was performed. As seen in Figure 7A, distinct but overlapping footprints were seen with the two proteins, the −50γ footprint encompassing the GGCTGGCT region of the promoter. Methylation interference of the partially purified −50γ showed contact on the sense strand at seven guanine residues (including the four contained in the GGCTGGCT motif) and at six on the antisense strand (Figure 7B). The same region of the Galago promoter lacks six of the contact guanines. A summary of the protein binding data is depicted in Figure 8.

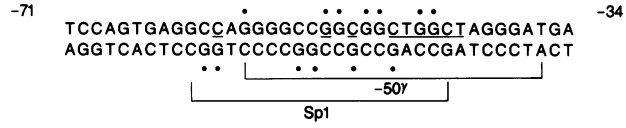

Fig. 8. Summary of footprinting data of Spl and the −50γ protein. Continuous open brackets indicate bases footprinted by Sp1 and −50γ. Closed circles indicate the guanines which are contacted by −50γ on the sense and antisense strands. The bases mutated or deleted in the Galago promoter are underlined.

Discussion

Competition between the genes of the β-globin locus for activating sequences in the LCR has been evoked as an integral mechanism in hemoglobin switching in man (Blom *et al.*, 1989). Initial evidence in transgenic mice suggested that both γ- and β-genes were silenced by this mechanism (Behringer *et al.*, 1990; Enver *et al.*, 1990). However, recent studies indicate that correct temporal expression of γ- and ε-genes can occur independently of other genes within the locus, presumably mediated by stage specific factors binding to sequences flanking the genes (Shih *et al.*, 1990; Dillon and Grosveld, 1991). The competition model of developmental regulation does hold for the β-gene as silencing of this gene in the fetal stage of development is not autonomous, relying on the presence of an active γ-gene (Dillon and Grosveld, 1991; Hanscombe *et al.*, 1991).

The studies reported here support the model of β-gene

silencing by the γ-gene in the fetal stage and identify stage selector sequences within the γ-promoter capable of mediating this effect. Our studies were performed utilizing the HS2 region of the LCR in view of the powerful erythroid-specific enhancer, active in transient assays, contained within this fragment (Tuan *et al.*, 1989; Ney *et al.*, 1990). Human erythroleukemia cells (K562) which express the γ- but not β-genes served as a surrogate for fetal stage erythroid cells (Lozzio and Lozzio, 1975; Rutherford *et al.*, 1979; Benz *et al.*, 1980). Developmental regulation of β-promoter mediated transcription in K562 cells was lost when the promoter was linked to HS2 (Figure 1). In keeping with a competition model, restoration of the developmental specificity of expression occurred when the γ-promoter was linked *in cis* in the presence of a single HS2 enhancer, but not when each promoter was provided with its own enhancer fragment (Figure 2B). Recent work by Hanscombe *et al.* (1991) indicates that the relative position of the β-gene with respect to the LCR is pivotal in the competition between γ- and β-genes. The suppression of β activity by the γ-promoter in our single enhancer construct (Figure 2B) was thus even more significant as the orientation we used would potentially bias the competition against the γ-promoter.

Competition between promoters for activating sequences as a mechanism of developmental regulation is not unique to the human globin system; indeed, the work of Choi and Engel (1988) provided the rationale for our experimental design. These investigators demonstrated in the chicken globin system that ε- and β-globin genes compete for the β-globin enhancer, with resultant silencing of the ε-gene in definitive erythroid cells. The β-globin enhancer is similar to HS2 in that it functions as a locus control element conferring erythroid tissue specificity and site independent expression on linked genes (Hesse *et al.*, 1986; Reitman *et al.*, 1990). The two enhancers also have a similar distribution of protein binding sites, raising the possibility that HS2 may play a role analogous to that of the chicken β-globin enhancer in developmental regulation (Reitman and Felsenfeld, 1988). The competitive advantage of the chicken β-promoter for the enhancer in the adult stage is endowed by an genetic element within the promoter, the SSE. We have demonstrated that the region in the human γ-promoter between nucleotides −53 and −35 relative to the transcriptional start site functions as a human SSE, allowing preferential interaction of the γ-promoter with HS2 when in competition with the β-promoter in a fetal environment (Figure 3). The studies in the chicken utilized pure primary cells in either primitive or definitive stages of development and thus examined the transition between these stages. Although our studies have demonstrated competitive regulation of the β-promoter by the γ-promoter only at a static developmental time point, i.e. the fetal stage of development (analgous to the chicken embryonic stage), a recent study by Morley *et al.* (1991) in an adult erythroid cell model (MEL) has shown that competition for HS2 between γ- and β-promoters does not occur at that stage of development. These results indicate that the γ-promoter silencing of the β-promoter in the fetal environment is a specific developmental phenomenon. The fetal specificity of the SSE was further highlighted by the use of chimeric β-promoter constructs, analogous to the approach of Lin *et al.* (1987). The insertion of the 18 bp γ-sequence into the corresponding position in the β-promoter resulted in a dramatic increase in activity compared with the wild-type β-promoter in K562 cells (Figure 4B and Figure 5A and B).

