

# Context-dependent EKLF responsiveness defines the developmental specificity of the human $\epsilon$ -globin gene in erythroid cells of YAC transgenic mice

Keiji Tanimoto,<sup>1</sup> Qinghui Liu,<sup>1,3</sup> Frank Grosveld,<sup>2</sup> Jörg Bungert,<sup>1,3</sup> and James Douglas Engel<sup>1,5</sup>

<sup>1</sup>Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, Illinois 60208-3500, USA; <sup>2</sup>Department of Cell Biology and Genetics, Erasmus University, Rotterdam 3000DR, Holland

We explored the mechanism of definitive-stage  $\epsilon$ -globin transcriptional inactivity within a human  $\beta$ -globin YAC expressed in transgenic mice. We focused on the globin CAC and CAAT promoter motifs, as previous laboratory and clinical studies indicated a pivotal role for these elements in globin gene activation. A high-affinity CAC-binding site for the erythroid krüppel-like factor (EKLF) was placed in the  $\epsilon$ -globin promoter at a position corresponding to that in the adult  $\beta$ -globin promoter, thereby simultaneously ablating a direct repeat (DR) element. This mutation led to EKLF-independent  $\epsilon$ -globin transcription during definitive erythropoiesis. A second 4-bp substitution in the  $\epsilon$ -globin CAAT sequence, which simultaneously disrupts a second DR element, further enhanced ectopic definitive erythroid activation of  $\epsilon$ -globin transcription, which surprisingly became EKLF dependent. We finally examined factors in nuclear extracts prepared from embryonic or adult erythroid cells that bound these elements *in vitro*, and we identified a novel DR-binding protein (DRED) whose properties are consistent with those expected for a definitive-stage  $\epsilon$ -globin repressor. We conclude that the suppression of  $\epsilon$ -globin transcription during definitive erythropoiesis is mediated by the binding of a repressor that prevents EKLF from activating the  $\epsilon$ -globin gene.

[Key Words: globin; switching; transgenic; YAC; EKLF; DRED]

Received May 30, 2000; revised version accepted September 18, 2000.

The human  $\beta$ -globin genes are organized with the embryonic  $\epsilon$ -globin gene located most 5', followed by the two fetal  $\gamma$ -globin genes (A $\gamma$  and G $\gamma$ ), whereas the adult  $\delta$ - and  $\beta$ -globin genes are at the 3' end of the locus (Fig. 1A, top; Stamatoyannopoulos and Neinhuis 1994). The  $\epsilon$ -globin gene is expressed during the first six weeks of gestation in primitive, nucleated erythroid cells of the yolk sac, while the  $\gamma$ - and  $\beta$ -globin genes are silent. The first switch occurs when the site of erythropoiesis shifts to definitive hematopoietic cells of the fetal liver, when the  $\epsilon$ -globin gene is silenced and the  $\gamma$ -globin genes are concomitantly activated. Gradually at around the time of birth, the site of hematopoiesis changes once more to the bone marrow and spleen, where the adult  $\beta$ -globin gene is transcribed and the  $\gamma$ -globin genes are reciprocally silenced.

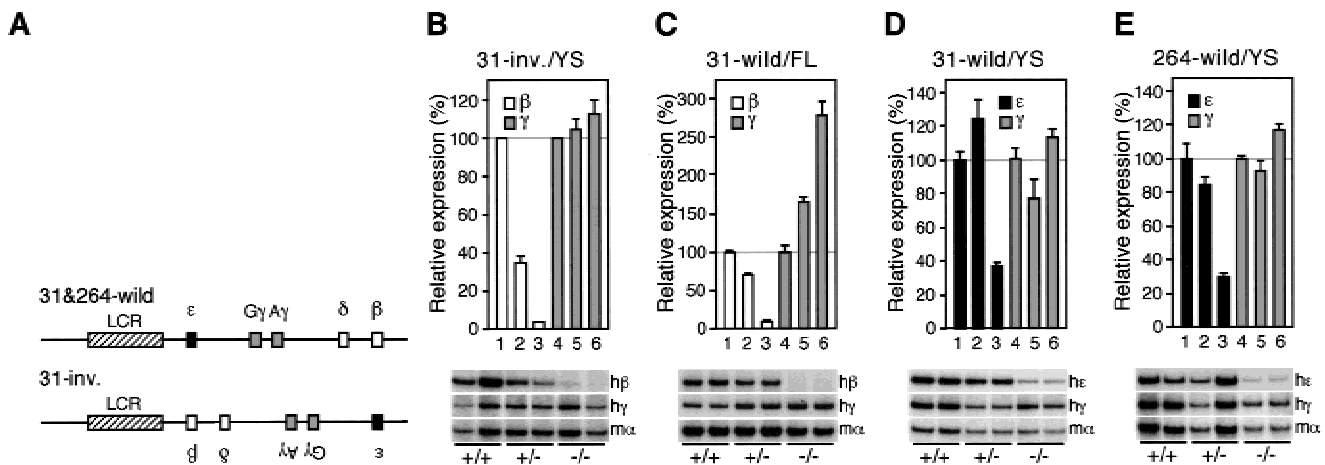
The individual genes of the human  $\beta$ -globin locus are developmentally regulated through their proximal promoters, although in transgenic mice the expression of individual globin genes is usually low and quite variable. This variability is due, in part, to integration position effects, which can be overcome by linking the genes *in cis* to the locus control region (LCR), a regulatory element composed of multiple DNase I hypersensitive sites (HS sites) (Townes et al. 1985; Chada et al. 1986; Kollias et al. 1986; Forrester et al. 1987; Grosveld et al. 1987). Consequently, high-level, position-independent expression is observed when the human  $\gamma$ - or  $\beta$ -globin genes are linked to the LCR (Grosveld et al. 1987; Enver et al. 1989; Behringer et al. 1990). However, these experiments also showed that the temporal regulation of individual LCR-linked globin genes was compromised. Only when tandemly linked to the LCR were the  $\gamma$ - and  $\beta$ -globin genes expressed at their proper developmental stages, indicating that the fetal ( $\gamma$ ) to adult ( $\beta$ ) switch is controlled by promoter competition (i.e., competition of the two promoters for a shared enhancer, in this case the LCR [Choi and Engel 1988; Behringer et al. 1990; Enver et al. 1990]). We recently showed that  $\epsilon$ -globin gene activation is also regulated competitively during primitive erythropoiesis (Tanimoto et al. 1999a). In contrast to temporally

Present addresses: <sup>3</sup>Department of Pharmacology, University of Illinois College of Medicine, Chicago, IL 60612, USA; <sup>4</sup>Department of Biochemistry and Molecular Biology, University of Florida, Gainesville, FL 32610, USA.

<sup>5</sup>Corresponding author.

E-MAIL [d-engel@northwestern.edu](mailto:d-engel@northwestern.edu); FAX (847) 467-2152.

Article and publication are at [www.genesdev.org/cgi/doi/10.1101/gad.822500](http://www.genesdev.org/cgi/doi/10.1101/gad.822500).



**Figure 1.** Evaluation of EKLK contribution to the activity of human globin genes in YAC transgenic mice. (A) All of the lines examined are single copy transgenes; lines 31-wild (C,D) and 264-wild (E) bear wild-type (top) and line 31-inv. (B) bears inverted (bottom) human  $\beta$ -globin YAC transgenes (Tanimoto et al. 1999a). (B–E) Total RNA was prepared from embryos derived from the intercross of male (EKLK<sup>+/-</sup>::YAC Tg<sup>+/-</sup>) and female (EKLK<sup>+/-</sup>) animals. The EKLK genotype of each embryo was determined by allele-specific PCR analysis. Samples of the various EKLK genotypes were analyzed individually at least three times. The expression levels of the  $\beta$ -like globin genes, normalized to that of the endogenous mouse  $\alpha$ -globin gene, were internally compared and the relative level of expression (with expression of EKLK<sup>+/+</sup> mice set at 100) was statistically analyzed. Data were collected from more than one litter. The average and standard deviation (S.D.) is graphically depicted for the genotypic group of more than two animals and the average is shown for two samples (at least two samples were analyzed for each group). Bars 1 and 4 in each histogram represent the expression level of the  $\beta$ -like globin genes in EKLK<sup>+/+</sup>, 2 and 5 in EKLK<sup>+/-</sup>, and 3 and 6 in EKLK<sup>-/-</sup> mice. Representative results of RT-PCR for human  $\epsilon$  (h $\epsilon$ ),  $\gamma$  (h $\gamma$ ),  $\beta$  (h $\beta$ ), and mouse  $\alpha$  (m $\alpha$ ) in different EKLK backgrounds (+/+, +/-, or -/-) are shown below each panel. (B)  $\beta$ - (open) and  $\gamma$  (shaded)-globin gene expression in the embryonic yolk sac; (C)  $\beta$ - and  $\gamma$ -globin gene expression in the 14.5-dpc fetal liver; (D,E)  $\epsilon$ - (black) and  $\gamma$  (shaded)-globin gene expression in the embryonic yolk sac.

