The EMBO Journal vol.15 no.21 pp.5888–5896, 1996

Erythroid Krüppel-like factor (EKLF) contains a multifunctional transcriptional activation domain important for inter- and intramolecular interactions

Xiaoyong Chen[1] and James J.Bieker[1,2,3]

[1]Brookdale Center for Molecular Biology and [2]Department of Biochemistry, Mount Sinai School of Medicine, One Gustave L.Levy Place, New York, NY 10029, USA

[3]Corresponding author

Erythroid Krüppel-like factor (EKLF) is a red cell-restricted transcriptional activator that plays a dominant role in establishing high levels of β-globin gene expression during erythroid ontogeny. Although its DNA binding domain belongs to the well-studied class of Krüppel-like zinc fingers, its proline-rich activation region has not been thoroughly examined. We have analyzed this region by monitoring the functional effects of its mutagenesis upon EKLF activity *in vivo* and *in vitro*. First, using co-transfection assays, we find that the transactivation region contains discrete stimulatory and inhibitory subdomains. Second, *in vitro* binding assays indicate that the inhibitory domain exerts its effect *in cis* by interfering with DNA binding. Third, *in vivo* competition assays demonstrate that EKLF interacts with a positive-acting cellular factor, and that the domain responsible for this *trans* interaction lies within a 40 amino acid sequence that is coincident with the EKLF minimal transactivation domain. Finally, site-directed mutagenesis of this domain implies that conformation and/or phosphorylation status of its central core may be critical for such interactions. These results point towards post-translational steric and/or allosteric control of EKLF function that may be important not just for its DNA binding ability, but also for its potential to interact with other proteins that fully establish the correct stereospecific array leading to efficient switching of β-globin transcription during development.

Keywords: EKLF/erythropoiesis/β-globin synthesis/proline-rich domain/transcription factor

Introduction

The combinatorial interplay between cell type-specific proteins bound at proximal and distal promoter elements provide a suitable model for partly explaining tissue-specific regulation of gene expression (Struhl, 1991). These interactions are thought to control formation of the stereospecific array that leads to maximal levels of transcription from its linked gene(s) (Tjian and Maniatis, 1994). Such a model of transcriptional activation becomes important for conceptualizing what may be occurring within the β-like globin gene cluster during the developmental regulation of its individual members (Townes and Behringer, 1990; Felsenfeld, 1992; Engel, 1993; Bungert *et al.*, 1995; Wijgerde *et al.*, 1995). This cluster encompasses ~70 kb that includes the locus control region (LCR) and β-globin gene variants (reviewed in Orkin, 1990). A small group of erythroid-restricted transcriptional activators that bind to a limited number of recurring sites within the LCR and the individual globin promoters may be involved in formation of the appropriate stereostructure that leads to developmentally controlled expression and switching of globin expression (Felsenfeld, 1992; Engel, 1993; Higgs and Wood, 1993; Wijgerde *et al.*, 1995).

One of these cell-specific components is Erythroid Krüppel-like Factor (EKLF) (Miller and Bieker, 1993). EKLF binds via its three zinc fingers to the CACCC element within the adult β-globin gene (Miller and Bieker, 1993; Feng *et al.*, 1994). Transfection of K562 cells (Donze *et al.*, 1995) and murine genetic ablation studies (Nuez *et al.*, 1995; Perkins *et al.*, 1995) indicate that it is responsible for activating high levels of the β-globin transcript. In addition, it contains a proline-rich transactivation domain that is specifically required for LCR-driven promoter- and red cell-specific activation of the β-globin gene (Bieker and Southwood, 1995). The requirement of a correct promoter context for demonstrating this specificity of action and effect has led us to question what the unique determinants are within the transactivation module that impart these particular properties to EKLF.

Prolines are considered 'helix-breakers' that lead to non-globular domains of unpredicted structure (Williamson, 1994). It is thought that such structures may provide a surface for reversible interactions with other proteins. This has clearly turned out to be the case. For example, the SH3 domain preferentially binds proline-rich consensus sequences (Yu *et al.*, 1994). Although EKLF has been denoted a 'proline-rich' transcription factor, it bears no sequence identity to other transcription factors of that genre (Mitchell and Tjian, 1989). However, proline-rich regions within other transcriptional activators have been shown to associate with the TFIIB component of the basal transcription machinery (Kim and Roeder, 1994), and with a specific subset of TBP-associated factors (TAFs) (Tanese *et al.*, 1991; Chiang and Roeder, 1995). The crucial role of EKLF in generating high levels of adult β-globin expression at the appropriate juncture in development make it of interest to undertake a molecular analysis of its proline-rich activation domain.

It is with this background in mind that we undertook a mutational analysis of the EKLF transactivation region using a number of approaches. First, we used transfection assays to localize the minimal domains of the region and analyze their effects on transactivation. Second, we used *in vitro* binding assays to ascertain the *cis*-acting effects imparted by one of these domains. Finally, we used *in vivo* competition assays to determine whether EKLF interacts *in trans* with other transcriptional proteins and to localize this effect. Our results demonstrate that the EKLF trans-

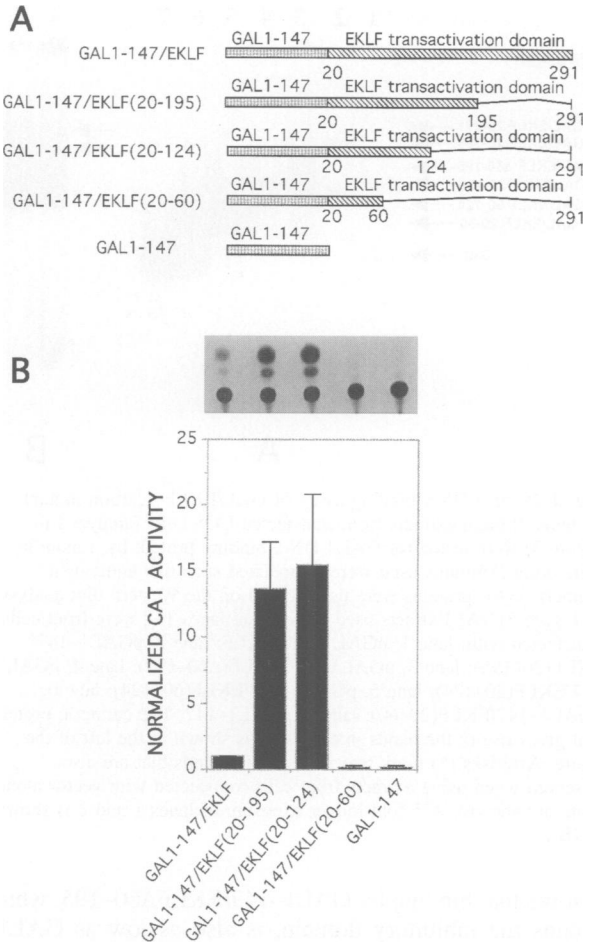

Fig. 1. Deletional analysis of the EKLF transactivation region. Transient co-transfection assays of 32DEpo1 cells were performed with pG5BCAT reporter and the indicated expression vectors (**A**). The normalized results from multiple experiments are shown in (**B**), along with an autoradiograph of the thin layer plate from one experiment.

