Transcriptional activation by hypersensitive site three of the human β-globin locus control region in murine erythroleukemia cells

Sara Pruzina [1], Michael Antoniou, Jacky Hurst, Frank Grosveld [1],*, Sjaak Philipsen [1]

National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Abstract

In this paper we describe a complete deletional analysis of hypersensitive site three (HS3) of the human β-globin Locus Control Region (LCR). The previously defined core fragment consists of 6 footprinted regions, with multiple binding sites for the erythroid-specific factor GATA-1 and G-rich motifs that can interact with ubiquitous factors such as Sp1 and TEF-2. We show in this paper that the 5′ half of this fragment is the most important for activity in murine erythroleukemia (MEL) cells. A fragment containing footprints 1–4 can stimulate transcription of a linked human β-globin gene to levels of about 40% of that obtained with footprints 1–6. Constructs containing either footprints 1–3 or 3–6 cannot be distinguished from the β-globin gene alone. We further show that binding sites for the erythroid-specific factor NF-E2 can co-operatively interact with parts of the HS3 core fragment, and that HS3 requires elements upstream from −103 in the human β-globin promoter for full activity. The importance of these results is discussed in the context of the regulation of the genes in the human β-globin cluster.

Keywords: Transcription activation; beta-Globin; Deletion analysis; (Erythroleukemia cell); (Mouse)

1. Introduction

The human β-like globin genes are arranged in the same order as their developmental expression pattern, with the embryonic ε-globin at the 5′ end, the two fetal γ genes in the middle and the adult β-globin gene at the 3′ end [1]. All the genes are controlled by the Locus Control Region (LCR) which is located upstream of the ε gene, and covers approximately 13 kb of DNA. The presence of the LCR is an absolute requirement for the high level expression of the genes in the cluster [2,3]. The linkage of a human β-globin transgene to the LCR results in high level, erythroid specific expression in mice [4], in sharp contrast to constructs lacking LCR sequences which show a low, variable expression level [5–8]. The most characteristic feature of the LCR is its ability to override position

effects when it is stably integrated into chromatin. This results in expression levels that are dependent on the copy number of the transgene and independent of the site of integration in the host genome [4]. This indicates that it can open up silent chromatin in the erythroid environment, with the possible exception of heterochromatic regions (D.R. Greaves and F.G., unpublished data).

The LCR is characterized by the presence of four erythroid-specific, developmentally stable DNaseI hypersensitive sites [9]. We and others have previously shown that the activity of the LCR resides in these hypersensitive site (HS) regions and identified the core elements as 200–300 bp fragments [10–14]. There is strong evidence suggesting that each HS retains the essential properties of the LCR, although each site varies with regard to the level of transcriptional stimulation per gene and its developmental specificity [15,16].

In this paper, we have focussed our attention on the core fragment of HS3, which is transcriptionally the most active individual site of the LCR [15,17]. It is a 225 bp fragment which we have functionally defined in both transgenic mice and murine erythroleukemia (MEL) cells. It confers position independent expres-

* Corresponding author. Fax: +31 10 4360225.
E-mail: grosveld@hb.fgg.eur.nl.
[1] Present address: MGC, Department of Cell Biology and Genetics, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.

sion to a linked β-globin gene at an expression level of 40–50% of that achieved with the full LCR [12]. In contrast, this site is not active in a typical enhancer assay by transient transfections [18,19].

With in vitro DNaseI footprinting and gel-retardation analysis we found that this fragment has a modular structure [12] (Fig. 1B), composed of three footprints for the erythroid-specific transcription factor GATA-1 [20,21] interspersed by G-rich sequences which bind ubiquitous factors such as Sp1 [22] and TEF-2 [23]. Using PCR we have made sequential deletions of these footprints from the 5' and the 3' end, and analyzed these for activity in MEL cells. We then added binding sites for the erythroid-specific factor NF-E2 [24,25] to a subset of these deletions, in order to assess whether interactions occur between NF-E2 sites and the HS3 core fragments. The role of NF-E2 sites has been studied intensely in the context of HS2, another DNaseI hypersensitive site of the human β-globin LCR [10,14,26,27]. This site is transcriptionally less active than HS3 and has an entirely different architecture [14]. Most notable is a double binding site for NF-E2, which is required for high levels of transcription but which is dispensable for position independent expression [10,27]. There is a consensus sequence for NF-E2 just upstream from the core fragment of HS3. We were therefore interested to see whether binding sites for this protein can functionally substitute for deletions in HS3, or synergize with the entire core fragment.

Finally, we addressed the interplay of HS3 and upstream sequence elements of the human β-globin gene promoter by testing it in the context of the non-erythroid H-2K reporter gene [28–30].