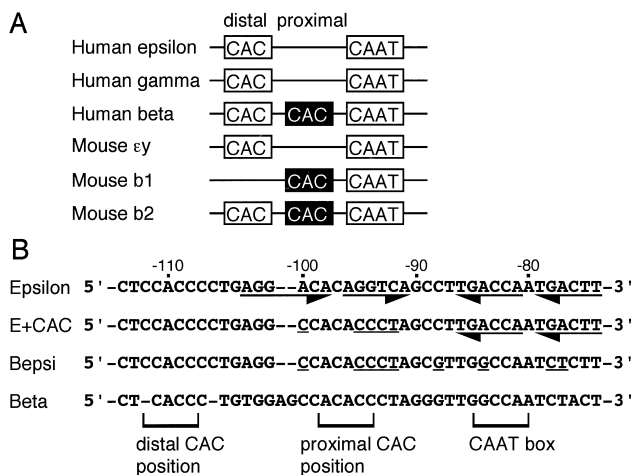
inappropriate expression of the single  $\gamma$ - or  $\beta$ -globin genes, when the  $\epsilon$ -globin gene is joined to the LCR, it is both activated and silenced autonomously (Raich et al. 1990).

A priori, the  $\epsilon$ -globin gene could be silenced by two nonexclusive mechanisms, one involving active transcriptional repression and the other a simple lack of activation, in definitive erythroid cells. Many previous studies focused on identifying negative regulatory elements in the  $\epsilon$ -globin promoter and more 5' sequences (Raich et al. 1992, 1995; Peters et al. 1993; J. Li et al. 1998; Q. Li et al. 1998). Raich et al. (1995) reported that mutating the combination of YY1 and GATA factor binding sites in the distal promoter region of the  $\epsilon$ -globin gene lead to its derepression in adult erythroid cells of transgenic mice. However, when this putative silencer was deleted from a human  $\beta$ -globin YAC containing the whole locus, no derepression of  $\epsilon$ -globin transcription was observed in transgenic animals, underscoring the critical need to analyze postulated regulatory elements within the context of the entire locus (Liu et al. 1997).

A second possibility is that the  $\epsilon$ -globin gene is inactive in definitive cells simply because it lacks appropriate activating factors that bind to its promoter region, an hypothesis that has not been widely explored. The CAC element is found in erythroid as well as many other genes. Each of the human  $\beta$ -like globin genes have one ( $\epsilon$ - and  $\gamma$ -globin) or two ( $\beta$ -globin) CAC boxes in their promoters (Fig. 2A). The critical nature of this ele-

ment to  $\beta$ -globin transcription was discovered many years ago, when specific human  $\beta$ -thalassemias were shown to be associated with mutations in the proximal CAC box of the adult gene (Orkin et al. 1982, 1984). Several proteins interact with CAC sequences in vitro (Hartzog and Myers 1993). These include the ubiquitous transcription factor Sp1 (Gumucio et al. 1991; Yu et al. 1991; Hartzog and Myers 1993) and the erythroid Krüppel-like factor (EKLK; Miller and Bieker 1993). EKLK shows extensive homology to Sp1 and binds to the proximal CAC box in the adult  $\beta$ -globin gene promoter with high affinity and can activate  $\beta$ -globin transcription in vitro (Miller and Bieker 1993; Feng et al. 1994).

Targeted disruption of the EKLK gene resulted in a thalassaemic phenotype leading to severe anemia and death at around 14 days post coitus (dpc), thereby demonstrating an indispensable role for EKLK in  $\beta$ -globin gene expression (Nuez et al. 1995; Perkins et al. 1995). The role of EKLK in human globin gene regulation was analyzed by crossing transgenic mice carrying the human  $\beta$ -globin locus into the EKLK-null mutant background (Perkins et al. 1996; Wijgerde et al. 1996). These experiments showed first, that the human embryonic  $\epsilon$ - and fetal  $\gamma$ -globin genes are expressed at normal levels in EKLK<sup>-/-</sup> embryos, and second, that adult  $\beta$ -globin transcription was undetectable. Thus, although all of the human  $\beta$ -like globin genes have potential EKLK-binding sites in their promoters, the EKLK mutant studies concluded that only the adult  $\beta$ -globin gene is affected by EKLK loss of function in vivo.



**Figure 2.** Comparison of the proximal promoter regions of  $\beta$ -type globin genes. (A) Positions of distal (open box) or proximal (solid box) CAC boxes are shown relative to the CAAT box in human  $\epsilon$ ,  $\gamma$ , and  $\beta$ , and murine  $\epsilon y$ , b1, and b2 genes. For the human  $\gamma$  gene, the  $\gamma$  distal CAAT box and its upstream region was used for alignment. (B) Sequence alignment of human embryonic  $\epsilon$ - and adult  $\beta$ -globin gene promoter regions as well as the mutant  $\epsilon$ -globin promoters examined in this study (E + CAC and Bepsi). Mutated nucleotides in the two promoters are underlined. Positions corresponding to the distal CAC, proximal CAC, and CAAT boxes are bracketed. Putative binding sites for COUP-TFII in K562 and DRED in MEL cells are underlined with arrows.

The late time of death of the EKLK mutant mice (14 dpc) was initially perplexing, because EKLK is expressed and functional in embryonic erythroid cells (Southwood et al. 1996). One plausible hypothesis for why ablation of EKLK did not affect expression of the  $\epsilon$ - or  $\gamma$ -globin promoters in primitive erythroid cells could be that their CAC boxes are not functional binding sites for EKLK in vivo. In accord with this hypothesis, it was shown that the CAC box of the adult  $\beta$ -globin gene has an eightfold higher affinity for EKLK than does the  $\gamma$  gene CAC box (Donze et al. 1995). However, it was later shown that this difference in binding affinity alone cannot explain the lack of  $\gamma$  gene activation by EKLK: When the 9-bp CAC sequence of the  $\beta$ -globin gene promoter was converted to the corresponding  $\gamma$  gene sequence, EKLK activated the two promoters equally (Asano and Stamatoyannopoulos 1998). Thus, the CAC sequence per se is not the sole determinant of differential  $\gamma$ -globin activation by EKLK. This conclusion provided the intriguing possibility that transcriptional output from EKLK-dependent binding sites might be contextually dependent on nearby binding sites for other factors.

Our primary goal on embarking on this study was to determine why EKLK does not activate the CAC site in the  $\epsilon$ -globin gene promoter during definitive erythropoiesis and what molecular mechanisms regulate this phenomenon. We presumed that these studies might also clarify why the  $\epsilon$ -globin gene is inactive in definitive erythroid cells. To address these questions, we proposed to define what the  $\epsilon$ -globin promoter is lacking for EKLK-

dependent activation in definitive erythroid cells. Our approach was to transfer specific binding sites from the  $\beta$ -globin gene that are known to be active at the definitive stage into the  $\epsilon$ -globin promoter. We generated two different  $\epsilon$ -globin promoter mutants, one containing the adult  $\beta$ -globin high-affinity EKLK-binding site (E + CAC), and a second incorporating the same mutation with a modified CAAT box, which even more closely mimics the adult  $\beta$ -globin gene promoter (Bepsi). We found that both mutated  $\epsilon$ -globin genes continued to be expressed in definitive erythroid cells, indicating that mutation of only 5 bp of the  $\epsilon$ -globin gene promoter (E + CAC), even within the context of a 150-kb YAC, markedly derepressed  $\epsilon$ -globin transcription. Surprisingly, when this same E + CAC mutant promoter was analyzed in mice lacking the EKLK gene, derepression was not abolished but was instead enhanced during definitive erythropoiesis, demonstrating that the effect of the modified CAC box on  $\epsilon$ -globin transcription is not mediated by EKLK. In contrast, the increase in definitive erythroid-specific expression from the combined CAC + CAAT mutant (Bepsi) was no longer observed in mice missing EKLK; that is, this mutant surprisingly adopted EKLK dependence in the presence of a second mutation adjacent to the CAC sequence. We conclude that the detailed arrangement of *cis* regulatory elements within the  $\epsilon$ -globin promoter mediate its EKLK responsiveness through a stage-specific silencing mechanism.

## Results

### *EKLK activates $\beta$ -globin transcription in primitive and definitive erythroid cells*

Previous work showed that EKLK is indispensable for adult  $\beta$ -globin gene transcription in definitive murine fetal liver erythroid cells. The same studies showed that EKLK is not required for  $\epsilon$ - or  $\gamma$ -globin gene expression in primitive erythroid cells of the embryonic yolk sac (Perkins et al. 1996; Wijgerde et al. 1996), although it is present and active during the primitive stage (Southwood et al. 1996; Tewari et al. 1998). We wished to first confirm that EKLK is active in the embryonic yolk sac, and then second, to test whether or not EKLK was required for activation of the  $\beta$ -globin gene in the yolk sac environment.

Previously, we generated transgenic mice in which the globin genes were inverted with respect to the LCR (Fig. 1A) and found that the adult  $\beta$ -globin gene in the inverted locus is expressed abundantly in primitive cells (Tanimoto et al. 1999a). To test whether embryonic expression of the adult  $\beta$ -globin gene requires EKLK, we bred the inverted human  $\beta$ -globin YAC transgene into the EKLK-null mutant background and analyzed globin mRNA expression by semiquantitative RT-PCR. In the locus-inverted transgenes, activation of the adult  $\beta$ -globin and suppression of the fetal  $\gamma$ -globin genes during the primitive stage (in the 9.5-day yolk sac) is regulated by EKLK in a gene dosage-dependent manner (Fig. 1B). We detect a similar dependence on EKLK for regulation of the adult  $\beta$ -globin gene during the fetal liver stage in

mice bearing the wild-type locus (Fig. 1C). The reciprocal increase in expression of the  $\gamma$ -globin genes (presumably caused by competition) is more dramatic in the wild type than in the genes-inverted locus. These results confirm that EKLK is active in primitive erythroid cells of the embryonic yolk sac and that it is essential for adult  $\beta$ -globin transcription in the embryonic erythroid environment, in accord with previous conclusions (Guy et al. 1998).