Several studies have demonstrated that stage-specific elements responsible for fetal expression of the γ-genes in tissue culture cells and transgenic mice reside in the γ-promoter (Kollias *et al.*, 1986; Donovan-Peluso *et al.*, 1987; Rutherford and Nienhuis, 1987; Trudel *et al.*, 1987). Perez-Stable and Constantini (1990), using truncation mutants of the γ-promoter, localized this activity to the −383 to −136 region. Although these experiments were performed in the absence of LCR components and a competing β-promoter, and thus were not analogous to our studies, it is possible that this region contains a redundant element thereby explaining the small rise in β-activity that we observed with the truncation from −260 to −167 (Figure 3), and the less than maximal effect observed with mutation of the −53 to −35 region in the context of the −260 γ-promoter (Figure 4B). We are currently evaluating the −260 to −167 region of the γ-promoter using our dual promoter assay system.

Interaction between stage-specific *cis*-acting sequences and *trans*-acting regulatory factors is important in developmental regulation (Emerson *et al.*, 1987; Lewis *et al.*, 1988; Gallarda *et al.*, 1989; Jackson *et al.*, 1989). However, none of the known human nuclear regulatory factors have been shown to influence the temporal specificity of globin gene expression. The protein(s) binding to the −50 region reported in these studies may play a role analogous to that of the chicken protein NF-E4 which binds to the SSE in that species (Gallarda *et al.*, 1989). Both the −50γ and NF-E4 proteins bind in the proximal promoter to sequences rich in guanines, although there is no direct homology in their consensus sequences. NF-E4 also binds to the chicken β-globin enhancer, presumably mediating close contact between the enhancer and SSE. In preliminary gel mobility shift assays we have demonstrated binding of the −50γ protein to an enhancer-containing probe (data not shown). Further evaluation of this interaction is under way. The −50γ protein has relative erythroid and developmental specificity, suggesting that levels of the factor may be important in determining its effect. The very low level of the −50γ protein seen in MEL extracts would explain the inability of the γ-gene to suppress a linked β-gene in the presence of HS2 when stably transfected into that cell line (Morley *et al.*, 1991).

The potential relevance of the −50γ protein to developmental regulation is further highlighted by studies comparing sequence homology in the γ-promoter of species with γ-genes that are expressed only embryonically (including the Galago) to those with a fetal stage of expression (Tagle *et al.*, 1988; Gumucio, in press). This phylogenetic footprinting demonstrates a breakdown in homology between these two groups in the −50 region, including multiple deletions and mutations affecting the −50γ protein contact bases in species without a fetal stage of γ-gene expression (Figure 8). The significance of this divergence is emphasized by two results. Firstly, the Galago γ-promoter probe is incapable of binding the partially purified −50 protein (Figure 6, lane 12). Secondly, the −53 to −35 region of the Galago γ-promoter (unlike the human γ-promoter) is incapable of conveying fetal properties when introduced into the β-promoter (Figure 5B). These findings would predict that β-gene expression in the Galago would be observed in the fetal period due to lack of competition

from an active γ-gene, a premise confirmed in the studies of Tagle *et al.* (1988).

The proximity of Sp1 and the −50γ protein's binding sites raises the possibility that competition between the two proteins for the overlapping sites may be important (Figure 6, lane 2, and Figures 7 and 9). A precedent for this type of interaction is seen with the PAL and CACCC proteins which bind to overlapping sites in the chicken β-promoter in an anticooperative manner, thereby influencing the developmental expression of the β-globin gene (Jackson *et al.*, 1989). Further characterization of the interaction between Sp1 and the −50 protein awaits purification of the latter.

Our findings, coupled with the evolutionary data outlined in this manuscript, support the conclusion that the proximal γ-promoter contains an SSE which regulates globin gene expression through interaction with a developmentally specific protein. Additional testing of this hypothesis involving mutation of the SSE in transgenic mice is currently in progress.