2. Materials and methods

2.1. Plasmid constructions

All the oligonucleotides used for cloning purposes were phosphorylated with polynucleotide kinase. Plasmid GSE 1758 was made by cutting the plasmid GSE 1273 [31], containing the human β-globin gene as a 5 kb BglII fragment, with NotI and HpaI, blunting and ligating to a polylinker containing EcoRV, NotI, ClaI, HindIII, XhoI, SpeI, Asp718 and SalI sites, recreating the HpaI site at −815. (Fig. 1). HS 3 deletions were made by the polymerase chain reaction (PCR), using the following primers: to generate 3' deletions, oligo fp1, sense strand was used in all the reactions, in combination with either oligo fp2, 3, 4 or 5 (antisense), to create constructs containing footprints 1–2, 1–3, 1–4 and 1–5, respectively. 5' deletions were made by combining oligo fp6, antisense, with either oligo 2, 3, 4, or 5 (sense) to obtain constructs containing footprints 2–6, 3–6, 4–6 and 5–6, respectively. PCR was performed as recommended by the supplier of Taq polymerase (Anglian Biotechnology). PCR products were gel purified, blunted and ligated to HindIII linkers (3' deletions) or Asp718 linkers (5' deletions). After digestion with the appropriate restriction enzymes they were cloned in their natural orientation relative to the human β-globin gene in GSE 1758. Constructs containing dimers and trimers of footprints 5–6 were generated by digesting the construct containing the monomer with XhoI and SalI. The purified fragment was ligated to itself, digested with XhoI and SalI, and dimers and trimers were isolated on a 4% PAA gel. These fragments were cloned in GSE 1758 digested with XhoI
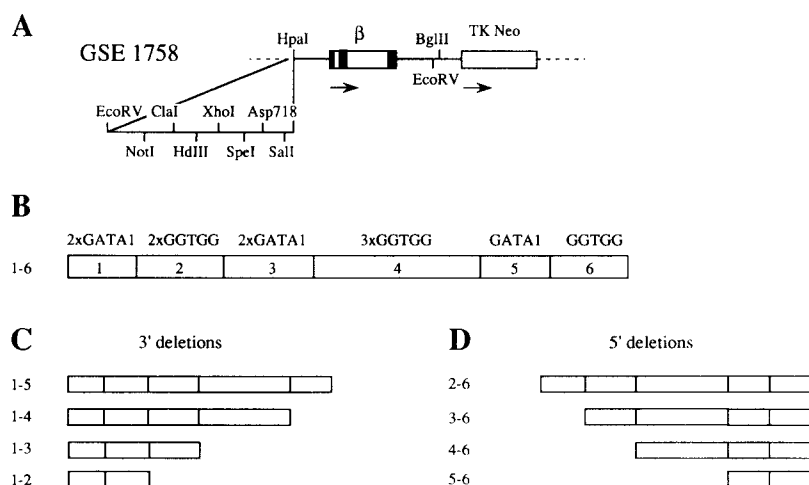

Fig. 1. The β-globin test plasmid and HS3 deletions. (A) GSE 1758 is a plasmid containing the human β-globin gene as a HpaI-BglII fragment, a TK promoter driving a neomycin resistance gene for selection of transfected cells and a polylinker at −815 in the β-globin gene with the indicated restriction sites. (B) Schematic representation of the 225 bp core fragment of HS3. Footprints are labelled 1 to 6. The occurrence of GATA-1 motifs and the sequence 'GGTGG' is indicated above each footprint. (C, D) Sequential deletion of footprints from the 3' (C) or 5' (D) end of the core fragment.

and SalI. All the constructs were checked by sequencing using a Sequenase kit (USB).

Plasmids containing the non-erythroid mouse H-2K gene have been described [28–30]. The HS3 core fragment containing footprints 1–6 was inserted at position −815 in the promoter of the human β-globin gene. In the minimal promoter construct, nucleotides −103 to −265 were deleted from the promoter region [30].

Sequences of oligonucleotides used:

For PCR:

oligo fp1, sense: aaGCTGGTGTGCCAGATGTGTC-TATCAGAGGTTCCAGGGAGG;

oligo fp 2, sense: aaATCAGAGGTTCCAGGGAGG-GTGGGGTGGGGTCAGGGCTGGCCAC;

oligo fp 2, antisense: ttTGGCCAGCCCTGACCCCA-CCCCACCCTCCCTGGAACCTCTGAT;

oligo fp 3, sense: aaTGGCCACCAGCTATCAGGGC-CCAGATGGGTTATAGGCTGG;

oligo fp 3, antisense: ttCCAGCCTATAACCCATCT-GGGCCCTGATAGCTGGTGGCCA;

oligo fp 4, sense: aaGCTCAGATAGGTGGTTAGGT-CAGGTTGGTGGTGCTGGGTGGAGTCCATGAC-TCCCAGGAGCCAG;

oligo fp 4, antisense: ttCTGGCTCCTGGGGAGTCAT-GGACTCCACCCAGCACCACCAACCTGACCTAA-CCA CCTATCTGAGC;

oligo fp 5, sense: aaCCATGACTCCCAGGAGC-CAGGAGAGATAGACCATGAGTAGAG;

oligo fp 5, antisense: ttCTCTACTCATGGTCTAT-CTCTCCTGGCTCCTGGGAGTCATGG;

oligo fp 6, antisense: ttCTATGCTGTGCCTCCCCCA-CCTTTCCCATGTCTCCCCTCTA;

HS3-NF-E2, sense: cgGAATTTTGACTCAGCAAA-CACAAGACCCTCat;

HS3-NF-E2, antisense: cgatGAGGGTCTTGTGTTT-GCTGAGTCAAAATTC;

HS2-NF-E2, sense: cgatAAGCACAGCAATGCTGA-GTCATGATGAGTCATGCTGAGGCTTa;

HS2-NF-E2, antisense: agcttAAGCCTCAGCATGAC-TCATCATGACTCAGCATTGCTGTGCTTat.