*EKLK is required for efficient  $\epsilon$ -globin transcription in the embryonic yolk sac*

We next examined the requirement for EKLK on primitive erythroid  $\epsilon$ -globin transcription in transgenic mice. In homozygous EKLK-null mutant mice, the  $\epsilon$ -globin transcript level declined precipitously (by 60%–70%) when compared with its transcription in controls (the normal configuration of the transgene in a wild-type background), demonstrating that EKLK is also required for full activation of the  $\epsilon$ -globin gene during the primitive stage (Fig. 1D,E). However, in contrast to  $\beta$ -globin transcription, activation of  $\epsilon$ -globin transcription is not EKLK-dose dependent nor is its effect on  $\epsilon$ -globin transcription as severe. These results imply that EKLK is an essential factor for  $\beta$ -globin transcription during both the primitive and definitive stages but that it plays only an ancillary role in  $\epsilon$ -globin transcription in the yolk sac (Tewari et al. 1998).

*Generation of chimeric  $\epsilon$ -globin gene promoters*

We found that EKLK is essential for both embryonic- and adult-specific globin gene expression. Assuming that EKLK exerts its function through the CAC box present in the globin gene promoters, some mechanism must allow EKLK to distinguish between these promoters during development. Because small differences in the sequence of the CAC motif cannot account for the difference in stage-specific expression of the genes (Asano and Stamatoyannopoulos 1998), we examined the broader sequence context surrounding the CAC sites.

Previous studies showed that the integrity of the promoter-proximal CAC and CAAT motifs is crucial for  $\beta$ -globin transcription (Myers et al. 1986). When we aligned the human and mouse embryonic and adult  $\beta$ -type globin gene promoters with the CAAT site of the human adult  $\beta$ -globin gene (Fig. 2A), one clearly distinctive feature was that the adult murine and human  $\beta$ -globin promoters have a CAC site situated in a gene-proximal position, whereas the position of the CAC elements in the  $\epsilon$  (and human  $\gamma$ -globin) promoters is more distal to the CAAT box. A number of human  $\beta^+$ -thalassemias are associated with mutations in the proximal, but not the distal, CAC site of the  $\beta$ -globin gene (Treisman et al. 1983). These observations lead us to suspect that the spacing between the CAC and CAAT motifs (or between them and other factor recognition sites in this region) might play a role in determining the binding specificity of EKLK to the various globin gene promoters.

To test this hypothesis and, at the same time, to iden-

tify the minimum requirement for the  $\epsilon$ -globin gene to be activated in definitive cells, we generated two mutant promoters by substitution of adult  $\beta$ -globin promoter sequences into the wild-type  $\epsilon$ -globin gene. In the first mutant, we changed five nucleotides between the distal  $\epsilon$ -globin CAC box and the CAAT box to create a new, more  $\beta$ -like, proximal CAC box (Fig. 2B, called E + CAC), while maintaining identical spacing to the wild-type  $\epsilon$ -globin gene. For the second mutation, we substituted four additional nucleotides around the  $\epsilon$ -globin CAAT box into the E + CAC mutant, creating a more adult  $\beta$ -like CAAT sequence as well (Fig. 2B, Bepsi).

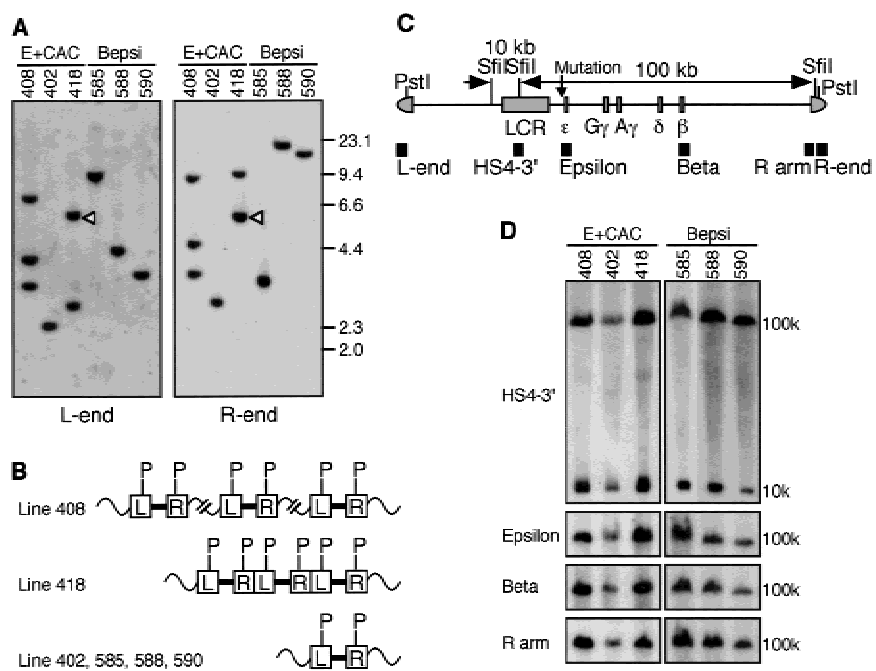
*YAC transgenic mice with site-specific mutations in the human  $\epsilon$ -globin promoter*

We introduced the two  $\epsilon$ -globin promoter mutations (E + CAC and Bepsi) separately into the 150-kb human  $\beta$ -globin YAC (A201F4.3) using homologous recombination in yeast (Gaensler et al. 1991; Bungert et al. 1995; Tanimoto et al. 1999b). Successful mutagenesis was confirmed by Southern blot analysis of YAC DNA (data not shown). We then purified the two mutant YACs from pulsed field gels (PFG) and injected them into fertilized mouse oocytes to generate transgenic animals. Tail DNA was prepared from offspring and screened using left (L) and right (R) YAC vector arm-specific PCR primers (data not shown; Bungert et al. 1995). Offspring with both L and R YAC vector arm sequences were then analyzed by detailed Southern blot hybridizations.

We generated three independent transgenic lines for each YAC (nos. 402, 408, and 418 for the E + CAC mutation and nos. 585, 588, and 590 for Bepsi). End-fragment analysis of thymus DNA recovered from each of these lines (Fig. 3A) revealed that four of the lines (402, 585, 588, and 590) carry single-copy YACs and that the other two (408 and 418) bear three copies (Fig. 3A,B). In E + CAC line 418, two end fragments were detected by both vector arm probes, one of which (R and L arm junction fragments) hybridized to both, indicating that the three copies had integrated in a direct head-to-tail repeat array. Further analysis using a fragment from the endogenous GATA-2 locus as a control confirmed this copy number analysis (data not shown).

In the YAC A201F4.3, the whole globin locus is contained within two *Sfi*I restriction enzyme fragments. One (10-kb) fragment contains LCR HS5 and HS4, and the other (100 kb) spans a region from within the LCR (5' to HS3), including all the  $\beta$ -like globin genes ( $\epsilon$ ,  $\gamma$ , and  $\beta$ ; Fig. 3C; Tanimoto et al. 1999b). High-molecular-weight DNA was prepared from the thymi of each line, embedded in agarose, and digested with *Sfi*I. After pulsed-field gel electrophoresis (PFGE) and transfer, we used probes spanning the locus and YAC vector arm (HS4–3',  $\epsilon$ -globin gene,  $\beta$ -globin gene, and the right arm, inside the *Sfi*I site; Fig. 3C) to detect either the 10- or 100-kb DNA fragments on Southern blots. All of the probes detected single bands of the expected sizes in all six lines (Fig. 3D), indicating that each carries intact, unfragmented copies of the transgene.