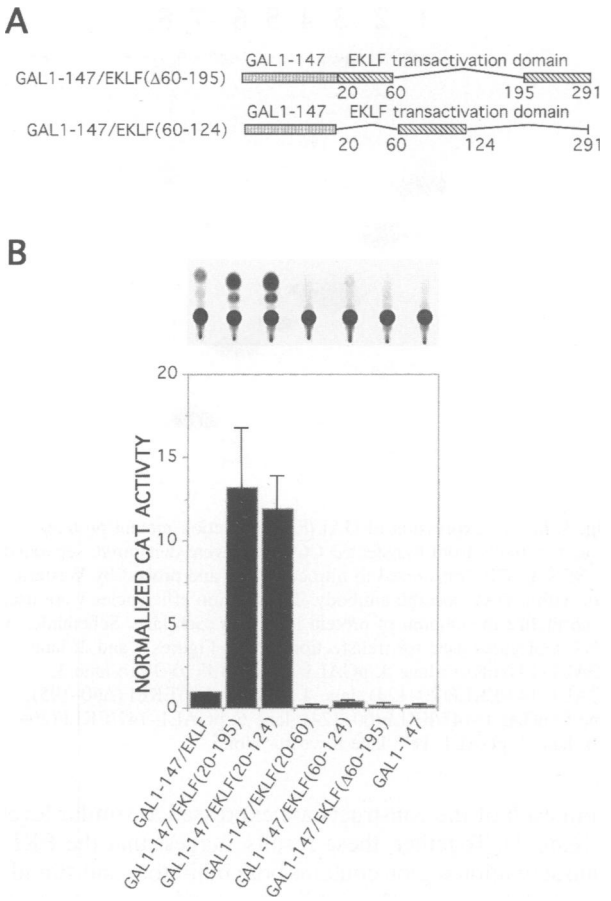

Fig. 2. Effects of internal deletions upon EKLF transactivation. (**A**) Schematic of constructs used in addition to those in Figure 1 for transient assays. The normalized results from multiple experiments are shown in (**B**), along with an autoradiograph of the thin layer plate from one experiment.

activation region is a mixture of domains that are distinguished by their ability to exert intramolecular effects and intermolecular associations.

Results

The EKLF transactivation region contains stimulatory and inhibitory subdomains

Our previous study demonstrated that the proline-rich region of EKLF (amino acids 20–291) can act as a transcriptional activator in both non-erythroid and erythroid cells (Bieker and Southwood, 1995). These studies were facilitated by the modular property of this region, such that linking it to the GAL1–147 DNA binding domain yielded a chimeric factor that activated transcription via binding to GAL4 DNA binding sites. Starting with this construct, we generated a series of deletion mutants to delimit the region responsible for the transcriptional activation. The initial deletions were constructed by serially removing portions of the EKLF transactivation domain from the C-terminal end (Figure 1A). The transcriptional activity of these chimeric proteins was examined by co-transfection with the pG5BCAT reporter plasmid (Lillie and Green, 1989) into mouse bone marrow 32DEpo1 cells (Migliaccio *et al.*, 1989). This transfectable cell line retains erythroid properties as it expresses GATA1 and induces β-globin after treatment with erythropoietin (Migliaccio *et al.*, 1989; Kreider *et al.*, 1993).

The results with GAL1–147/EKLF(20–195) (Figure 1B) immediately revealed the presence of an inhibitory domain located at amino acids 195–291, since removal of this region boosted transactivation 14-fold when compared with the full-length construct (after normalization to an internal control). Removal of an additional 71 amino acids [to residue 124; GAL1–147/EKLF(20–124)] had no effect, but removal to residue 60 [GAL1–147/EKLF(20–60)] completely abolished transactivation. The importance of amino acids 60–124 for transactivation was verified by the fact that an internal deletion that removed this portion is also deficient in transactivation [Figure 2B; GAL1–147/EKLF(Δ60–195)]. However, the sequence requirement for activation is complex, as a construct that solely comprises amino acids 60–124 only has weak (but detectable) activity [Figure 2B; GAL1–147/EKLF(60–124)]. This residual activity retains ~25% that of the full-length construct (GAL1–147/EKLF). To exclude the possibility that the results could be explained by differing expression of the mutated proteins, COS7 cells were transfected with each of the constructs, and protein extracts were prepared and analyzed by Western blot analysis with an anti-GAL1–147 antibody. The analysis shows that the chimeric proteins

Fig. 3. *In vivo* expression of GAL/EKLF deletion mutant proteins. Protein extracts from transfected COS cells were denatured, separated by SDS–PAGE, transferred to nitrocellulose, and probed by Western blot with a GAL-specific antibody. Transfection efficiencies were used to normalize the amount of protein loaded in each lane. Schematics of DNA molecules used for transfections are in Figures 1 and 2: lane 1, pGAL1–147/EKLF; lane 2, pGAL1–147/EKLF(20–195); lane 3, pGAL1–147/EKLF(20–124); lane 4, pGAL1–147/EKLF(Δ60–195); lane 5, pGAL1–147/EKLF(60–124); lane 6, pGAL1–147/EKLF(20–60); lane 7, pGAL1–147; lane 8, vector alone.