2.2. Transfection of MEL cells

About $5 \cdot 10^6$ cells were transfected by electroporation with 5–25 μg of plasmid DNA which was linearized with PvuI. The transfections were split imme-
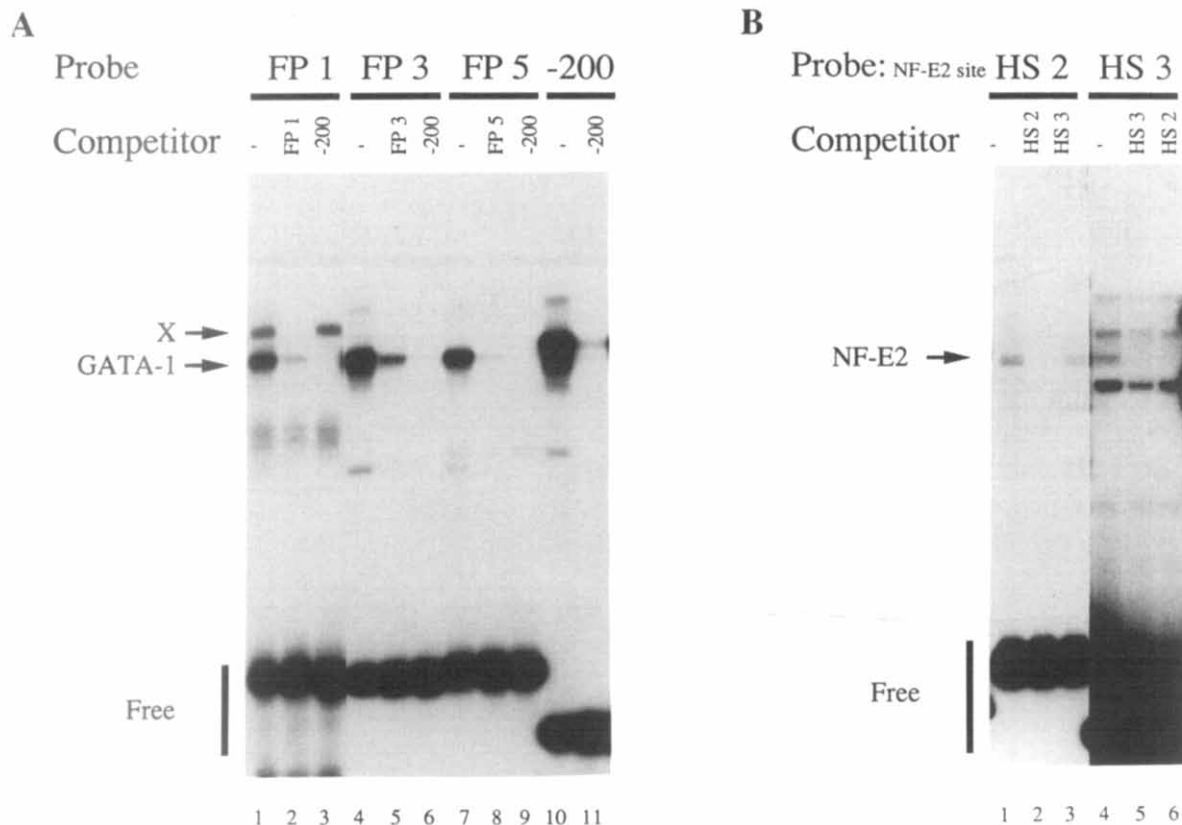

Fig. 2. Band shift assay of GATA-1 and NF-E2 binding sites. Probes and competitors are, as indicated above the lanes, bandshifts were performed using MEL nuclear extract. The positions of free probes and retarded complexes specific for GATA-1 and NF-E2 are shown; 'X' denotes a ubiquitous protein binding to footprint 1 only. Note that in panel B, the bandshift with the HS3 probe was exposed 10 times longer than the bandshift with the HS2 probe. '−200' refers to the oligonucleotide covering the GATA-1 site around position −200 in the β-globin promoter [30].

diately after electroporation to generate three independent populations for each construct. After selection in G418, cells were induced by the addition of DMSO (2% final concentration) to the media and further incubated for four days. DNA was isolated from uninduced cells, while RNA was isolated from induced cells [29].

2.3. DNA analysis

Genomic MEL DNA was cut with *Bgl*II, Southern blotted and probed for the human β-globin transgene, with the mouse Thy-1 gene as a loading control [4]. Only populations which harboured the intact transgene were further analyzed.
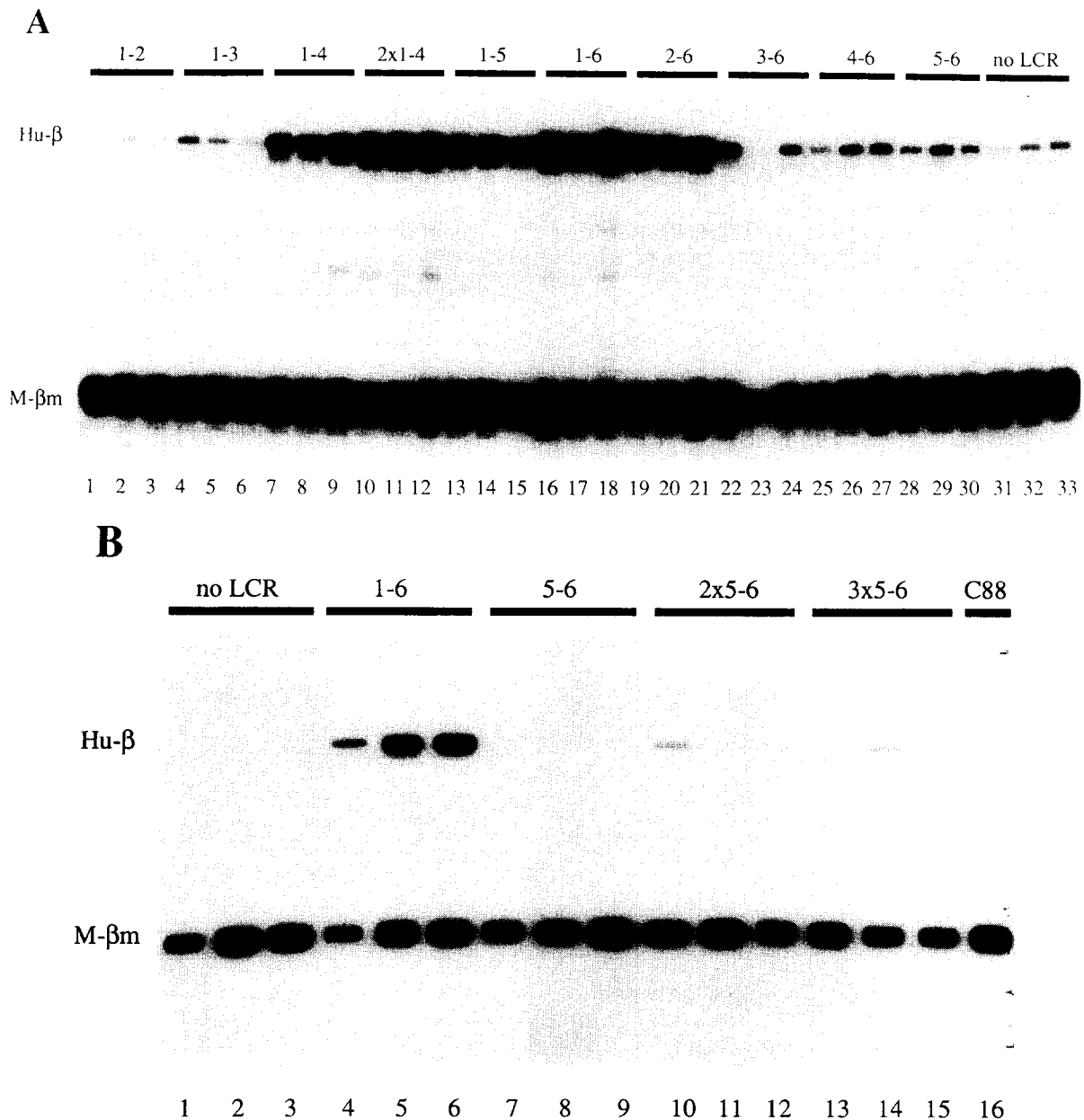