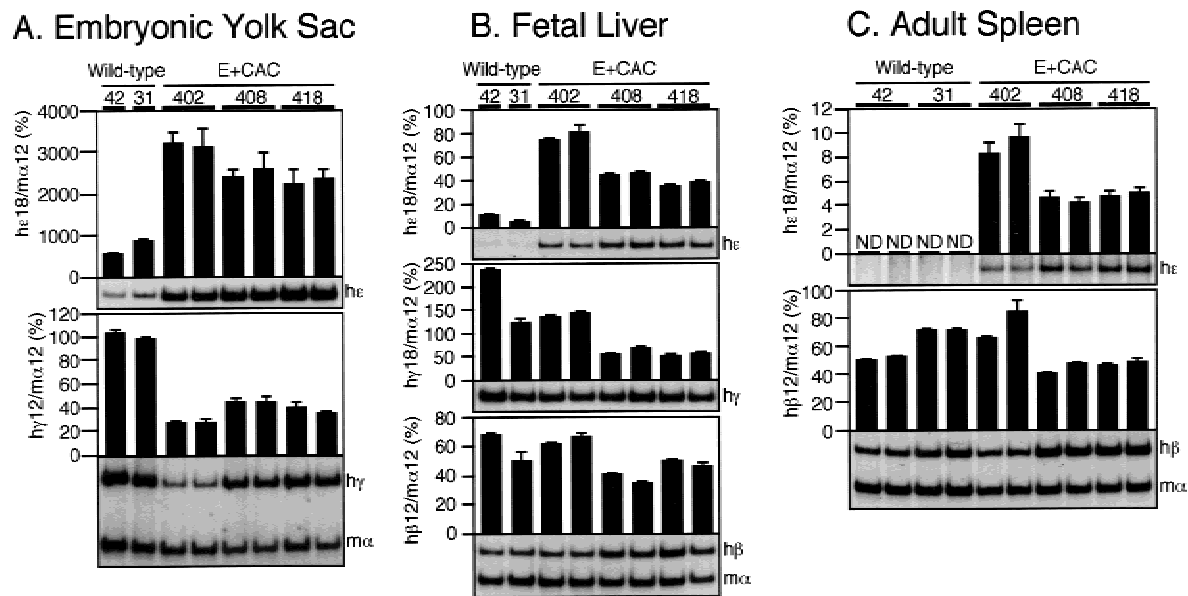
**Figure 3.** Structural analysis of human  $\beta$ -globin YACs in transgenic mice. (A) End-fragment analysis of  $\beta$ -globin YAC transgenic mice. Genomic DNA from transgenic thymus was digested with *Pst*I and blotted to nylon membranes. A 723-bp *Pst*I-*Alw*NI fragment (L-end, derived from the left YAC vector arm), and a 639-bp *Pvu*II-*Ava*I fragment (R-end, from the right YAC arm) were used as probes. These fragments are located outside of the *Pst*I restriction enzyme sites within the YAC vector arms (C), thus enabling analysis of the ends of the transgene at the genomic integration site. Using this strategy, it is possible to determine the arrangement of multiply integrated copies of the transgene (e.g., whether they integrate in a head-to-head, head-to-tail, or tail-to-tail configuration; Liu et al. 1998). Head-to-tail junction fragments are indicated by open triangles (in line 418). DNA size markers are shown at right (in kb). The identity of the transgenic lines are indicated by the number on top. (B) Structural organization of the human  $\beta$ -globin YACs in transgenic mice. Each of the four lines, 402, 585, 588, and 590 (bottom), yield end fragments with different sizes (A) for the L- and R-end probes, indicating that each of these lines harbor single copies of the transgene. Line 418 (middle) shows two bands with both L- and R-end probes and one of these is the size of a H-T junction fragment (open triangle in A). The intensity of this junction band is stronger than that of others, indicating that this line carries more than two copies of the transgene. Further internal comparison, using a fragment of the GATA-2 locus as a probe, confirmed that this line has three intact copies of the transgene (data not shown). Line 408 (top) has three fragments for both probes and none of them hybridized to both L and R probes. In subsequent generations, all of these fragments segregated together, indicating that three copies of the transgene are integrated close to one another but not directly linked. The results were confirmed by internal fragment assay using the GATA-2 probe (data not shown). (C) Schematic representation of the human  $\beta$ -globin YAC (A201F4.3) indicating the positions of *Sfi*I and *Pst*I restriction enzyme sites. The whole  $\beta$ -globin locus is contained within two *Sfi*I restriction enzyme fragments (10 kb and 100 kb, as indicated). The positions of the probes used for Southern blot analysis are depicted by the solid boxes. Mutations shown in Fig. 2B (5 and 9 nucleotides each for E + CAC and Bepsi, respectively) were introduced into the YAC by homologous recombination in yeast. (D) Integrity of the human  $\beta$ -globin YAC transgenes. Thymus cells from transgenic mice were embedded in agarose plugs and digested with *Sfi*I at 50°C (Tanimoto et al. 1999b). The DNA was then separated by PFGE, blotted to nylon membranes, and hybridized separately to probes (indicated at left) from the  $\beta$ -globin locus or from the right YAC vector arm (see C). The sizes of the expected bands are shown at right.



The E + CAC mutation enhances  $\epsilon$ -globin transcription at all developmental stages

To analyze the expression of the individual human globin genes in transgenic mice, we performed multiplex semiquantitative RT-PCR as detailed previously (Tanimoto et al. 1999b). At least four independent samples were analyzed for each line at the different developmental stages (9.5-dpc yolk sac and 14.5-dpc fetal liver or adult spleen). The expression level of each human  $\beta$ -type globin mRNA was internally normalized to that of mouse  $\alpha$ -globin and then divided by the transgene copy number to obtain expression values per gene copy. Intact, single copy wild-type human  $\beta$ -globin YAC transgenic lines (31 and 42; Tanimoto et al. 1999b) were used as the controls. In the embryonic yolk sac, expression of the E + CAC mutant  $\epsilon$ -globin gene increased three- to fourfold when compared with wild-type  $\epsilon$ -globin (Fig. 4A, top). At the same time, expression of the  $\gamma$ -globin genes declined to 30%–40% of their wild-type level (Fig. 4A, bottom), indicating (as expected) that both the  $\gamma$ - and

$\epsilon$ -globin genes are competitively regulated in primitive stage erythroid cells (Tanimoto et al. 1999b).  $\epsilon$ -globin transcription from the E + CAC mutant promoter was not silenced in the fetal liver (14.5 dpc; Fig. 4B, top) or the adult spleen (Fig. 4C, top), stages at which the wild-type  $\epsilon$ -globin gene is normally undetectable, thereby fulfilling one prediction of the original hypothesis under which this work was initiated. However, the reduction in expression of the competing  $\gamma$ - (Fig. 4B, middle) and  $\beta$ -globin (Fig. 4B, bottom, and C, bottom) genes in definitive erythroid cells was not as significant as it was during the embryonic stage. This may be due to the fact that the latter genes are normally highly expressed in the fetal liver and adult spleen, respectively, and thus the  $\epsilon$ -globin gene is unable to compete as effectively as it does in the yolk sac, where  $\gamma$ -globins provide the only competing promoters. This fact makes the observation of continued  $\epsilon$ -globin transcription during both of the definitive erythroid stages (fetal liver and adult spleen) even more significant, as allochronic expression of  $\epsilon$  in these tissues takes places within the strongly competitive definitive



**Figure 4.** Expression of the human  $\beta$ -type globin genes in the E + CAC mutant YAC. (A)  $\epsilon$ - and  $\gamma$ -globin gene expression in the embryonic yolk sac. RNA was isolated from the yolk sacs of two (except for wild-type lines 42 and 31; Tanimoto et al. 1999a) independent litters for each line at 9.5 dpc. Expression of human  $\epsilon$  ( $h\epsilon$ )- and human  $\gamma$  ( $h\gamma$ )/mouse  $\alpha$  ( $m\alpha$ )-globin genes was analyzed separately by semiquantitative RT-PCR because of the difficulty in comparing low levels of  $\epsilon$  with the more abundant  $\gamma/\alpha$  expression at the same cycle number (see Materials and Methods for more details). The signals for  $\epsilon$ -globin at 18 cycles and  $\gamma/\alpha$ -globin at 12 cycles were quantitated by PhosphorImager, and the ratios of  $\epsilon/\alpha$  (*top*) and  $\gamma/\alpha$  (*bottom*) were calculated (the mouse  $\alpha$  signal at 12 cycles was set at 100 and the values are normalized by transgene copy numbers) and illustrated as histograms. For each sample, the mean  $\pm$  standard deviation (SD) from at least three independent experiments was statistically analyzed; the SD is shown for each. Representative data are shown beneath each panel. (B)  $\epsilon$  (*top*)-,  $\gamma$  (*middle*)-, and  $\beta$  (*bottom*)-globin gene expression in the fetal liver. RNA was isolated from the fetal liver of 14.5-dpc embryos. Expression of  $\epsilon$ -,  $\gamma$ -, and human  $\beta$  ( $h\beta$ )/ $\alpha$ -globin genes was analyzed separately by RT-PCR. The signal for  $\epsilon$ - and  $\gamma$ -globin at 18 cycles and  $\beta/\alpha$ -globin at 12 cycles was quantified and the ratio of  $\epsilon/\alpha$ ,  $\gamma/\alpha$ , and  $\beta/\alpha$  was calculated ( $\alpha$  signal at 12 cycles was set at 100). (C)  $\epsilon$  (*top*)- and  $\beta$  (*bottom*)-globin gene expression in the anemic adult spleen. Two animals ( $\sim$ 4 wk old) representing each transgenic line were made anemic and RNA was isolated from the spleen. Expression of  $\epsilon$ - and  $\beta/\alpha$ -globin genes was analyzed separately by RT-PCR. No  $\gamma$ -globin expression was observed after 18 cycles of amplification. The signals for  $\epsilon$ -globin at 18 cycles and  $\beta/\alpha$ -globin at 12 cycles were quantified, and the ratios of  $\epsilon/\alpha$  and  $\beta/\alpha$  were calculated ( $\alpha$  signal at 12 cycles was set at 100).