Materials and methods

DNA construction

pUC007, a pUC-based plasmid with a novel polylinker, was used as the vector in all constructions. The polylinker contains *Xho*I, *Bgl*II, *Xba*I, *Hin*dIII, *Sma*I, *Bam*HI, *Apa*I, *Sph*I and *Sal*I sites subcloned into the *Hin*dIII and *Eco*RI sites of pUC9 (Ney *et al.*, 1990).

HS2 was subcloned as a 1.45 kb *Kpn*I (blunted)−*Bgl*II fragment into the *Bgl*II (blunted) and *Bam*HI sites of pUC007. The γ-globin promoter was subcloned as a *Hin*fI(−260 relative to the cap site)−*Alu*I(+36) fragment into the *Bgl*II and *Hin*dIII sites of pUC007 via an intermediate step into the *Hin*cII site of pUC9. The coding region of the firefly luciferase gene (pSV232A-LA delta5′; deWet *et al.*, 1987) was then linked 3′ to the γ-globin promoter as a *Hin*dIII−*Bam*HI fragment in pUC007. An *Xho*I−*Aat*II fragment of this plasmid containing the chimeric γ-globin promoter/luciferase gene was subcloned into the *Sal*I−*Aat*II sites of pUC007 3′ to HS2. The β-globin promoter was subcloned as a *Bam*HI (−1.4 kb relative to the cap site)−*Nco*I(blunted to +38) fragment into pUC9. Sequences from the *Bam*HI site to the *Acc*I site (at −385) were removed and the ends blunted and religated. The coding sequence of the CAT or firefly luciferase gene was then linked 3′ to the β-globin promoter as a *Hin*dIII−*Sal*I(blunted) fragment in the *Sma*I(blunted) site in pUC007. An *Xho*I−*Aat*II fragment of this plasmid containing the chimeric β-globin promoter−reporter gene was subcloned into the *Sal*I−*Aat*II sites of pUC007 3′ to HS2. The dual promoter (HS2 βCAT/γluciferase, HS2βCAT/βluciferase) and dual HS2 (HS2 βCAT/HS2 γluciferase) constructs were obtained by the *Xho*I/*Sal*I-*Aat*II strategy described above.

Truncation mutants of the γ-promoter (−167, −137 and −53) were made by linearizing a plasmid containing the promoter, brief digestion with BAL-31, blunting of the nibbled ends and religation utilizing linkers. The −35 truncation was made by subcloning an oligonucleotide with a 5′ *Bgl*II site and 3′ *Hin*dIII site into the same sites in the −53 truncation mutant, thereby removing the wild-type sequences between −53 and −35. The truncations were then subcloned using the *Xho*I/*Sal*I−*Aat*II strategy 3′ to either the HS2 fragment or an HS2 β-promoter−CAT gene fragment, as described above.

Oligonucleotides

Oligonucleotides were synthesized on an Applied Biosystems DNA Synthesizer model 380B using phosphoramidite chemistry. The oligonucleotides were chromatographed on Sephadex G-25 columns (Pharmacia) and purified using denaturing polyacrylamide gel electrophoresis. Complementary strands were annealed in 10 mM Tris−HCl, pH 7.5, 50 mM NaCl, and 10 mM $MgCl_2$ by heating to 95° C for 2 min and slow cooling to room temperature. When subcloned into plasmids, sequence was verified by the chain termination method using Sequenase (USB).

Cell lines and nuclear extracts

Human erythroleukemia (K562) cells, murine erythroleukemia cells (MEL) and suspension HeLa S3 cells were grown in Improved Minimal Essential Medium (IMEM, Biofluids) with 10% fetal calf serum. Cells were cultured in suspension in microcarrier flasks (Wheaton) and harvested in late log phase. Extracts were prepared by the method of Dignam *et al.* (1983) in the presence of PMSF, aprotinin, pepstatin and leupeptin, dialyzed against 20 mM HEPES, pH 7.8, containing 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF and 20% glycerol and stored in small aliquots at −70°C. Protein concentration was determined by the Bradford method (Bradford, 1976) using a kit from Bio-Rad and bovine serum albumin as a standard.