Fig. 3. Funtional analysis of HS3 deletions in MEL cells. 20 μg of total RNA from induced MEL populations, transfected with the constructs as indicated on top of the figure, was used for quantitative S1 analysis. The samples were hybridized to probes specific for the 5' ends of mouse β-major (M-βm; 96 nt protected) and human β-globin (Hu-β; 153 nt protected fragment) mRNA. They were then digested with S1 nuclease, and the protected fragments were separated on 6% sequencing gels. The protected bands were quantitated with the phosphorimager. Relative specific activities of the probes was 1 : 1.

2.4. RNA analysis

Total RNA from MEL cells was isolated as described [12]. 10–20 μg per sample was used for quantitative S1 nuclease analysis with mixed DNA probes specific for the 5′ ends of human β-globin [28] and mouse β-major mRNA [6], or β-H-2K hybrid mRNA [29]. Probe excess was demonstrated by using three times the amount of RNA of one particular sample. Expression data from S1 sequencing gels were quantitated using a Phosphor Imager (Molecular Dynamics).

2.5. Gel mobility shift assays

Gel mobility shift assays were done essentially as described previously [21,30], using 5–10 μg nuclear extract prepared from MEL cells per reaction. The oligonucleotides shown above were used as probes and competitors. Competitors were added in 100-fold molar excess prior to the addition of the extract.

3. Results

3.1. Deletional analysis of HS3

We have previously established that the majority of the activity of HS3 is retained on a 225 bp fragment, which contains 6 footprints, labelled 1–6 in Fig. 1. Footprints 1, 3 and 5 have consensus motifs for the erythroid-specific factor GATA-1, while footprints 2, 4 and 6 have the pentanucleotide sequence GGTGG in common. Fig. 2A shows that footprints 1, 3 and 5 are capable of binding GATA-1 in a gel retardation assay using nuclear extract from MEL cells. In addition, footprint 1 also binds a protein designated 'X', which is not erythroid specific since it is also present in HeLa cells (not shown). As we have shown previously, footprints 2, 4 and 6 bind Sp1, TEF-2 and a number of other unknown proteins; none of these appear to be erythroid specific [12].

We made sequential deletions of these footprints by PCR to obtain fragments containing footprints 1–5, 1–4, 1–3, 1–2 (3′ deletions) and 2–6, 3–6, 4–6, 5–6 (5′ deletions). These were cloned 815 bp upstream from the cap site of the human β-globin gene in the plasmid GSE 1758, which contains a neomycin resistance marker to allow selection of transformed cells with G418 (Fig. 1). After the fidelity of the resulting clones was confirmed by sequencing, plasmids were linearized with PvuI and stable transformants obtained by electroporating MEL cells and selection in G418 containing media. The cells were induced to differentiate by the addition of DMSO and harvested after 4 days. DNA and RNA were isolated, transgene integrity was monitored by Southern blotting and expression levels

were assayed by quantitative S1 nuclease protection analysis. The results are shown in Fig. 3A and summarized in Fig. 6.

It is clear that deletion of footprints from the 5′ end has a more severe effect than deletion of footprints from the 3′ end. The 5′ deletions show that removal of footprint 1 (construct 2–6 in Figs. 3A and 6A) results in the loss of 70% of the activity of the entire core fragment. This construct still expresses at a higher level than the gene alone. However, the additional deletion of footprint 2, leaving footprints 3–6, is indistinguishable from the control (Fig. 6A). The result with the 3′ deletions is that a construct without footprints 5 and 6 (construct 1–4 in Figs. 3A and 6A) expresses at 40% of the level obtained with footprints 1–6, significantly higher than the human β-globin gene without an LCR. This is confirmed by the expression level obtained with the construct containing a dimer of this fragment (construct 2 × 1–4 in Figs. 3A and 6A), which approaches the level obtained with all 6 footprints. Together these data imply that the 5′ half of HS3 is the most important for activity and that footprints 1, 2, 4, 5 and 6 are functional since their deletion results in decreased expression levels. We cannot assess the role of footprint 3 from these experiments, but an internal deletion of this sequence results in an inactive construct (not shown), suggesting that it does have functional significance.

From these data it appeared that multimers of GATA-1 sites and GGTGG motifs can functionally reconstitute the activity of HS3, and we therefore addressed this question directly by testing a monomer, dimer and trimer of footprint 5 and 6, the simplest combination of a GATA-1 site and a GGTGG motif found in HS3 (Fig. 1). The results (Figs. 3B and 6B) show that this does not restore activity. We have also attempted to restore activity by linking together oligonucleotides of footprints 1, 2, 3 and 4. However, this too failed (data not shown). This inactive construct harbours all the sequence information present in the active construct containing footprints 1–4 (Figs. 3A and 6A), but the spacing between the footprints has changed.

We conclude that simple combinations are inactive and do not retain the properties of HS3.