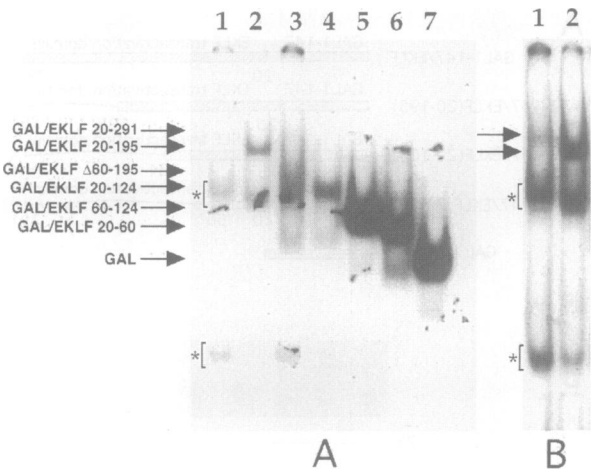

Fig. 4. *In vitro* DNA binding assay of GAL/EKLF deletion mutant proteins. Protein extracts from transfected COS cells (analyzed in Figure 3) were tested for GAL4 DNA binding protein by a mobility shift assay. Amounts used were normalized such that equivalent chimeric GAL proteins were used (based on the Western blot analysis of Figure 3). (**A**) Extracts used to generate lanes 1–7 were from cells transfected with: lane 1, pGAL1–147/EKLF; lane 2, pGAL1–147/EKLF(20–195); lane 3, pGAL1–147/EKLF(Δ60–195); lane 4, pGAL1–147/EKLF(20–124); lane 5, pGAL1–147/EKLF(60–124); lane 6, pGAL1–147/EKLF(20–60); lane 7, pGAL1–147. The chimeric protein that gives rise to the bands in each lane is shown to the left of the figure. Asterisks (*) mark two non-specific bands that are also observed when using extracts from cells transfected with vector alone (data not shown). A 5-fold longer exposure of lanes 1 and 2 is shown in (**B**).

from each of the constructs are expressed at similar levels (Figure 3). Together, these results suggest that the EKLF transactivation region contains both inhibitory and stimulatory subdomains. The inhibitory subdomain is at the carboxyl end and encompasses amino acids 195–291. On the other hand, the stimulatory subdomain is located at the amino-terminus. Specifically, the segment that contains amino acids 60–124, although able to generate detectable transactivation by itself, requires the additional contribution of the adjacent amino acids 20–60 for maximal transcriptional activity that is 35-fold higher. As a result, we feel that EKLF amino acids 20–124 define the minimal activation subdomain.

The EKLF inhibitory subdomain affects DNA binding ability in vitro

We next addressed how these subdomains were performing their function. Specifically, we considered whether the activation and inhibitory domains act by inter- or intramolecular means. Two approaches were undertaken to address this issue. One mechanism by which the inhibitory domain might be exerting its function would be by interfering, *in cis*, with efficient DNA binding by the full-length protein. We used gel mobility shift assays to determine the relative binding by each of the mutated GAL/EKLF proteins to a GAL4 DNA binding sequence. The protein extracts from the transfected COS7 cells described above (Figure 3) contain an equivalent amount of each chimeric protein. However, DNA binding assays using these same extracts demonstrate that, although all proteins bind the DNA target site, they do not do so equally well (Figure 4). In fact, under the conditions of this assay, removal of the inhibitory domain substantially increases the DNA binding ability [compare GAL1–147/EKLF with GAL1–147/EKLF(20–195)]. Although not obvious in the figure, use of different gel conditions resolve that binding by GAL1–147/EKLFΔ60–195, which retains the inhibitory domain, is also as low as GAL1–147/EKLF (data not shown). Binding is further increased by additional removal of EKLF sequences, such that the GAL1–147/EKLF(20–60) and GAL1–147/EKLF(60–124) constructs bind as well as GAL1–147 alone. We draw two conclusions from these data. First, the low level of *in vivo* transactivation by GAL1–147/EKLF(60–124) and non-detectable CAT activity generated by GAL1–147/EKLF(20–60) are not due to their low expression or low *in vitro* DNA binding, but are a function of the suboptimal transactivation domain configuration. Second, the increase in transactivation by GAL1–147/EKLF(20–195) relative to that of the full-length chimera can be accounted for by the increase in DNA binding that results from removal of amino acids 196–291. This is consistent with the idea that amino acids 196–291 of EKLF must be exerting a *cis*, or intramolecular, inhibitory effect upon the DNA binding and transactivational activity of the full-length protein.

Establishment of an in vivo competition assay

The second approach was directed at an alternate (although non-exclusive) scenario whereby the EKLF activation domain exerts its effects *in trans* by interacting with a cellular factor. To address this issue we used an *in vivo* competition assay (Gill and Ptashne, 1988; Meyer *et al.*, 1989; Colgan *et al.*, 1993), based on the idea that co-transfection with increasing amounts of full-length EKLF will sequester limiting proteins with which it interacts by protein–protein interactions. Because native EKLF does not bind to the reporter (pG5BCAT), any effects observed upon its normal transactivation by the chimeric GAL1–147/EKLF activator will necessarily result from native

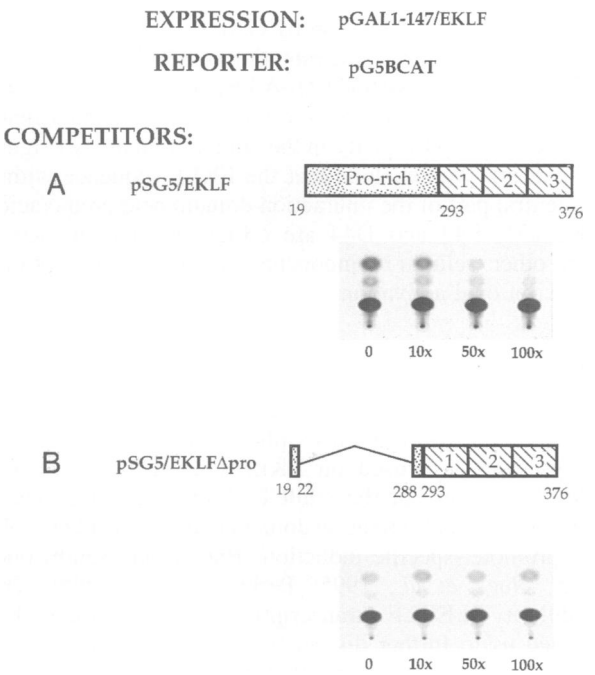

Fig. 5. Establishment of an *in vivo* competition assay. Transfection of 32DEpo1 cells was performed with a constant amount of pGAL1–147/EKLF and pG5BCAT, and with increasing amounts of pSG5/EKLF (**A**) or pSG5/EKLFΔpro (**B**) as indicated. A schematic diagram of the protein product from each of these constructs is shown. Total DNA concentration was kept constant by using pSG5 vector DNA as 'filler'. The autoradiograph of the thin layer plate from one typical CAT assay is shown for each experiment.