DNA transfection and transient assays

Plasmid DNA was prepared by Triton X-100 lysis and purification over two cesium chloride gradients. Supercoiled DNA was introduced into K562 cells in log phase growth by electroporation (Potter *et al.*, 1984) using a Bio-Rad gene pulser apparatus. 4×10^7 cells were mixed with equimolar amounts of the test plasmids in the presence of carrier plasmid (pUC007) to a total of 50 μg in a final volume of 425 μl and electroporation at 960 μF and 0.20 V was performed. In initial experiments an internal control plasmid RSV-GPT was included to assess transfection efficiency; however, with the dual reporter constructs transcriptional factor availability was limiting. In these experiments, multiple repeats using at least two different preparations of DNA, standardization of extract protein concentration and comparison of luciferase values between constructs served as experimental controls. After transfection, cells were grown in the presence and absence of 20 μM hemin and harvested after 48 h. Cell lysates were prepared by repeated freeze−thaw cycles. Luciferase activity was measured on a Monolight 2001 luminometer (Analytical Luminescence Laboratories) (de Wet *et al.*, 1987). The assay was linear from 200 to 6×10^6 light units. Samples giving non-linear values were diluted and re-assayed. CAT activity was determined as previously described (Gorman *et al.*, 1982).

Gel mobility shift assay

Probes for these assays were prepared from oligonucleotides subcloned into pUC007, released by *Hin*dIII and *Bam*HI digestion, labelled with Klenow DNA polymerase and purified on 12% polyacrylamide gels. Assays were performed with 20 000 c.p.m. of probe added last to a 20 μl reaction containing varying amounts of nuclear extract, 500 ng of poly[d(I−C)], 6 mM $MgCl_2$, 60 mM KCl and 100 μg of bovine serum albumin (Fried and Crothers, 1981; Strauss and Varshavsky, 1984). In some assays double-stranded oligonucleotides added in 200 molar excess served as non-radiolabelled competitors. After the reaction mixture had been incubated at 4°C for 15 min and at 25°C or 10 min, samples were electrophoresed in 4% non-denaturing polyacrylamide gels in 0.5 × Tris−borate−EDTA buffer for 90 min at 10 V/cm. The gels were dried and exposed to film for 6 h.

In situ copper−phenanthroline footprinting

After gel mobility shift assay, as described above, *in situ* copper−phenanthroline footprinting was performed as described by Kuwabara and Sigman (1987).

Methylation interference

Probes for methylation interference studies were lightly methylated with dimethyl sulfate. Following gel mobility shift, the gel was autoradiographed at 4°C, the band of interest excised from the gel and the DNA eluted and purified. The DNA was cleaved at dG and dA residues, and analyzed on 12% sequencing gels as described by Maxam and Gilbert (1980).

References

Behringer,R., Ryan,T.M., Palmiter,R.D., Brinster,R.L. and Townes,T.M. (1990) *Genes Dev.*, **4**, 380−389.

Benz Jr,E.J., Murnane,M.J., Tonkonow,B.L., Berman,B.W., Mazur,E.M., Cavallesco,C., Jenko,T., Snyder,E.L., Forget,B.G. and Hoffman,R. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 3509−3513.

Blom van Assendelft,G., Hanscombe,O., Grosveld,F. and Greaves,D.R. (1989) *Cell*, **56**, 969−977.

Bradford,M.M. (1976) *Anal. Biochem.*, **72**, 248−254.

Chada,K., Magram,J., Raphael,K., Radice,G., Lacy,E. and Constantini,F. (1985) *Nature*, **314**, 377−380.

Chada,K., Magram,J. and Constantini,F. (1986) *Nature*, **319**, 685−689.

Choi,O.B. and Engel,J.D. (1988) *Cell*, **55**, 17−26.

deWet,J.R., Wood,K.V., DeLuca,M., Helinski,D.R. and Subramani,S. (1987) *Mol. Cell. Biol.*, **7**, 725−737.

Dignam,J.D., Lebovitz,R.M. and Roeder,R.G. (1983) *Nucleic Acids Res.*, **11**, 1475−1489.

Dillon,N. and Grosveld,F. (1991) *Nature*, **350**, 252−254.

Donovan-Peluso,M., Acuto,S., Swanson,M., Dobkin,C. and Bank,A. (1987) *J. Biol. Chem.*, **262**, 17051−17057.