3.2. Interactions of HS3 with NF-E2 binding sites

We next investigated the potential interactions between HS3 and NF-E2, a factor that has been shown to be essential for the function of HS2 of the human β-globin LCR. There is a single consensus for NF-E2 just upstream from the HS3 fragment containing footprints 1–6, and we wished to establish first whether NF-E2 can bind to this sequence. We analyzed this by gel retardation, using the well characterized dimer
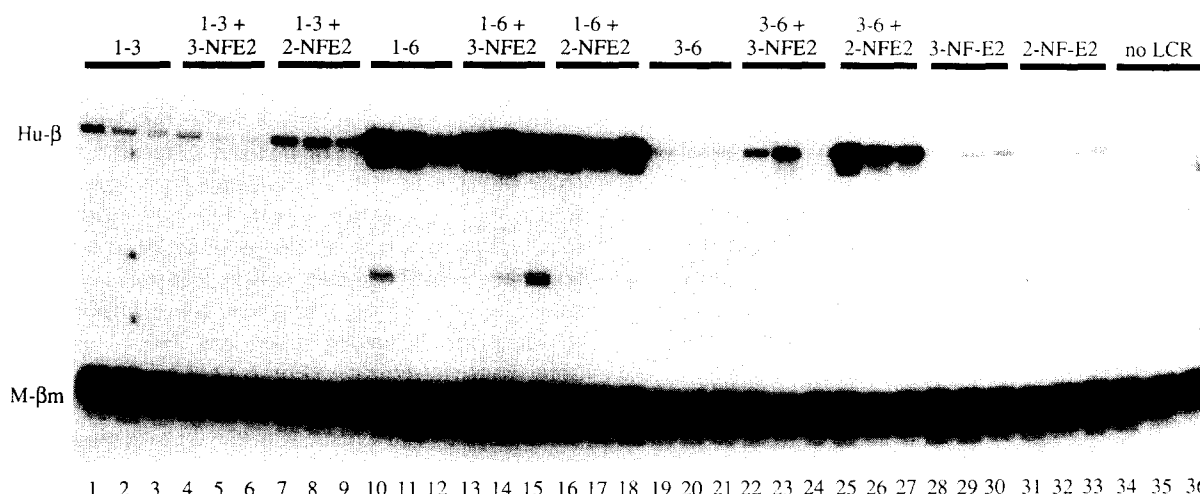
Fig. 4. Functional analysis of the addition of NF-E2 bindingsites to HS3 constructs. 3-NF-E2 is the NF-E2 site located 30 bp upstream of the core fragment of HS3; 2-NF-E2 is the NF-E2 dimer site of HS2. These were added to the HS3 constructs indicated. Other details are as in Fig. 3.

NF-E2 binding site of HS2 as a control [14]. As shown in Fig. 2B, the HS3 sequence is a genuine NF-E2 binding site, but it binds much less efficiently than the HS2 sequence. The reason for this is not known, but it might reflect the difference between a single (HS3) and a dimer (HS2) binding site.

Having shown that the site binds NF-E2 in vitro, we made a series of constructs in which we added the HS2 or HS3 NF-E2 binding sites to the plasmids containing footprints 1–3, 1–6 and 3–6, i.e., the entire core fragment and the largest 3' and 5' deletion constructs that did not give expression levels above background (Figs. 3A and 6). These constructs were then assayed in MEL cells as described above. The result is shown in Figs. 4 and 6C.

A moderate increase in the expression level is observed when just the NF-E2 binding sites are cloned upstream of the β-globin gene, which is in agreement

with previously published data [14,27]. Addition of the HS3 NF-E2 binding site to footprints 1–6 or 1–3 has no obvious effect on the expression level obtained with these constructs, while addition of the HS2 dimer site to these constructs results in a small increase of transcription.

Interestingly, when the NF-E2 binding sites are cloned upstream of footprints 3–6, we find a significant increase in the expression level. As expected, the HS2 NF-E2 dimer site gives the most pronounced effect, stimulating transcription approximately 10-fold. The level obtained with this construct is identical to that observed with footprints 1–6 (Fig. 6C), indicating that the NF-E2 dimer site can functionally replace footprints 1–2 in the MEL assay.

However, when the NF-E2 dimer site is tested in the context of footprints 1–6, we do not find such a synergistic increase. We therefore conclude that NF-E2 does
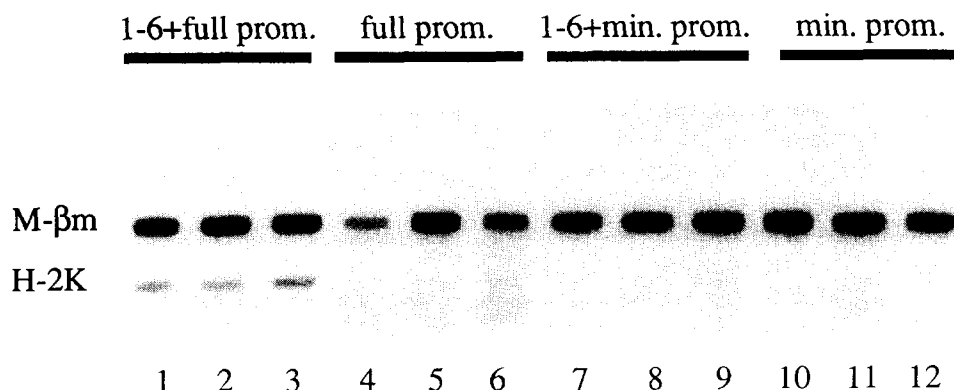

Fig. 5. Promoter requirements of HS3. The promoter region up to −815 of the human β-globin gene was fused to the mouse H-2K reporter gene, and the core fragment of HS3 was cloned at −815. To obtain the minimal promoter construct, nucleotides −103 to −265 were deleted from the promoter. These constructs were then tested in stably transfected MEL cells. For S1 analysis, probes specific for the 5' ends of β-H-2K hybrid mRNA (85 nt protected fragment) and mouse β-major (M-βm) mRNA were used. Relative specific activities of the probes: H-2K/M-βm = 10:1.

not play a major part in the function of HS3 in MEL cells.

3.3. The human β-globin gene promoter and transcriptional activation by HS3

The promoter and enhancer of the human β-globin gene contain binding sites for erythroid and ubiquitous transcription factors [21,30]. It is possible that the transcriptional activation observed with the HS3/β-globin constructs in this study is due, at least in part, to these local genetic elements. We therefore investigated the requirements of HS3 for promoter elements by linking footprints 1–6 to the β-globin promoter (up to

−815) and cloning this in front of a non-erythroid H-2K reporter gene [28,30].