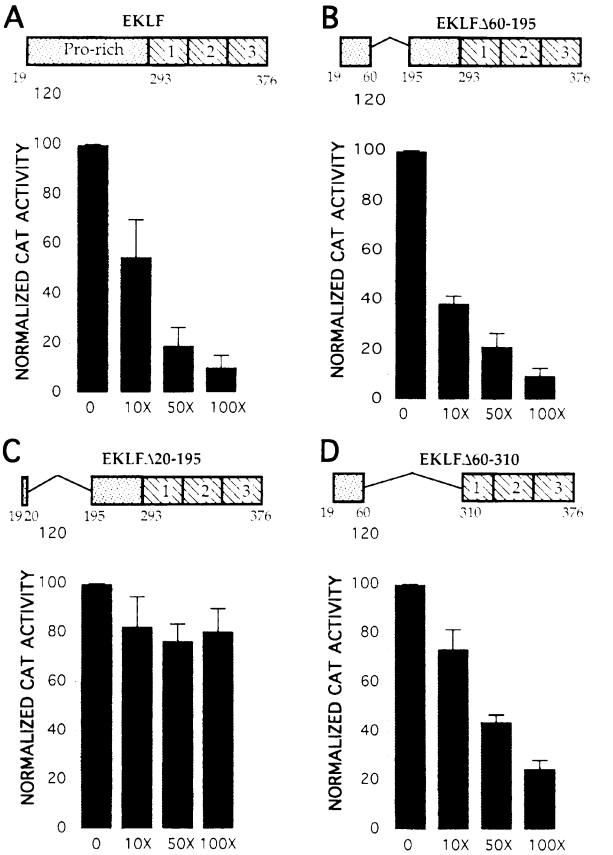

Fig. 6. Localization of the EKLF region responsible for *in vivo* competition. *In vivo* competition was performed as in Figure 5, except the total DNA concentration was kept constant by using pSG5/EKLFΔpro DNA as filler. Normalized data from multiple experiments are shown using pSG5/EKLF (**A**), pSG5/EKLFΔ60–195 (**B**), pSG5/EKLFΔ20–195 (**C**) or pSG5/EKLFΔ60–310 (**D**) as competitor DNA. Each of these constructs is capable of yielding equivalent amounts of protein (data not shown).

EKLF's depletion of a transcriptionally required cellular factor. 32DEpo1 cells were co-transfected with activating construct GAL1–147/EKLF, reporter pG5BCAT, and increasing amounts of competitor pSG5/EKLF across a 100-fold range. The total amount of DNA was kept constant by the addition of pSG5. Results of such an experiment (Figure 5A) indicate that the CAT activity decreases in proportion to the amount of full-length EKLF that was included. This suggests that full-length EKLF is competing for a positive-acting factor that is required for efficient transactivation by GAL1–147/EKLF at the pG5BCAT promoter. It is also consistent with the idea from our previous data that the inhibitory domain is operating *in cis*, rather than interacting with a negative-acting factor.

We next repeated this experiment using increasing amounts of pSG5/EKLFΔpro (Bieker and Southwood, 1995), which encodes a protein that only contains the zinc fingers. As shown in Figure 5B, this construct was not able to compete with GAL1–147/EKLF for activation of pG5BCAT. The results in Figure 5A and B immediately localized the competitive effect to the EKLF transactivation domain. As a result we felt justified in using this assay to address the intermolecular issue for EKLF transactivation that was raised above.

A cellular factor is involved in the EKLF transcriptional activating mechanism

Although the zinc finger region had no effect on the competition assay, we took the added precaution of using pSG5/EKLFΔpro (rather than pSG5) as filler DNA for our deletion studies. This effectively normalizes the total amount of zinc finger region to a constant level within each experiment.

Using these modified conditions, competition with full-length EKLF results in a >90% decrease in transcriptional activation by GAL1–147/EKLF as before (Figure 6A). This remains true when a portion of the minimal transactivation domain is deleted (EKLFΔ60–195; Figure 6B). However, when an additional deletion that results in complete removal of the minimal activation domain is tested, competitive ability is abolished (EKLFΔ20–195; Figure 6C). This pointed to EKLF amino acids 20–60 as being responsible for competition. Their importance was further verified by testing an additional construct that contains only this portion of the transactivation region (EKLFΔ60–310). In this case, competitive ability is retained to nearly the same level as that of the full-length protein (Figure 6D). Each of these constructs is capable of expressing an equivalent level of protein (data not shown; see Materials and methods). Together, these data show that amino acids 20–60 define an interaction domain that is critical for competition, and suggest that it disrupts GAL1–147/EKLF activation by sequestering a positive activator. As this sequence is part of the minimal activation domain, these data are consistent with the idea that the mechanism of this activation is *in trans*, unlike that of the

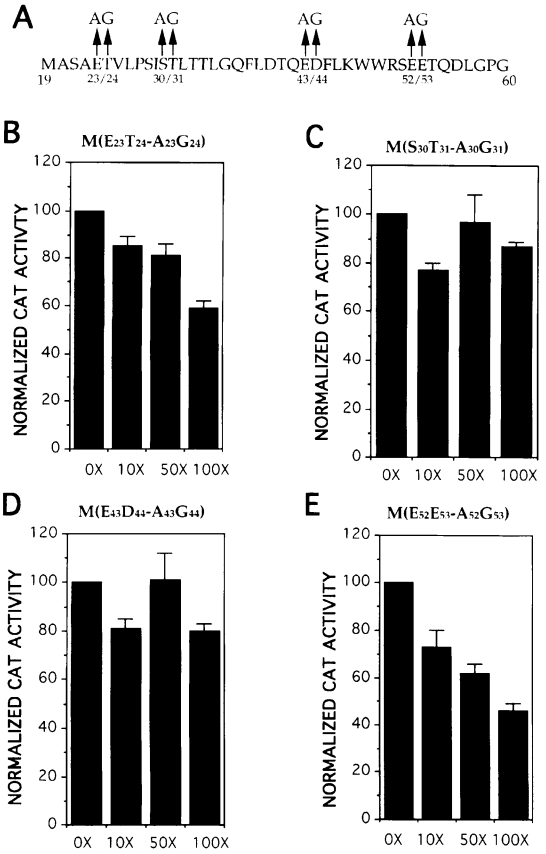

Fig. 7. Effects of site-directed mutagenesis of the EKLF interaction domain upon *in vivo* competition. (**A**) The sequence of the first 41 amino acids of EKLF are shown along with the location of the mutated pairs of amino acids for the pSG5/EKLFΔ60–195-derived constructs used in (B–E). Normalized data from multiple experiments are shown using pSG5/EKLFΔ60–195/M(E23T24–A23G24) (**B**), pSG5/EKLF(Δ60–195)/M(S30T31–A30G31) (**C**), pSG5/EKLF(Δ60–195)/M(E43D44–A43G44) (**D**) or pSG5/EKLF(Δ60–195)/M(E52E53 A52G53) (**E**) as competitor DNA. Each of these constructs is capable of yielding equivalent amounts of protein (data not shown).

inhibitory domain, which is not involved in the competitive interaction between EKLF and a cellular factor.