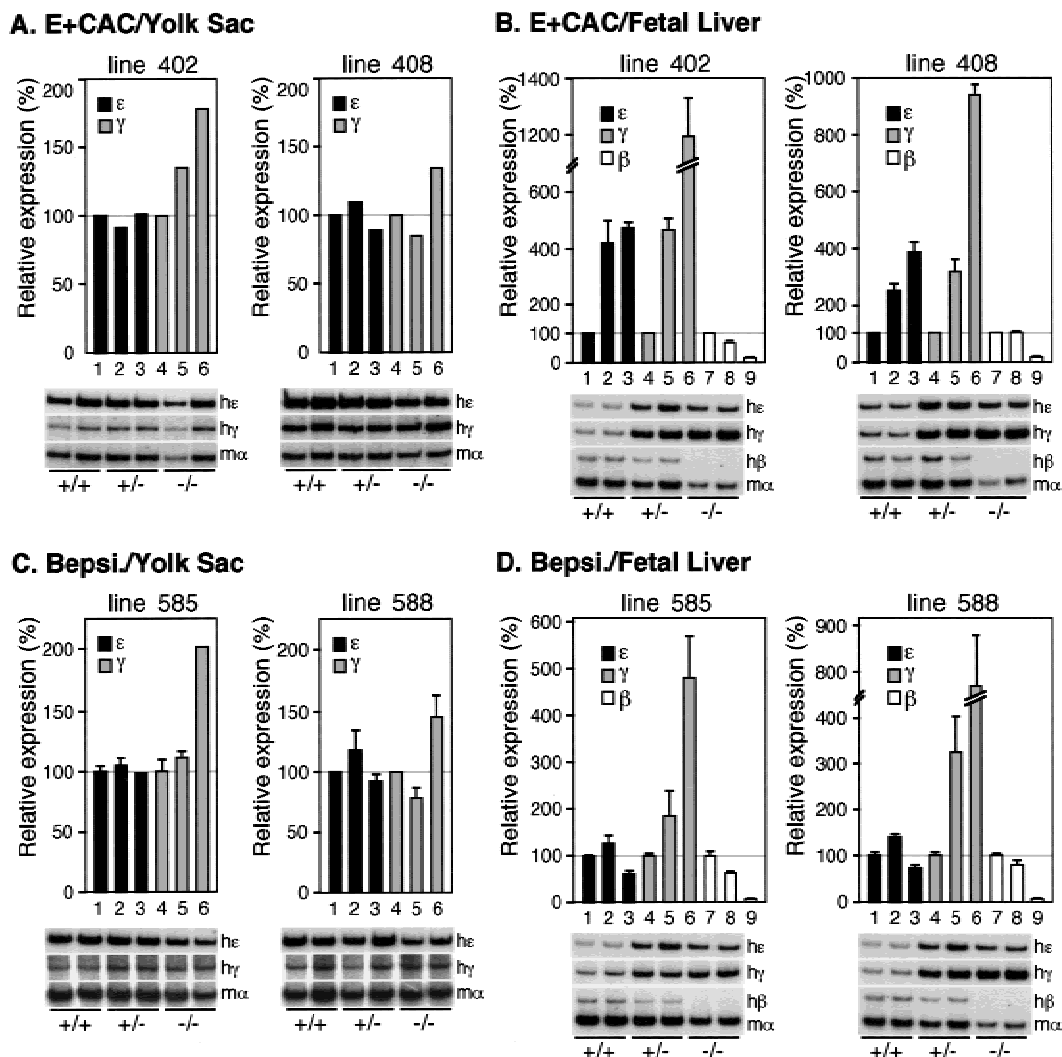
erythroid environment. Thus, the 5-bp E + CAC mutation in the  $\epsilon$ -globin gene promoter leads to a significant increase in  $\epsilon$ -globin gene transcription during primitive erythropoiesis and derepresses its transcription during definitive erythropoiesis.

#### *Derepression of $\epsilon$ -globin transcription by the E + CAC mutant promoter during definitive erythropoiesis is EKLK independent*

Because the effect of the E + CAC mutation was to increase expression of the  $\epsilon$ -globin gene at all erythroid developmental stages, we suspected that positive factors, such as Sp1 or EKLK, might be recruited to the newly generated proximal CAC site in a developmental stage-independent manner. To test this hypothesis, we crossed two of the E + CAC transgenic lines with mice in which the EKLK gene had been disrupted by germ-line gene targeting and then examined the effect of the loss of EKLK activity on E + CAC mutant  $\epsilon$ -globin gene transcription. In the embryonic yolk sac (Fig. 5A), no significant difference in  $\epsilon$  mRNA accumulation was observed in either EKLK heterozygous or homozygous mutant back-

grounds, demonstrating that EKLK does not enhance  $\epsilon$ -gene expression from the E + CAC promoter during the primitive stage. The fact that we did not observe any reduction in  $\epsilon$ -globin mRNA accumulation from the E + CAC promoter mutant in EKLK-deficient mice (in contrast to the dramatic reduction in  $\epsilon$ -globin transcripts in EKLK mutant mice bearing the wild-type locus; Fig. 1D,E) demonstrates that there is a significant EKLK-independent contribution of the E + CAC mutation to the enhancement of  $\epsilon$ -globin promoter activity (see Discussion).

The E + CAC mutant  $\epsilon$ -globin gene is expressed at higher levels in the fetal liver of EKLK mutant mice than it is in wild-type mice, as is also true for the  $\gamma$ -globin genes (Fig. 5B). Taken together, these results show that neither the mutant  $\epsilon$ -globin gene promoter nor the  $\gamma$ -globin gene promoters are EKLK dependent. One interpretation of this result is that the lack of competition by the  $\beta$ -globin gene (because of its requirement for EKLK) leads to an increase in competitive expression of any genes that are expressed in an EKLK-independent manner. Therefore EKLK appears to be effecting globin gene switching by preferentially activating the  $\beta$ -globin



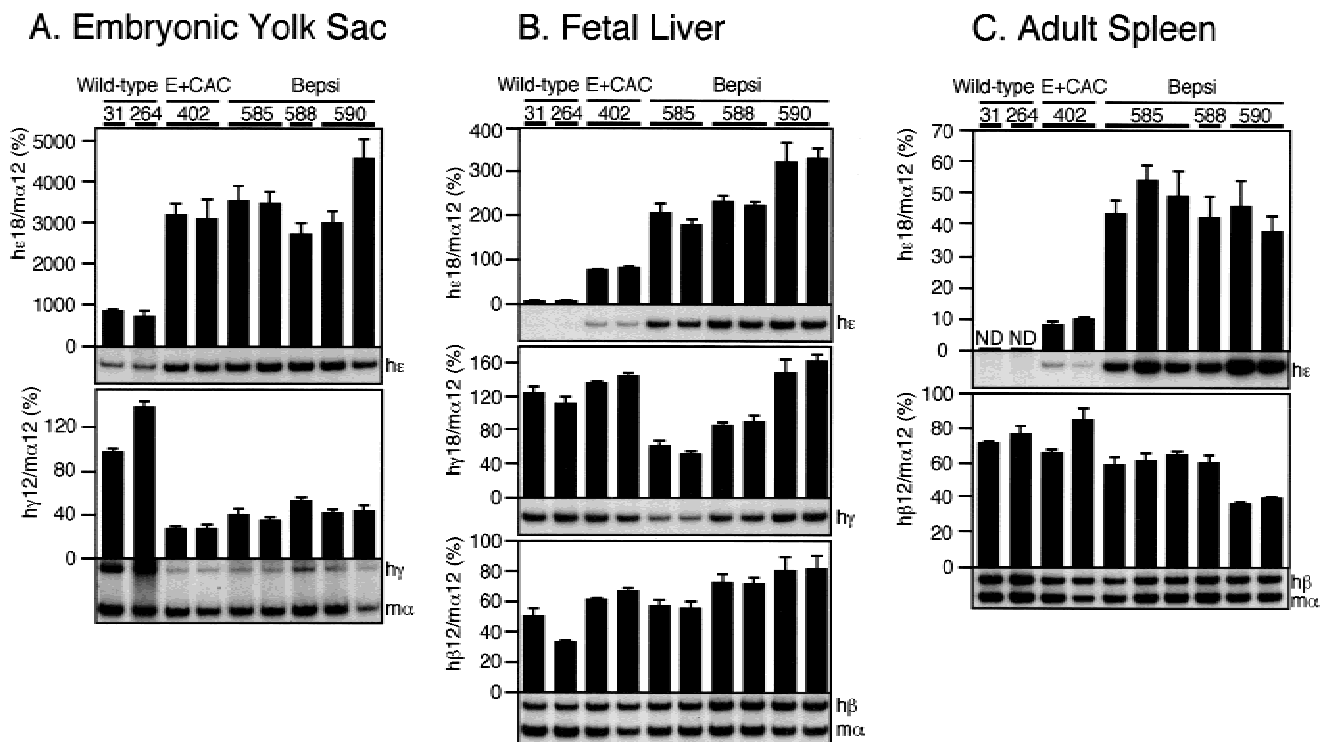
**Figure 5.** Enhanced activities of the Bepsi, but not E + CAC, promoters require EKLf. Two lines of mutant animals from each construct (E + CAC and Bepsi) were used to analyze the expression of human  $\beta$ -like globin genes in EKLf-null background. Total RNA was prepared and analyzed as described in the legend to Fig. 1. The expression levels of the  $\beta$ -like globin genes in EKLf<sup>+/+</sup> background was set at 100. The average and standard deviation (SD) is shown graphically for the genotypic group of more than two animals and only an average is shown for that of two (at least two samples were analyzed for each group). Expression of  $\beta$ -like globin genes in EKLf<sup>+/+</sup> (bars 1 and 4); in EKLf<sup>+/-</sup> (bars 2 and 5), and in EKLf<sup>-/-</sup> (bars 3 and 6) mice is shown. Representative results of RT-PCR for human  $\epsilon$  (h $\epsilon$ ),  $\gamma$  (h $\gamma$ ),  $\beta$  (h $\beta$ ), and mouse  $\alpha$  (m $\alpha$ ) in different EKLf mutant backgrounds (+/+, +/-, or -/-) are shown below each panel. (A) Embryonic (9.5-dpc) expression of  $\epsilon$ - and  $\gamma$ -globin genes in the E + CAC transgenic mice. (B) Fetal liver (14.5-dpc) expression of  $\epsilon$ -,  $\beta$ -, and  $\gamma$ -globin genes in the E + CAC transgenic mice. (C) Embryonic (9.5-dpc) expression of  $\epsilon$ - and  $\gamma$ -globin genes in the Bepsi transgenic mice. (D) Fetal liver (14.5-dpc) expression of  $\epsilon$ -,  $\beta$ -, and  $\gamma$ -globin genes in the Bepsi transgenic mice.

gene promoter in definitive erythroid cells, thereby competitively suppressing expression of the (potentially active) mutant  $\epsilon$  and wild-type  $\gamma$ -globin gene promoters.