We have shown previously that, in the absence of the LCR, elements between −103 and −265 are required for inducibility of the human β-globin promoter in MEL cells [28,30]. However, in the presence of the microlocus LCR [31], promoter elements upstream from −103 appear to be redundant, while the presence of the minimal promoter is sufficient to obtain the full level of expression [29]. The minimal promoter consists of a CACC box at −95, a CCAAT box at −80, and a TATAA box at −30 [30]. We deleted nucleotides −103 to −265 to investigate whether footprints 1–6 can function through the minimal promoter.
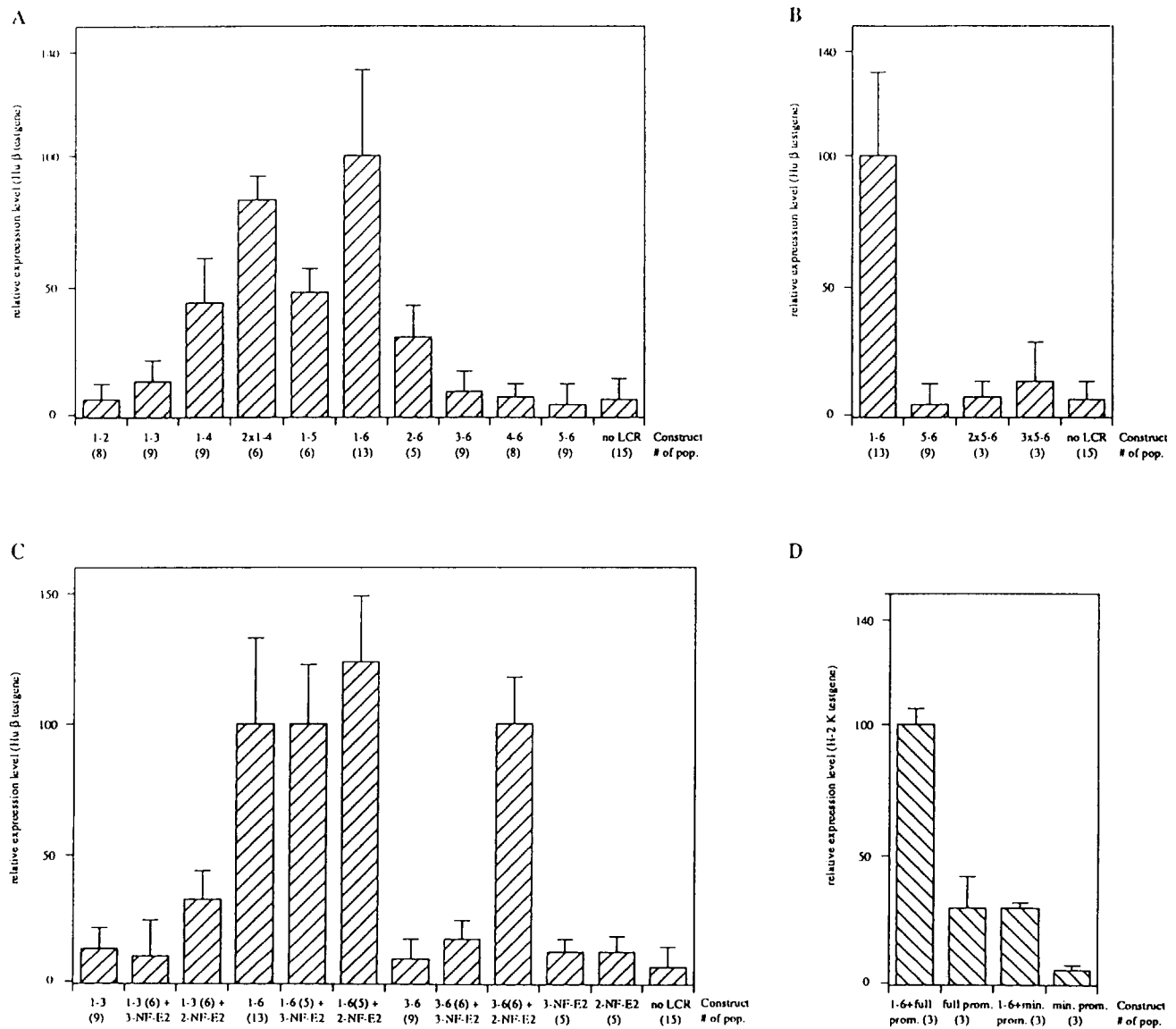

Fig. 6. Quantitative analysis of HS3 constructs in MEL cells. Expression data were quantitated from S1 gels using the phosphorimager. The expression level obtained with the construct containing footprints 1–6 was arbitrarily set at 100%. The number of populations analyzed for each construct is indicated between parentheses. Error bars represent standard deviations.

The results obtained with the constructs containing footprints 1–6 show that HS3 is capable of activating the full β-globin promoter when attached to the H-2K gene (Figs. 5 and 6D). This result is in agreement with that reported for the microlocus LCR [29]. In this earlier study, no difference was found between constructs containing the full or the minimal promoter of the β-globin gene. A 2-fold reduction in the expression level of the minimal promoter construct was observed after deletion of HS3 from the microlocus, leaving HS1, 2 and 4. Interestingly, we observe a more severe reduction when we test the HS3 core fragment with the the minimal promoter/H-2K construct. The expression level is five times lower than that obtained with the full promoter, demonstrating that in MEL cells HS3 interacts with protein binding sites in the upstream promoter from −103 to −265. Most notably, the upstream promoter of the β-globin gene contains a strong GATA-1 binding site at −200 [30]. It is important to recognise however that HS3 can activate both the full and the minimal promoter when attached to the non-erythroid H-2K gene.