Directed mutagenesis of the interaction domain

In order to further define the amino acid residues that are important for EKLF's intermolecular interaction, the interaction domain segment that encompassed amino acids 20–60 was analyzed by site-directed mutagenesis. Inspection of its sequence indicates that it is rich in acidic amino acids, and that it contains 10 serine and threonine residues. Acidic domains are commonly found within transcriptional activators, and serine/threonine phosphorylation is also known to regulate transcriptional activity. Four pairs of amino acids were mutated on construct pSG5/EKLF(Δ60–195): Glu23/Thr24 (E23T24), Ser30/Thr31 (S30T31), Glu43/Asp44 (E43D44) and Glu52/Glu53 (E53E53), were all changed to Ala/Gly (Figure 7A). All four mutant proteins can be expressed to a similar level as the parental construct (data not shown; see Materials and methods). These four mutant constructs were then used as competitors in the *in vivo* competition assay. The results show that mutation of E23T24 or E52E53 (EKLF(Δ60–195)/M(E23T24–A23G24) or EKLF(Δ60–195)/M(E52E53–A52G53) results in proteins that still have a significant ability to depress activation by GAL1–147/EKLF (Figure 7B and E). In contrast, the mutation of S30T31 or E43D44 [EKLF(Δ60–195)/M(S30T31–A30G31) or EKLF(Δ60–195)/M(E43D44–A43G44)] results in a protein that cannot decrease the CAT activity in the competition assay (Figure 7C and D). We conclude that the EKLF sequence within the central part of the interaction domain near amino acids S30, T31, E43 and D44 are critical for its interaction with other cellular components that results in optimal transcriptional activation.

Discussion

EKLF is a transcriptional activator that is critical for erythroid cell-specific transcription. Most of its functional analyses have focused on EKLF's interaction with the CACCC element at the adult β-globin promoter, where its transcriptional activation domain plays a key role in cell- and promoter-specific induction (Bieker and Southwood, 1995; Nuez *et al.*, 1995; Perkins *et al.*, 1995). The modularity of EKLF's transcription activation domain has enabled us to further dissect this 270 amino acid region into separate subdomains that illuminate their distinct functional properties and raise important implications about EKLF's mechanism of action.

EKLF contains a multipartite transcriptional activation region

Our results (summarized in Figure 8) suggest that the EKLF proline-rich region contains functionally distinct inhibitory (aa 195–291) and activation (aa 20–124) domains. The aa 20–124 minimal activation domain can be further divided into a subregion (aa 20–60) that itself does not activate, but interacts via intermolecular interactions with another cellular protein and confers optimal transactivation potential to its adjacent (aa 60–124) sequence. At the same time, the inhibitory domain (aa 195–291) operates intramolecularly; that is, it appears to prevent efficient binding by the DNA binding portion within the same molecule. The increase in *in vitro* DNA binding observed upon its removal from the full-length protein is reflected in its increased transactivation potential *in vivo*. Multipartite transcriptional activation domains have been observed within a number of transcription factors, including c-myb (Luscher and Eisenman, 1990), p53 (Hupp *et al.*, 1992), NGFI-A (Russo *et al.*, 1993) and C/EBPα (Nerlov and Ziff, 1994).

Mechanistic implications of a cis-acting EKLF inhibitory domain

The inhibitory domain may function by retaining EKLF in a state of lower affinity for its target, possibly by sterically or allosterically masking the DNA binding region—as occurs with the Ets-1 transcription factor (Petersen *et al.*, 1995). Upon activation, EKLF may undergo a conformational change such that the DNA binding region becomes readily available for binding to its cognate site, leading to an increase in transcriptional activity. Of course, this leaves open the question of the trigger for such a change. Because we saw no evidence for the inhibitory domain's ability to interact with another cellular factor, activation is not likely to occur in the same fashion as NGFI-A (Russo *et al.*, 1993). In that case,

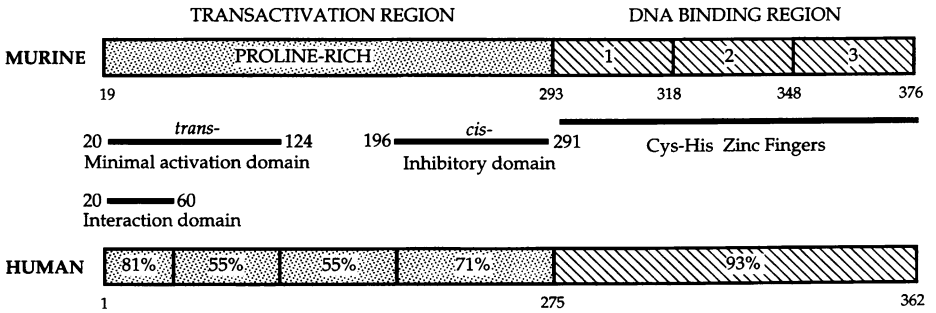

Fig. 8. Summary of EKLF subdomains and comparison with its human homologue. Schematic of the EKLF protein is shown along with demarcation of its various subdomains based on present and earlier studies. The amino acid sequence similarity of each subdomain is separately compared with the human EKLF sequence (Bieker, 1996). Additional comparisons within each domain (pI, helicity, protein kinase motifs) are described in the text.

in vivo competition experiments suggested that a titratable repressor normally prevents NGFI-A function, and that its removal triggers activation. In the present case, our data suggest that the control of EKLF activity more likely involves a *cis*-acting modification that changes its conformation and directly affects its ability to bind DNA independently of other proteins. For example, this phenomenon implies that a modification such as phosphorylation may be involved in opening up the protein, much as is the case with p53. p53 contains a carboxy-terminal regulatory domain whose removal or modification by casein kinase II, protein kinase C or cyclin-dependent kinases activates p53 DNA binding (Hupp *et al.*, 1992; Hupp and Lane, 1994; Wang and Prives, 1995). There are a number of potential phosphorylation sites within this region of EKLF. Alternatively, more subtle changes, such as redox potential (Hainaut and Milner, 1993), may exert an effect.

A striking observation is the inhibitory domain's ability to interfere with DNA binding even when fused to a non-related DNA binding module. However, it is uncertain whether the inhibitory domain itself is directly masking the DNA binding region, or whether it is simply a conduit that brings another part of the molecule in proximity to the DNA binding region. In any case, it is unlikely that the inhibitory domain affects the conformation of the adjacent EKLF zinc fingers, as these folded structures are inherently quite stable (Frankel *et al.*, 1987).