*The Bepsi mutation activates  $\epsilon$ -globin gene expression in a developmental stage-specific manner*

Because we did not observe EKLf-dependent activation of the E + CAC mutant promoter, we assumed that this mutant might be deficient (in comparison to the adult  $\beta$ -globin promoter) in some other way. Therefore, we next analyzed globin gene expression in the second  $\epsilon$ -glo-

bin substitution mutant promoter that was even more  $\beta$ -like (Bepsi). The level of  $\epsilon$ -globin gene transcription in the embryonic yolk sac was not significantly different in the E + CAC and Bepsi mutants (Fig. 6A). However, in the 14.5-dpc fetal liver (Fig. 6B) we observed a 2.5- to 4.2-fold increase in  $\epsilon$ -globin transcription from Bepsi when compared with the E + CAC mutant promoter, and the effect was even more pronounced in the adult spleen (from 4.7- to 5.4-fold higher; Fig. 6C). Because both of the mutants have identical CAC sites, definitive-stage-specific differences in their activities must be attributable to the additional mutation incorporated around the CAAT box in the Bepsi mutant.



**Figure 6.** Expression of the human  $\beta$ -globin genes in the Bepsi mutant YAC. Samples were analyzed in the same way as described in the legend to Fig. 4. (A)  $\epsilon$ - and  $\gamma$ -globin gene expression in the embryonic yolk sac. Samples from wild-type (lines 31 and 264; Tanimoto et al. 1999a) and E + CAC (line 402) transgenic mice were analyzed at the same time and typical data are shown. All of the reference lines as well as three of the Bepsi lines carry single copy transgenes, hence the data can be directly compared without normalization. The signals for  $\epsilon$ -globin at 18 cycles and  $\gamma/\alpha$ -globin at 12 cycles were quantified, and the ratios of  $\epsilon/\alpha$  (top) and  $\gamma/\alpha$  (bottom) were calculated (the mouse  $\alpha$  signal at 12 cycles was set at 100) and illustrated as histograms. For each sample, the mean  $\pm$  standard deviation (SD) from at least three independent experiments was statistically analyzed; the SD is shown for each. Representative data are shown beneath each panel. (B)  $\epsilon$ -,  $\gamma$ -, and  $\beta$ -globin gene expression in the fetal liver. The signal for  $\epsilon$ -globin and  $\gamma$ -globin at 18 cycles and  $\beta/\alpha$ -globin at 12 cycles was quantified and the ratio of  $\epsilon/\alpha$ ,  $\gamma/\alpha$ , and  $\beta/\alpha$  was calculated ( $\alpha$  signal at 12 cycles was set at 100). (C)  $\epsilon$ - and  $\beta$ -globin gene expression in the anemic adult spleen. The signals for  $\epsilon$ -globin at 18 cycles and  $\beta/\alpha$ -globin at 12 cycles were quantified, and then the ratios of  $\epsilon/\alpha$  and  $\beta/\alpha$  were calculated ( $\alpha$  signal at 12 cycles was set at 100).

#### Definitive stage-specific activation of the Bepsi mutant promoter is EKLK dependent

To determine whether the Bepsi mutant promoter was EKLK dependent, we analyzed the Bepsi mutant YACs in the EKLK-null mutant background. At the primitive stage (Fig. 5C), there was no difference in the level of  $\epsilon$ -globin mRNA in the wild-type versus the EKLK-null mutant backgrounds. We therefore concluded that the increase in  $\epsilon$ -globin transcription of the Bepsi mutant (in comparison to the wild-type locus) does not require primitive-stage EKLK activity. We anticipated this result because the E + CAC promoter, which has similar activity to Bepsi in primitive cells, is also EKLK independent (Fig. 5A). However, in definitive fetal liver erythroid cells, we found that Bepsi promoter activity (which is much more active than the E + CAC promoter at this stage; Fig. 6B) was now strongly EKLK dependent (Fig. 5D). As  $\beta$ -globin mRNA diminished in the EKLK-null mutant background,  $\gamma$  gene expression increased significantly, whereas  $\epsilon$  expression decreased by 30%–40%. Because  $\epsilon$ -globin transcription in the E + CAC mutant locus increases by ~fourfold in EKLK-null mice in the de-

finite fetal liver (Fig. 5B), diminished  $\epsilon$ -globin accumulation in the Bepsi mutant at the same stage gains added significance. Because expression of the E + CAC and Bepsi mutants in the absence of EKLK is almost identical during both the yolk sac and fetal liver stages, we conclude that the activation of Bepsi above that of the E + CAC promoter during definitive erythropoiesis is absolutely dependent on EKLK activity. Thus in contrast to the E + CAC mutant, which is not directly affected by EKLK loss in definitive cells, the additional CAAT mutation in the mutant  $\epsilon$ -globin promoter now confers EKLK responsiveness. These data show that the context of the CAC-binding site, rather than its sequence, confers the specificity for binding of the EKLK protein to the  $\beta$ -type globin genes.

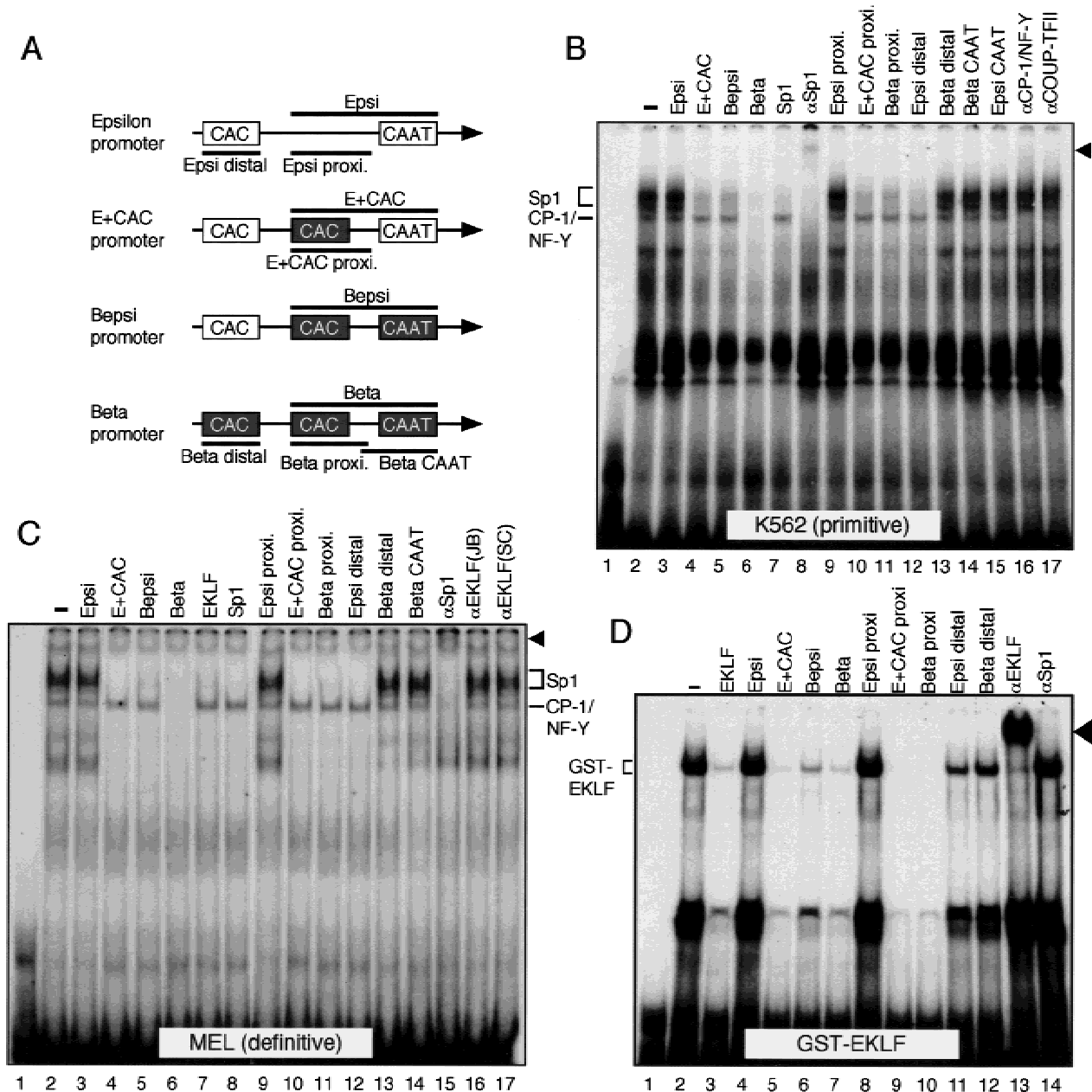
#### The E + CAC and Bepsi mutants bind to transcription factors that interact with the adult $\beta$ -globin gene promoter