4. Discussion

In this paper we have investigated some of the parameters governing the activity of HS3 of the human β-globin LCR. We have analyzed the core fragment by progressive removal of footprinted areas from either the 5' or the 3' end. The results show that the 5' half of the core fragment is the most important for expression. The removal of footprint 1, a GATA-1 site, results in the loss of 70% of the activity of HS3. Subsequent removal of footprint 2 abolishes expression to the level observed with the gene alone. Deletions from the 3' end of footprints 6 or 5 and 6 results in a reduction of expression to 47 and 40%, respectively, of that obtained with the full site. An important conclusion from these data is that both GATA-1 sites (footprints 1, 3 and 5) and the G-rich motifs (footprints 2, 4 and 6) have a functional role, since we observe an effect on expression if we remove either type of footprint. Moreover, in a parallel study in transgenic mice we have shown by point mutagenesis that the GATA-1 sites in footprints 1 and 3 and the G-rich motif in footprint 2 are absolutely required for the LCR activity of HS3 [32]. These data are consistent with the results we report here. Interestingly, in vivo footprinting data have been recently published for HS3 [33]. Orkin and co-workers [32] have demonstrated that protein-DNA interactions occur in vivo at footprints 1 to 5, in agreement with our results. The only discrepancy between this report and the in vivo footprinting data occurs at footprint 6. Footprint 6 is not footprinted in vivo, but its deletion leads to a reduced expression level in our

assay (Figs. 3A and 6A). This discrepancy may be explained by the fact that footprint 6 is a relatively weak protein binding site which as a result is occupied too infrequently to be detected with the in vivo footprinting technique.

From our previous work the interesting question arose whether HS3 simply depends on a certain number of footprints containing the GATA-and the G-rich motif, or whether in addition these need to be in a particular arrangement. The minimum number of footprints required to obtain activity in the MEL transfection system used in this study is four, as demonstrated by the construct containing footprints 1–4. This construct consists of GATA-1 binding sites in footprints 1 and 3 and G-rich sequences in footprints 2 and 4. Duplication of this fragment results in an additive increase in the level of expression. However, when we tried to mimic this situation by testing 1, 2 and 3 copies of footprints 5 and 6 (a GATA-1 site and a G-rich motif), we did not restore expression to the levels observed with either footprints 1–4 or 1–6. The expression level remained indistinguishable from that obtained with the gene alone. In a separate attempt we combined a synthetic GATA-1 site with a consensus for the G-rich motif and inserted 1 to 6 copies of this module into the β-globin test plasmid. Again, we failed to obtain expression levels above background (data not shown). In addition, we cloned oligonucleotides covering footprints 1, 2, 3 and 4 into the test plasmid. This construct now contains all the sequence information present in the construct containing footprints 1–4, but the spacing between the footprints has changed. This construct was also inactive. Although these are three lines of negative evidence, they support our notion that the protein binding sites in HS3 have to be in a particular spatial arrangement in order to create a functional entity. Activity is therefore achieved through synergistic interactions between the individual elements only if the appropriate spacing requirements have been fulfilled.

The construct with footprints 1–4 may not be the minimal construct with LCR activity due to limitations of the transfection assay. Since transfected cells have to be selected for resistance to G418, i.e., expression of the neomycin gene, the populations obtained are biased for productive integrations of the constructs. This results in an artificially high expression level of non-LCR containing constructs. Our work in transgenic mice [32] has shown that constructs containing footprints 1–3 or 3–6, which are indistinguishable from the background in the MEL cell system, are the smallest constructs with LCR-like properties.

The only erythroid-specific factor found to interact with the core fragment of HS3 so far is GATA-1. We therefore determined whether another erythroid-specific factor, NF-E2 [24,25], could play a part in

transcriptional activation by HS3. In HS2 of the human β-globin LCR, a dimer binding site for NF-E2 has been shown to be responsible for the high levels of transcriptional activation achieved with this site [10,26,27]. HS3 contains a single consensus binding site for NF-E2 30 bp upstream of the core fragment containing footprints 1–6. Both the HS2 and HS3 NF-E2 binding sites are footprinted in vivo [33–35], but functional data to establish the role of NF-E2 in the activity of HS3 have not, as yet, been reported. It was therefore of interest to see whether the addition of NF-E2 binding sites to HS3 constructs would have any effect in vivo. Addition of the HS3 NF-E2 site did not appear to have any effect on the expression of constructs containing footprints 1–6 and 1–3. This would argue that NF-E2 is not important in the context of HS3, and is consistent with the observation that HS2 and HS3 do not synergize when tested in cis [17,29]. However, addition of the HS2 NF-E2 dimer site to the otherwise inactive construct containing footprints 3–6 does result in a synergistic enhancement of transcription. We conclude from this that NF-E2 can substitute at least partly for the functions retained in footprints 1 and 2, and that its activity may be obscured in the presence of these footprints in the MEL system. Interestingly, when we originally mapped the core fragment of HS3, we noticed that although it gave the full level of activity in transfected MEL cells, the expression level in transgenic mice was reduced to about 60% of that obtained with a larger fragment, which includes the NF-E2 site. We suggested that this reduction might be caused by the loss of this NF-E2 motif, and the results presented here support this view. NF-E2 is not as vital for the functioning of HS3 as it is for HS2. Addition of the HS2 dimer site to the HS3 fragment containing footprints 1–6 results in a moderate increase of expression (Figs. 4 and 6C). In the case of HS2, deletion of this site results in a dramatic decrease in the expression level [10,14,27]. This highlights a fundamental difference between these two elements of the LCR, which have a totally different architecture [12,14]. It suggests that the stimulation of transcription by NF-E2 is highly dependent on the context of other surrounding binding sites.

Finally, we started investigating the promoter requirements of HS3, by asking whether the minimal β-globin promoter [30] is sufficient to obtain the level of expression obtained with the full promoter, as has been shown to be the case for the microlocus LCR [29]. The data show that HS3 requires the intact promoter to function fully. This paradox may be explained by distance effects. In humans, the LCR is situated 50 kb upstream of the human β-globin gene, which is much more than in our constructs. A large reduction in distance between the LCR and the β-globin promoter may circumvent the requirement for the upstream pro-

moter, since this reduction might facilitate interactions between the LCR and the promoter. Our results indicate that the upstream promoter could be functional in vivo. We have recently been able to generate transgenic mice containing the intact 70 kb human β-globin locus [36]. This has opened up the way to investigate these questions in the closest possible approximation to the in vivo situation.

In addition, it should be noted that the promoter of the TK-neo gene might compete with the globin promoter for HS3. The TK promoter may therefore compete more successfully for HS3 with the minimal promoter than it does with the full promoter, resulting in an increased expression level of the TK-neo gene, at the expense of the transcription of the β-globin gene. This suggests that the upstream promoter is involved in the switch from fetal to adult globin expression, a process in which the relative position of the genes in the gene cluster and their ability to compete for the LCR have been proposed to be key factors [37–39].