Biological implications of EKLF protein–protein interactions

At the opposite end of the transactivation region, protein interactions are established that are important for the overall critical function of EKLF, which is to transcriptionally activate its target genes. A significant portion of the minimal activation domain provides an interacting surface for association with a cellular factor, likely as part of the transcriptional activation machinery. Protein–protein interactions are known to be important for regulation of transcriptional activity (Tjian and Maniatis, 1994). This directly implies that deciphering the participants in such protein–protein interactions will further elucidate the role of EKLF in red cell-specific activation. In particular, such interactions may have significant relevance to the regulation of globin cluster expression, specifically with respect to adult β-globin switching. Developmental expression studies reveal that EKLF is expressed early during erythroid ontogeny, both at the very beginning of blood island formation in the yolk sac and in the fetal liver primordia (Southwood *et al.*, 1996). Yet *in vitro* and *in vivo* data indicate that EKLF does not interact well with embryonic/fetal globin CAC sites (Bieker, 1994; Donze *et al.*, 1995). In fact, *in vivo* studies have suggested that EKLF is an important factor for the switch from fetal to adult globin (Donze *et al.*, 1995). In addition, genetic ablation of EKLF leads to embryonic death from a profound β-thalassemia at the time of the switch to adult β-globin expression (Nuez *et al.*, 1995; Perkins *et al.*, 1995). As EKLF protein is present in both primitive and definitive erythroid tissues (Southwood *et al.*, 1996), its mere presence must not be sufficient to trigger adult β-globin expression. Although developmentally regulated post-translational modifications may be playing an important role in establishing EKLF function, an additional scenario is that other proteins may need to be expressed or become accessible for full empowerment of the adult β-globin promoter within the adult red cell. Generation of a high-efficiency β-globin promoter may thus involve recruitment by the already-present EKLF protein of other transcription factors to that region of the globin cluster. This line of thinking makes it imperative to isolate the cellular factor(s) with which the EKLF interaction domain is associating, as it may shed light on the mechanism of globin switching during development. If this protein is a component of the LCR, then a critical link to theoretical models of globin gene regulation will have been attained (Townes and Behringer, 1990; Engel, 1993; Wijgerde *et al.*, 1995).

Directed mutagenesis of the interaction domain indicates that its central region must be critical for this function. In addition to changing the acidic nature of this region, mutation of E43D44 disrupts a potential casein kinase II site (T41Q42E43D44). As a result, phosphorylation may be playing a role in EKLF's activation mechanism. Phosphorylation of activation domains are known to potentiate the transcriptional function associated with CREB (Gonzalez and Montminy, 1989), c-Jun (Binetruy *et al.*, 1991; Pulverer *et al.*, 1991), Elk-1 (Marais *et al.*, 1993) and ATF-2 (Livingstone *et al.*, 1995).

There is recent evidence that EKLF can associate with GATA1 to synergize transcriptional activity *in vivo* (Merika and Orkin, 1995; Gregory *et al.*, 1996). These studies monitored activation of natural and artificial GATA/CACCC site-containing reporters that were co-transfected

with GATA1 and EKLF expression vectors into *Drosophila* cells. The *in vitro* association of these two proteins was also demonstrated, and one of these studies (Merika and Orkin, 1995) further indicated that the GATA1 DNA binding region can associate with the EKLF zinc finger region. Although we found no evidence that the EKLF DNA binding region played a role in protein competition *in vivo*, the experimental design of our study was significantly different. In particular, the present experiments were focused on using a mammalian erythropotential cell to dissect the EKLF transactivation region by means of an *in vivo* competition assay. The end-point of this assay was disruption of transcriptional activation *in vivo*. As a result, putative EKLF/GATA1 interactions as suggested by the co-transfection data (Merika and Orkin, 1995; Gregory *et al.*, 1996), rather than by *in vitro* 'pull-down' data, are of most relevance to the present observations. In addition, the *in vivo* competition assay is designed to deplete a limiting factor; it is unlikely that GATA1 is limiting in 32DEpo1 cells (Kreider *et al.*, 1993). As a result, we feel the two sets of data are not in conflict, and in fact may suggest that the EKLF transactivation and DNA binding regions both play a role in generating productive protein–protein complexes.

Structural correlates to functionally important motifs

It is instructive to compare the sequence of these murine subdomains to those in the human EKLF homologue (Figure 8; based on Bieker, 1996). Excluding the DNA binding region, the interaction domain has the highest sequence identity (81%). All the sites that were disrupted are conserved between mice and humans, including the putative casein kinase II site. The inhibitory domain also retains a relatively high level of sequence identity (71%). However, other regions, including the amino acids adjacent to the interaction domain (which together with it constitute the minimal activation domain), are not as highly conserved (55%), although there are extended motifs within this domain that are identical.

Continuing this comparison between homologues further, we find biochemical and predicted secondary structure differences between subdomains that are conserved in both the human and murine sequences. For example, both the interaction domain and its adjacent minimal activation domain are quite acidic, with a pI ~3.7. This is in contrast to the inhibitory domain, which is considerably basic, with a pI ~10.7. The large number of prolines scattered throughout the transactivation region prevent the appearance of β-sheets and α-helices, with the only exception being that of the interaction domain, which is predicted to form two α-helical regions separated by a β-sheet (based on protein modeling algorithms; Chou and Fasman, 1978; Garnier *et al.*, 1978). Acidic domains are known to be strong activators of transcription; in some cases, the precise sequence appears to be unimportant for this effect (Ptashne, 1988). However, it is noteworthy that directed mutagenesis of amino acids within the EKLF interaction domain that do not change its overall acidity (i.e. S30T31) yet shortens its predicted β-sheet is sufficient to disrupt its ability to compete *in vivo*. This implies that acidity, *per se*, is not sufficient for EKLF activity, and that the sequence/secondary structure of aa 20–60 bears functional importance (Leuther *et al.*, 1993). Consistent with this, the other disrupting mutant pair (E43D44) shortens the predicted second α-helix.

These data provide a compelling basis to investigate modulations in EKLF steric/allosteric properties and phosphorylation status, and to determine how these changes influence its association with other proteins and its DNA binding site. Convergence of all these approaches will likely illuminate additional details relevant to the developmental regulation of the globin gene cluster, particularly with respect to a more complete understanding of β-globin gene switching during erythroid ontogeny.