To begin to decipher specific nuclear factor requirements for the recruitment of EKLK to the Bepsi mutant pro-



moter, we conducted EMSA. First, we determined whether the two mutant  $\epsilon$ -globin promoter fragments were capable of binding to factors that normally interact with the adult  $\beta$ -globin gene promoter. The radiolabeled probe corresponded to sequences in the proximal pro-

moter of the  $\beta$ -globin gene (Beta, Fig. 7A, bottom), whereas various subfragments of the wild-type and mutant promoters were used as competitors (Fig. 7A). The  $\beta$ -proximal fragment binds to virtually an identical set of proteins in both human K562 (primitive) and murine



**Figure 7.** Analysis of nuclear factor binding to the  $\beta$ -globin gene promoter. (A) Schematic representation of the proximal promoter region of  $\epsilon$ , E + CAC, Bepsi, and  $\beta$  genes and their fragments used as competitors in the EMSA experiments (thick lines). Positions of distal and proximal CAC as well as CAAT boxes are shown as open or black boxes (see Materials and Methods for detailed sequence information). (B) EMSA using cell extracts from K562 cells and a probe containing the proximal CAC and CAAT boxes from the adult  $\beta$ -globin promoter (Beta in A). A 200-fold molar excess of unlabeled DNA fragments was used as competitor unless stated otherwise. The supershift appearing in the presence of Sp1 antibody is indicated by an arrowhead (lane 8). (C) The same probe was used in EMSA with MEL cell extracts. The antibody/Sp1 supershift is depicted by an arrowhead (lane 15). (D) EMSA with affinity-purified GST-EKLF. An 80-fold molar excess of unlabeled competitor was used in this experiment. The protein/DNA complexes in lanes 11 and 12 disappeared when a 400-fold molar excess of competitor was added (Epsi distal and Beta distal; data not shown). The antibody/EKLF supershift (JB) is indicated by an arrowhead (lane 13).

MEL (definitive) erythroid cells (Fig. 7B,C). These binding proteins include CP-1/NF-Y (Fig. 7B, lane 16) and Sp1 (Fig. 7B, lanes 7,8 and 7C, lanes 8,15), in good agreement with previous reports (Hartzog and Myers 1993; Gillemans et al. 1998). A fragment corresponding to the proximal promoter of the wild-type  $\epsilon$ -globin gene (Epsi) failed to compete for the binding of these factors (Fig. 7B,C, lanes 3), whereas both of the mutant sequences (E + CAC and Bepsi) efficiently compete for Sp1, but not for CP-1/NF-Y. Sp1 binds to the proximal CAC site of the Beta fragment (Fig. 7B,C, lanes 11 and 14, respectively). Although the distal CAC site of  $\epsilon$ -globin can efficiently compete for Sp1 binding (Epsi distal; Fig. 7B,C, lanes 12), the distal CAC site of the adult  $\beta$  promoter does not (Beta distal; Fig. 7B,C, lanes 13). Inclusion of an antibody to CP-1/NF-Y eliminated the formation of the anticipated DNA/protein complex (Fig. 7B, lane 16;  $\alpha$ CP-1/NF-Y). Although CP-1/NF-Y was originally identified as a CCAAT box-binding protein, neither competition with the  $\epsilon$ - or  $\beta$ -globin CAAT sequences had any effect on complex formation (Fig. 7B, lanes 14,15 and 7C, lane 14). An antibody recognizing SP-1 interfered with the formation of both the CP-1/NF-Y and SP-1 complexes (Fig. 7B, lane 8 and 7C, lane 15) whereas a consensus sequence for Sp1 was unable to compete for CP-1/NF-Y binding (Fig. 7B, lane 7 and 7C, lane 8). The reason for this disparity is unclear.

Because EKLK gel shifts are partially obscured by Sp1 binding (Fig. 7C, lanes 16,17; see also Gillemans et al. 1998), we expressed EKLK in *Escherichia coli* as a GST-fusion protein (Bieker and Southwood 1995) to analyze the binding characteristics of EKLK to the various globin CAC sequences (Fig. 7D). The proximal  $\beta$ -globin CAC sequence was used as a high affinity EKLK binding site, and other globin promoter fragments were used as competitors. Specific binding of EKLK to the probe (lane 2) was disrupted by competition with Beta (lane 7) but not Epsi (lane 4). EKLK binds efficiently to both mutant  $\epsilon$ -globin promoter CAC boxes (E + CAC and Bepsi, lanes 5,6,8–10). The distal CAC elements of both the  $\epsilon$ - and  $\beta$ -globin promoters have lower affinity for EKLK than does the proximal  $\beta$  promoter CAC sequence (lanes 10–12, and data not shown). In summary, both E + CAC and Bepsi sequences bind to the same array of proteins that also interact with the adult  $\beta$ -globin gene promoter, including Sp1 and EKLK. Although this may explain why both the E + CAC and Bepsi promoters are activated in a developmental stage independent manner (because the  $\beta$ -globin promoter is potentially active in both primitive and definitive cells), these data do not explain why the E + CAC and Bepsi promoters have different activities and display a different dependence on EKLK in definitive erythroid cells.

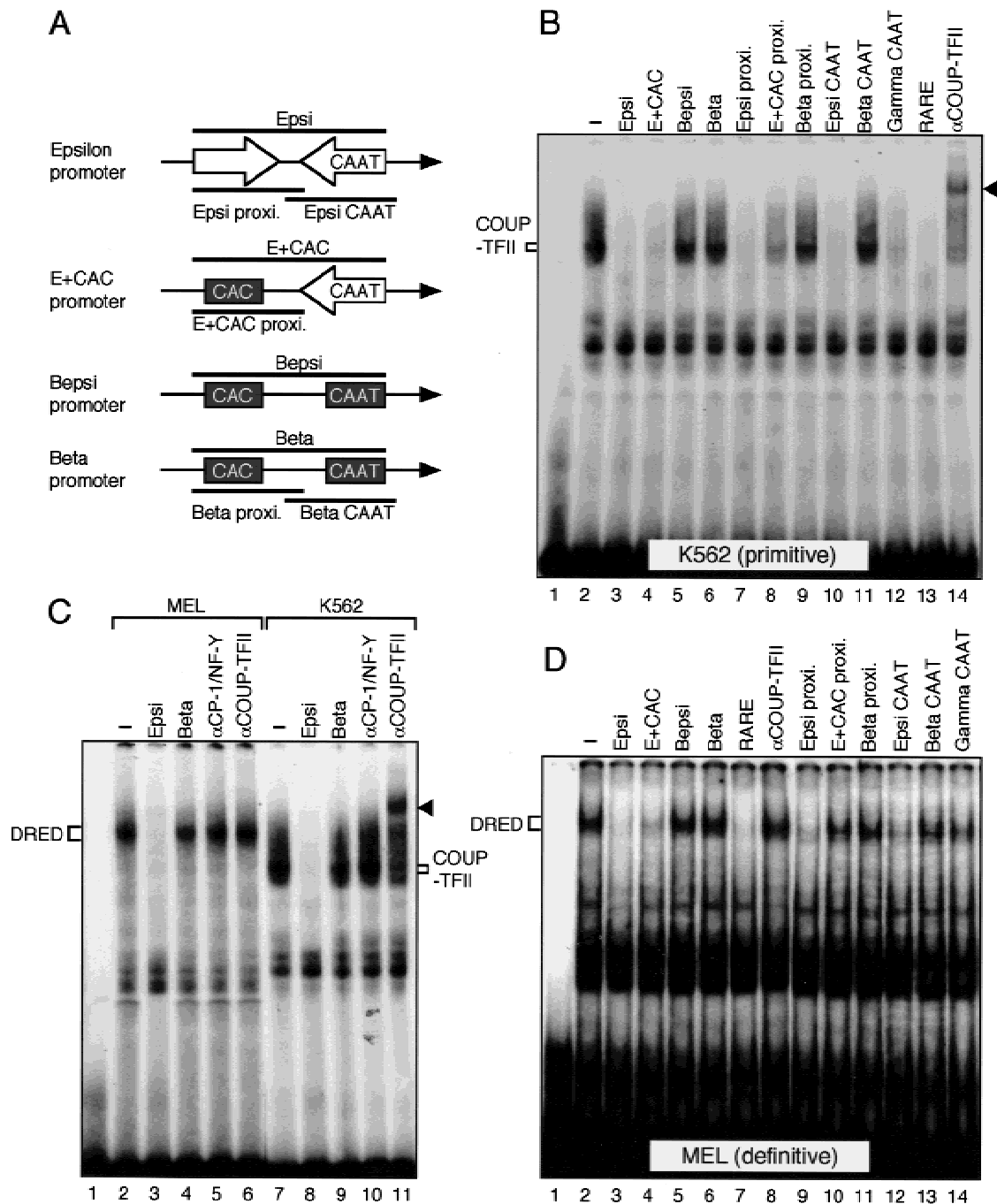
*E + CAC, but not Bepsi, binds to a definitive erythroid direct repeat nuclear factor*

Filipe et al. (1999) recently reported that a putative repressor protein, COUP-TFII, can bind to two sites in the  $\epsilon$ -globin promoter (Fig. 2B). To determine whether the

introduction of new sequences into the mutant  $\epsilon$ -promoter fragments might have disrupted these sites, we performed additional EMSA experiments using the wild-type  $\epsilon$ -globin promoter as a probe (Epsi; Fig. 8A, top). Incubation of the Epsi probe with K562 extracts led to the formation of one major and several minor complexes (Fig. 8B, lane 2). The major complex was eliminated by the RARE sequence (a strong COUP-TFII binding consensus; Fig. 8B, lane 13) and was supershifted by anti-COUP-TFII antibody (Fig. 8B, lane 14), confirming that this complex involves COUP-TFII. The corresponding region from the  $\beta$ -globin gene promoter (Beta; Fig. 8A, bottom) failed to compete (Fig. 8B, lane 6), confirming that COUP-TFII does not bind to the wild-type  $\beta$ -globin gene promoter.