As to the mechanism of transcriptional activation by HS3, there may be some common elements between HS3 and the promoter, notably the GATA-1 site at −200 and the CACC box at −95 in the promoter and the modular array of GATA-1 sites and GGTGG (CCACC) motifs in HS3. Since the most widely favoured model [40] for gene activation postulates direct interaction between the activating sequences and the promoter, with the intervening DNA looping out (the 'looping model'), interactions between HS3 and the β-globin promoter may occur through these common motifs. However, the spatial arrangement of these elements appears to be very different between the promoter and HS3, making it difficult to envisage how these interactions are achieved. We are currently designing assays to investigate the validity of the looping model both in vitro and in vivo.

Acknowledgements

References

[1] Collins, F.S. and Weissman, S.M. (1984) Prog. Acid Res. Mol. Biol. 31, 315.
[2] Driscoll, M.C., Dobkin, C.S. and Alter, B.P. (1989) Proc. Natl. Acad. Sci. USA 86, 7470.
[3] Kioussis, D., Vanin, E., deLange, T., Flavell, R.A. and Grosveld, F.G. (1983) Nature 306, 662.

[4] Grosveld, F., Blom van Assendelft, G., Greaves, D. and Kollias, G. (1987) Cell 51, 975.

[5] Chada, K., Magram, J., Raphael, K., Radice, G., Lacy, E. and Costantini, F. (1985) Nature 314, 377.

[6] Kollias, G., Wrighton, N., Hurst, J. and Grosveld, F. (1986) Cell 46, 89.

[7] Perez-Stable, C. and Costantini, F. (1990) Mol. Cell. Biol. 10, 1116.

[8] Townes, T.M., Lingrel, J.B., Chen, H.Y., Brinster, R.L. and Palmiter, R.D. (1985) EMBO J, 4, 1715.

[9] Tuan, D., Solomon, W., Li, Q. and London, I.M. (1985) Proc. Natl. Acad. Sci. USA 82, 6384.

[10] Caterina, J., Ryan, K., Pawlik, R., Palmiter, R., Brinster, R., Behringer, R. and Townes, T. (1991) Proc. Natl. Acad. Sci. USA 88, 1626.

[11] Lowrey, H., Bodine, D. and Nienhuis, A. (1992) Proc. Natl. Acad. Sci. USA 89, 1143.

[12] Philipsen, S., Talbot, D., Fraser, P. and Grosveld, F. (1990) EMBO J. 9, 2159.

[13] Pruzina, S., Hanscombe, O., Whyatt, D., Grosveld, F. and Philipsen, S. (1991) Nucleic Acids Res. 19, 1413.

[14] Talbot, D., Philipsen, S., Fraser, P. and Grosveld, F. (1990) EMBO J. 9, 2169.

[15] Fraser, P., Hurst, J., Collis, P. and Grosveld, F. (1990) Nucleic Acids Res. 18, 3503.

[16] Fraser, P., Pruzina, S., Antoniou, M. and Grosveld, F. (1993) Genes Dev. 7, 106.

[17] Collis, P., Antoniou, M. and Grosveld, F. (1990) EMBO J. 9, 233.

[18] Hug, B., Moon, A. and Ley, T. (1992) Nucleic Acid Res. 20, 5771.

[19] Tuan, D.Y., Solomon, W.B., London, I.M. and Lee, D.P. (1989) Proc. Natl. Acad. Sci. USA 86, 2554–2558.

[20] Tsai, S., Martin, D.I.K., Zon, L.I., D'Andrea, A.D., Wong, G. and Orkin, S.H. (1989) Nature 339, 446.

[21] Wall, L., deBoer, E. and Grosveld, F. (1988) Genes Dev. 2, 1089.

[22] Dynan, W. and Tjian, R. (1983) Cell 35, 79.

[23] Xiao J.-H., Davidson, I., Macchi, M., Rosales, R., Vigneron, M., Staub, A. and Chambon, P. (1987) Genes Dev. 1, 794.

[24] Mignotte, V., Wall, L., deBoer, E., Grosveld, F. and Romeo, P.H. (1989) Nucleic Acids Res. 17, 37.

[25] Andrew, N.C., Erdjument-Bromage, H., Davidson, M.B., Tempst, P. and Orkin, S.H. (1993) Nature 362, 722.

[26] Sorrentino, B., Ney, P., Bodine, D. and Nienhius, A.W. (1990) Nucleic Acids Res. 18, 2721.

[27] Talbot, D. and Grosveld, F. (1991) EMBO J. 10, 1391.

[28] Antoniou, M., deBoer, E., Habets, G. and Grosveld, F. (1988) EMBO J. 7, 377.

[29] Antoniou, M. and Grosveld, F. (1990) Genes Dev. 4, 1007.

[30] DeBoer, E., Antoniou, M., Mignotte, V., Wall, L. and Grosveld, F. (1988) EMBO J. 7, 4203.

[31] Talbot, D., Collis, P., Antoniou, M., Vidal, M., Grosveld, F. and Greaves, D.R. (1989) Nature 338, 352.

[32] Philipsen, S., Pruzina, S. and Grosveld, F. (1993) EMBO J. 12, 1077.

[33] Strauss, E. and Orkin, S. (1992) Proc. Natl. Acad. Sci. USA 89, 5809.

[34] Ikuta, T. and Kan, Y.W. (1991) Proc. Natl. Acad. Sci. USA 88, 10188.

[35] Reddy, P.M.S. and Shen, C.-K.J. (1991) Proc. Natl. Acad. Sci. USA 88, 8676.

[36] Strouboulis, J., Dillon, N. and Grosveld, F.G. (1992) Genes Dev. 6, 1857.

[37] Dillon, N. and Grosveld, F. (1991) Nature 350, 252.

[38] Enver, T., Raich, N., Ebens, A.J., Papayannopoulou, T., Costantini, F. and Stamatoyannopoulos, G. (1990) Nature 344, 309.

[39] Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N. and Grosveld, F. (1991) Genes Dev. 5, 1387.

[40] Ptashne, M. (1986) Nature 322, 697.