Materials and methods

Plasmid constructions

Constructs pGAL1–147/EKLF, pSG5/EKLF, pSG5/EKLFΔpro have been previously described (Miller and Bieker, 1993; Bieker and Southwood, 1995). All inserts were routinely purified by agarose gel electrophoresis prior to ligation (Geneclean). pGAL1–147/EKLF(20–195) was generated by removing the *Sac*II–*Xba*I fragment from pGAL1–147/EKLF and ligating the filled-in (Klenow DNA polymerase) ends of the shortened parent clone. pGAL1–147/EKLF(20–124) was constructed by *Bam*HI digestion of pGAL1–147/EKLF (to isolate the EKLF transactivation region insert) followed by *Taq*I digestion of the isolated piece and ligation of the shortened EKLF fragment back into pGAL1–147. pGAL1–147/EKLF(20–60) was made by removing the *Apa*I–*Xba*I fragment from pGAL1–147/EKLF. pGAL1–147/EKLF(60–124) was constructed by removing the *Pst*I–*Apa*I fragment from pGAL1–147/EKLF(20–124). pGAL1–147/EKLF(Δ60–195) was made by removing the *Apa*I–*Sac*II fragment from pGAL1–147/EKLF. pSG5/EKLF(Δ60–195) was made by removal of the *Sac*II–*Apa*I fragment from pSG5/EKLF. pSG5/EKLF(10–60) was made by removal of the *Sac*I–*Apa*I fragment from pSG5/EKLF.

Transformer Site-Directed Mutagenesis Kit (Clontech) was used to mutate four pairs of amino acids within pSG5/EKLF(Δ60–195). For the E23T24 to A23G24 mutation, mutation primer 5′-TG GCC TCA GCT GCC GGC GTC TTA CCC TCC-3′ was used. For the S30T31 to A30G31 mutation, mutant primer 5′-TTA CCC TCC ATC GCC GGC CTC ACC ACC CTG -3′ was used. For the E43D44 to A43G44 mutation, mutation primer 5′-TTT CTG GAC ACC CAG GCC GGC TTC CTC AAG TGG TGG -3′ was used. For the E52E53 to A52G53 mutation, mutation primer 5′-GG TGG CGG TCT GCC GGC ACG CAG GAT TTG-3′ was used. All mutation primers contained a *Nae*I site, so all correctly generated mutants attained a newly created *Nae*I restriction site. The selection primer was 5′-GCTTTTTTGGAGTACTA-GGCTTTTGC-3′ in all cases, which changed a *Stu*I site to a *Sca*I site.

The reading frame and sequences of all constructs were confirmed by DNA sequence analysis. Due to the lack of a common epitope that could be recognized by our anti-EKLF antibody (Southwood *et al.*, 1996) in all the deleted/mutated expression constructs, the capabilities of these constructs to generate the correct protein product was assessed by a coupled *in vitro* transcription/translation assay (Promega) as recommended by the manufacturer.

Transfections and CAT assays

Transient transfection of 32DEpo1 cells (Migliaccio *et al.*, 1989) was performed by the DEAE–dextran method (Suzow and Friedman, 1993). 10 μg test DNA, 10 μg reporter DNA and 1 μg growth hormone DNA were used in each transfection. For competition assays, 0.1 μg test DNA, 1 μg reporter DNA and 0×, 10×, 50×, 100× (by mass) competitor DNA were used. Total DNA was kept constant by the addition of pSG5 or pSG5/EKLFΔpro. CAT assay was performed for 2 h using proteins extracted from transfected cells as described before (Miller and Bieker, 1993). The CAT activity presented in relevant figures is the average of multiple assays and after normalization to growth hormone levels (Miller and Bieker, 1993).

Immunoblots and gel shift analysis

The level of proteins expressed by transfected constructs was monitored with an anti-GAL1–147 primary antibody (Santa Cruz Biologicals) and a peroxidase labeled anti-rabbit secondary antibody (Amersham). Interactions were visualized by the ECL Western Blotting Analysis System (Amersham). Protein samples from transfected cells were

prepared as for CAT assays. Protein amounts were normalized to growth hormone levels from the same cells. SDS–polyacrylamide gel electrophoresis and electrotransfer to nitrocellulose were by established procedures (Sambrook *et al.*, 1989).

Gel shift analysis was performed as previously described using either a CAC site- or GAL4 site-containing oligonucleotide (Miller and Bieker, 1993).

Acknowledgements

We thank Dr Jon Licht for a critical reading of this manuscript. This work was supported by funds from the Cooley's Anemia Foundation to X.C. and PHS Grant DK46865 to J.J.B., who is a Scholar of the Leukemia Society of America. This is manuscript #218 from the Brookdale Center for Molecular Biology at the Mount Sinai Medical Center.

References

Bieker,J.J. (1994) Role of Erythroid Krüppel-like Factor (EKLF) in erythroid-specific transcription. In Stamatoyannopoulos,G. (ed.), *Molecular Biology of Hemoglobin Switching*. Intercept, Ltd, Andover, Vol. 1, pp. 231–241.

Bieker,J.J. (1996) Isolation, genomic structure, and expression of human Erythroid Krüppel-like Factor (EKLF). *DNA Cell Biol.*, **15**, 347–352.

Bieker,J.J. and Southwood,C.M. (1995) The Erythroid Krüppel-like Factor (EKLF) transactivation domain is a critical component for cell-specific inducibility of a β-globin promoter. *Mol. Cell. Biol.*, **15**, 852–860.

Binetruy,B., Smeal,T. and Karin,M. (1991) Ha-Ras augments c-Jun activity and stimulates phosphorylation of its activation domain. *Nature*, **351**, 122–127.

Bungert,J., Dave,U., Lim,K.C., Lieuw,K.H., Shavit,J.A., Liu,Q. and Engel,J.D. (1995) Synergistic regulation of human β-globin gene switching by locus control region elements HS3 and HS4. *Genes Dev.*, **9**, 3083–3096.

Chiang,C.M. and Roeder,R.G. (1995) Cloning of an intrinsic human TFIID subunit that interacts with multiple transcriptional activators. *Science*, **267**, 531–536.

Chou,P.Y. and Fasman,G.D. (1978) Empirical predictions of protein conformation. *Annu. Rev. Biochem.*, **47**, 251–256.

Colgan,J., Wampler,S. and Manley,J.L. (1993) Interaction between a transcriptional activator and transcription factor IIB *in vivo*. *Nature*, **362**, 549–553.

Donze,D., Townes,T.M. and Bieker,J.J. (1995) Role of Erythroid Krüppel-like Factor (EKLF) in human γ- to β-globin switching. *J. Biol. Chem.*, **270**, 1955–1959.

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

Felsenfeld,G. (1992) Chromatin as an essential part of the transcriptional mechanism. *Nature*, **355**, 219–224.

Feng,W.C., Southwood,C.M. and Bieker,J.J. (1994) Analyses of β-thalassemia mutant DNA interactions with erythroid Krüppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J. Biol. Chem.*, **269**, 1493–1500.

Frankel,A.D., Berg,J.M. and Pabo,C.O. (1987) Metal-dependent folding of a single zinc finger from transcription factor IIIA. *Proc. Natl Acad. Sci. USA*, **84**, 4841–4845.