Emerson,B.M., Nickol,J.M., Jackson,P.D. and Felsenfeld,G. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4786–4790.
Enver,T., Ebens,A.J., Forrester,W.C. and Stamatoyannopoulos,G. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 7033–7037.
Enver,T., Raich,N., Ebens,A.J., Papayannopoulou,T., Constantini,F. and Stamatoyannopoulos,G. (1990) *Nature*, **344**, 309–313.
Fried,M. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6506–6525.
Gallarda,J.L., Foley,K.P., Yang,Z. and Engle,J.D. (1989) *Genes Dev.*, **3**, 1845–1859.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Grosveld,F., van Assendelft,G.B., Greaves,D.R. and Kollias,G. (1987) *Cell*, **51**, 975–985.
Gumucio,D.L., *Mol. Cell. Biol.*, in press.
Gumucio,D.L., Rood,K.L., Blanchard-McQuate,K.L., Gray,T.A., Saulino,A. and Collins,F.S. (1991) *Blood*, **78**, 1853–1863.
Hanscombe,O., Whyatt,D., Fraser,P., Yannoutsos,N., Greaves,D., Dillon,N. and Grosveld,F. (1991) *Genes Dev.*, **5**, 1387–1394.
Hesse,J.E., Nickol,J.M., Lieber,M.R. and Felsenfeld,G. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 4312–4316.
Jackson,P.D., Evans,T., Nickol,J.M. and Felsenfeld,G. (1989) *Genes Dev.*, **3**, 1860–1873.
Kollias,G., Wrighton,N., Hurst,J. and Grosveld,F. (1986) *Cell*, **46**, 89–94.
Kuwabara,M.D. and Sigman,D.S. (1987) *Biochemistry*, **26**, 7234–7238.
Lewis,C.D., Clark,S.P., Felsenfeld,G. and Gould,H. (1988) *Genes Dev.*, **2**, 863–873.
Lin,H.J., Anagnou,N.P., Rutherford,T.R., Shimada,T. and Nienhuis,A.W. (1987) *J. Clin. Invest.*, **80**, 374–380.
Lozzio,C.B. and Lozzio,B.B. (1975) *Blood*, **45**, 321–334.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 449–560.
Mignotte,V., Wall,L., deBoer,E., Grosveld,F. and Romeo,P.H. (1989) *Nucleic Acids Res.*, **17**, 37–54.
Morley,B.J., Abbott,C.A. and Wood,W.G. (1991) *Blood*, **78**, 1355–1363.
Ney,P.A., Sorrentino,B.P., McDonagh,K.T. and Nienhuis,A.W. (1990) *Genes Dev.*, **4**, 993–1006.
Perez-Stable,C. and Constantini,F. (1990) *Mol. Cell. Biol.*, **10**, 1116–1125.
Potter,H., Weir,L. and Leder,P. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 7161–7165.
Reitman,M. and Felsenfeld,G. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6267–6271.
Reitman,M., Lee,E., Westphal,H. and Felsenfeld,G. (1990) *Nature*, **348**, 749–752.
Rutherford,T. and Nienhuis,A.W. (1987) *Mol. Cell. Biol.*, **7**, 398–402.
Rutherford,T., Clegg,J.B. and Weatherall,D.J. (1979) *Nature*, **280**, 164–165.
Ryan,T.M., Behringer,R.R., Martin,N.C., Townes,T.M., Palmiter,R.D. and Brinster,R.L. (1989) *Genes Dev.*, **3**, 314–323.
Shih,D.M., Wall,R.J. and Shapiro,S.G. (1990) *Nucleic Acids Res.*, **18**, 5465–5472.
Stamatoyannopoulos,G. and Nienhuis,A.W. (1987) In Stamatoyannopoulos,G., Nienhuis,A.W., Leder,P. and Majerus,P. (eds), *The Molecular Basis of Blood Diseases.* W.B.Saunders, Philadelphia, pp. 67–105.
Strauss,F. and Varshavsky,A. (1984) *Cell*, **37**, 889–901.
Tagle,D.A., Koop,B.F., Goodman,M., Slightom,J.L., Hess,D.L. and Jones,R.T. (1988) *J. Mol. Biol.*, **203**, 439–455.
Talbot,D., Collis,P., Antoniou,M., Vidal,M., Grosveld,F. and Greaves,D.R. (1989) *Nature*, **338**, 352–355.
Townes,T., Lingrel,J., Chen,H., Brinster,R. and Palmiter,R. (1985) *EMBO J.*, **4**, 1715–1723.
Trainor,C.D., Evans,T., Felsenfeld,G. and Boguski,M.S. (1990) *Nature*, **343**, 92–96.
Trudel,M., Magram,J., Bruckner,L. and Constantini,F. (1987) *Mol. Cell. Biol.*, **7**, 4024–4029.
Tuan,D., Solomon,W., Li,Q. and London,I.M. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6384–6388.
Tuan,D.Y., Solomon,W.B., London,I.M. and Lee,D.P. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 2554–2558.
Zon,L.I., Tsai,S-F., Burgess,S., Matsudaira,P., Bruns,G.A.P. and Orkin,S.H. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 668–672.

Received on January 14, 1992; revised on April 30, 1992