Garnier,J., Osguthorpe,D.J. and Robson,B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. *J. Mol. Biol.*, **120**, 97–120.

Gill,G. and Ptashne,M. (1988) Negative effect of the transcriptional activator GAL4. *Nature*, **334**, 721–724.

Gonzalez,G.A. and Montminy,M.R. (1989) Cyclic AMP stimulates somatostatin gene transcription by phosphorylation of CREB at serine 133. *Cell*, **59**, 675–680.

Gregory,R.C., Taxman,D.J., Seshasayee,D., Kensinger,M.H., Bieker,J.J. and Wojchowski,D.M. (1996) Functional interaction of GATA1 with Erythroid Krüppel-like Factor and SP1 at defined erythroid promoters. *Blood*, **87**, 1793–1801.

Hainaut,P. and Milner,J. (1993) Redox modulation of p53 conformation and sequence-specific DNA binding *in vitro*. *Cancer Res.*, **53**, 4469–4473.

Higgs,D.R. and Wood,W.G. (1993) Understanding erythroid differentiation. *Curr. Biol.*, **3**, 548–550.

Hupp,T.R. and Lane,D.P. (1994) Allosteric activation of latent p53 tetramers. *Curr. Biol.*, **4**, 865–875.

Hupp,T.R., Meek,D.W., Midgley,C.A. and Lane,D.P. (1992) Regulation of the specific DNA binding function of p53. *Cell*, **71**, 875–886.

Kim,T.K. and Roeder,R.G. (1994) Proline-rich activator CTF1 targets the TFIIB assembly step during transcriptional activation. *Proc. Natl Acad. Sci. USA*, **91**, 4170–4174.

Kreider,B.L., Orkin,S.H. and Ihle,J.N. (1993) Loss of erythropoietin reponsiveness in erythroid progenitors due to expression of the Evi-1 myeloid-transforming gene. *Proc. Natl Acad. Sci. USA*, **90**, 6454–6458.

Leuther,K.K., Salmeron,J.M. and Johnston,S.A. (1993) Genetic evidence that an activation domain of GAL4 does not require acidity and may form a β sheet. *Cell*, **72**, 575–585.

Lillie,J.W. and Green,M.R. (1989) Transcription activation by the adenovirus E1a protein. *Nature*, **338**, 39–44.

Livingstone,C., Patel,G. and Jones,N. (1995) ATF-2 contains a phosphorylation-dependent transcriptional activation domain. *EMBO J.*, **14**, 1785–1797.

Luscher,B. and Eisenman,R. (1990) New light on myc and myb. Part II. myb. *Genes Dev.*, **4**, 2235–2241.

Marais,R., Wynne,J. and Treisman,R. (1993) The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell*, **73**, 381–393.

Merika,M. and Orkin,S.H. (1995) Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Krüppel family proteins Sp1 and EKLF. *Mol. Cell. Biol.*, **15**, 2437–2447.

Meyer,M.E., Gronemeyer,H., Turcotte,B., Bocquel,M.T., Tasset,D. and Chambon,P. (1989) Steroid hormone receptors compete for factors that mediate their enhancer function. *Cell*, **57**, 433–442.

Migliaccio,G., Migliaccio,A.R., Broudy,V., Kreider,B., Rovera,G. and Adamson,J.W. (1989) Selection of lineage-restricted cell lines immortalized at different stages of hematopoietic differentiation from the murine cell line 32D. *J. Cell Biol.*, **109**, 833–841.

Miller,I.J. and Bieker,J.J. (1993) A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins. *Mol. Cell Biol.*, **13**, 2776–2786.

Mitchell,P.J. and Tjian,R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science*, **245**, 371–378.

Nerlov,C. and Ziff,E.B. (1994) Three levels of functional interaction determine the activity of CCAAT/enhancer binding protein-α on the serum albumin promoter. *Genes Dev.*, **8**, 350–362.

Nuez,B., Michalovich,D., Bygrave,A., Ploemacher,R. and Grosveld,F. (1995) Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature*, **375**, 316–318.

Orkin,S.H. (1990) Globin gene regulation and switching: circa 1990. *Cell*, **63**, 665–672.

Perkins,A.C., Sharpe,A.H. and Orkin,S.H. (1995) Lethal β-thalassemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature*, **375**, 318–322.

Petersen,J.M., Skalicky,J.J., Donaldson,L.W., McIntosh,L.P., Alber,T. and Graves,B.J. (1995) Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an α helix. *Science*, **269**, 1866–1869.

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

Pulverer,B.J., Kyriakis,J.M., Avruch,J., Nikolakaki,E. and Woodgett,J.R. (1991) Phosphorylation of c-jun mediated by MAP kinases. *Nature*, **353**, 670–674.

Russo,M.W., Matheny,C. and Milbrandt,J. (1993) Transcriptional activity of the zinc finger protein NGFI-A is influenced by its interaction with a cellular factor. *Mol. Cell. Biol.*, **13**, 6858–6865.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Southwood,C.M., Downs,K.M. and Bieker,J.J. (1996) Erythroid Krüppel-like Factor (EKLF) exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. *Dev. Dyn.*, **206**, 248–259.

Struhl,K. (1991) Mechanisms for diversity in gene expression patterns. *Neuron*, **7**, 177–181.

Suzow,J. and Friedman,A.D. (1993) The murine myeloperoxidase promoter contains several functional elements, one of which binds a cell type-restricted transcription factor, Myeloid Nuclear Factor 1 (MyNF1). *Mol. Cell. Biol.*, **13**, 2141–2151.

Tanese,N., Pugh,B.F. and Tijan,R. (1991) Coactivators for a proline-rich activator purified from the multisubnit human TFIID complex. *Genes Dev.*, **5**, 2212–2224.

Tjian,R. and Maniatis,T. (1994) Transcriptional activation: a complex puzzle with few easy pieces. *Cell*, **77**, 5–8.

Townes,T.M. and Behringer,R.R. (1990) Human globin locus activation region (LAR): role in temporal control. *Trends Genet.*, **6**, 219–223.

Wang,Y. and Prives,C. (1995) Increased and altered DNA binding of human p53 by S and G2/M but not G1 cyclin-dependent kinases. *Nature*, **376**, 88–91.

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

Williamson,M.P. (1994) The structure and function of proline-rich regions in proteins. *Biochem. J.*, **297**, 249–260.

Yu,H., Chen,J.K., Feng,S., Dalgarno,D.C., Brauer,A.W. and Schreiber,S.L. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell*, **76**, 933–945.

Received on May 20, 1996; revised on July 16, 1996