The  $\epsilon$ -globin promoter contains two binding sites for COUP-TFII (Filipe et al. 1999), one in a gene proximal position and the other overlapping the CAAT box (shown as open arrows in Fig. 8A). Each of these sites is comprised of a direct repeat sequence, AGGTCA, referred to as the DR (direct repeat; Glass 1994, no. 3334). Introduction of the CAC box into the proximal  $\epsilon$ -globin promoter position destroys one DR consensus site, thereby disrupting the binding of COUP-TFII to that element (Fig. 8B, lane 8). The E + CAC fragment (Fig. 8A) can still bind to the COUP-TFII factor because of the second DR left intact in the Epsi probe (Epsi CAAT, Fig. 8B, lane 10). However, when the combined CAC and CAAT boxes in the Bepsi mutant were examined (in which both DR elements are mutated, Fig. 8B, lane 5), we found that COUP-TFII no longer bound. Although this could theoretically account for the difference between the E + CAC and Bepsi promoter activities, it does not, as COUP-TFII is not expressed in definitive erythroid cells (Filipe et al. 1999).

Because the difference in COUP-TFII binding cannot account for the divergent activities of the EKLK and Bepsi promoters, we next asked whether we could identify definitive stage-specific factors that could interact with the wild-type  $\epsilon$ -globin gene promoter. We found that one major protein complex is formed on the wild-type  $\epsilon$ -globin promoter when incubated with MEL cell extracts (Fig. 8C, lane 2), and it exhibits a decidedly different mobility than the complex formed by COUP-TFII (Fig. 8, cf. lanes 2 and 7). Antibody supershift experiments confirmed that this definitive erythroid nuclear protein is neither COUP-TFII (lane 6) nor CP-1/NF-Y (lane 5). Antibodies that specifically recognize other nuclear hormone receptors (including RAR  $\alpha$ ,  $\beta$ , and  $\gamma$  or RXR  $\alpha$ ,  $\beta$ , and  $\gamma$ ) also failed to alter the intensity or mobility of this complex (data not shown; Materials and Methods). This MEL cell-enriched factor, which we refer to as the DRED (direct repeat erythroid-definitive) protein, does not bind to the  $\beta$ -globin gene promoter, as is also true for COUP-TFII (Fig. 8C, lanes 4,9). When the binding sites for the DRED protein were examined in greater detail using many different competitors, the results indicated that DRED binds to the identical repeat sequence as COUP-TFII in the  $\epsilon$  promoter (Fig. 8D, lanes 9,12). Importantly, the E + CAC or Bepsi mutant frag-



**Figure 8.** Analysis of nuclear factor binding to the  $\epsilon$ -globin gene promoter. (A) Schematic representation of the promoter region of the  $\epsilon$ , E + CAC, Bepsi, and  $\beta$  genes, and fragments used as competitors in the EMSA experiments (thick lines). Positions of proximal CAC (black box) as well as  $\epsilon$  (in an open arrow) or  $\beta$  (black box) CAAT boxes are shown. Putative binding sites for both COUP-TFII and DRED are indicated by the open arrow (each arrow represents a direct repeat sequence). (B) EMSA with K562 cell extracts and a probe containing proximal sequences and CAAT box of the  $\epsilon$ -globin promoter (Epsi in A). A 200-fold molar excess of unlabeled DNA was used as competitor unless otherwise stated. The shifted band generated after including an antibody recognizing COUP-TFII is indicated with an arrowhead (lane 14). (C) Comparison of proteins from MEL and K562 cells that interact with the Epsi fragment. A comparable amount of protein extract was used for analysis. A 200-fold molar excess of competitor DNA or antibodies were used to test binding specificity. The supershifted COUP-TFII/antibody complex is detectable only in K562 cells and indicated by an arrowhead. (D) An EMSA similar to that shown in B was performed with MEL cell extract.

ments either only incompletely disrupt this binding or have no effect, respectively (Fig. 8D, lanes 4,5). Because both the binding- and developmental stage-specificity of

the DRED protein fulfills requirements that might be expected of a definitive erythroid repressor of  $\epsilon$ -globin transcription, this factor may be responsible for  $\epsilon$ -globin

promoter silencing as well as its nonresponsiveness to EKLK in definitive erythroid cells (see Discussion).

## Discussion

The timing and cell-type specificity of transcription is regulated by positive and negative regulatory *cis* elements in both gene-proximal (promoter) and -distal (enhancer, silencer) regions. An additional factor regulating gene expression in a positional and developmental stage-specific manner operates in multigene loci, where the relative position of genes, both with respect to one another as well as to shared regulatory elements, is critical (Ohtsuki and Levine 1998; Kmita et al. 2000). For example, we recently found that the adult  $\beta$ -globin gene can be ectopically activated at the embryonic stage if the wild-type configuration of the genes is inverted with respect to the LCR (Fig. 1B; Tanimoto et al. 1999a). This result indicated that all of the  $\beta$ -like globin genes are capable of activity at the embryonic stage and that the gene order in the wild-type locus normally inhibits  $\beta$ -globin from being expressed in primitive erythroid cells. However, in the wild-type locus the  $\epsilon$ - and  $\gamma$ -globin genes are autonomously silenced at the definitive stage (Raich et al. 1990; Dillon and Grosveld 1991), and therefore the adult  $\beta$ -globin gene is the only active transcription unit.

To decipher the underlying molecular mechanisms responsible for human  $\beta$ -globin gene switching (developmentally regulated temporal control), we set out to identify specific DNA sequence elements that confer definitive erythroid stage expression. We principally focused on the EKLK protein and its binding site, the CAC *cis* element, because they are the only species known to be essential for definitive-stage-specific expression of the  $\beta$ -globin gene. For the initial strategy, we aligned the proximal promoter regions of all the human and murine  $\beta$ -type globin genes. The wild-type genes expressed in definitive cells all have a CAC box in a position proximal to the (globin) ubiquitous CAAT box (Fig. 2A). Because natural mutations that severely affect adult  $\beta$ -globin gene expression in adults are often found in the gene-proximal CAC element (but never with the distal CAC box), we assumed a priori that this motif was important for definitive-specific expression. To test this hypothesis, we first created a high-affinity EKLK-binding site in the  $\epsilon$ -globin gene promoter at the same relative position it occupies in the adult  $\beta$ -globin gene promoter. We found that the E + CAC mutant  $\epsilon$ -globin gene is expressed in definitive erythroid cells of human  $\beta$ -globin YAC transgenic mice. However, this elevated  $\epsilon$ -globin mutant transcription was not stage specific;  $\epsilon$  transcription increased by three- to fivefold in the yolk sac, the fetal liver, and the adult spleen.

One plausible explanation for the increase in activity of the E + CAC mutant was that the modification creates a binding site for a positive factor that is active at all developmental stages. The EMSA experiments suggested that either Sp1 or EKLK could be that factor. Because high-level  $\beta$ -globin transcription at all developmental

stages is mediated by EKLK binding to the proximal CAC site, we suspected that the newly created CAC box in the  $\epsilon$ -promoter would also recruit EKLK. Surprisingly, however, the E + CAC mutant gene did not differ in expression when EKLK was removed by genetic manipulation (Fig. 5A). Moreover, its activity was enhanced in the fetal liver in the absence of EKLK (because the  $\beta$ -globin gene, which would normally compete for LCR activity at the fetal liver stage, is not expressed in the absence of EKLK; Fig. 5B). The results demonstrate that EKLK does not mediate the enhanced expression from the E + CAC mutant  $\epsilon$ -globin gene promoter.

A growing list of factors have been identified recently that are all able to bind to CAC sequences; these related proteins are called BKLF (Crossley et al. 1996), GKLF (Shields et al. 1996), LKLF (Anderson et al. 1995), UKLF (Matsumoto et al. 1998) and FKLF (Asano et al. 1999). BKLF is expressed in erythroid cells at even higher levels than EKLK (Crossley et al. 1996). FKLF was isolated from fetal liver erythroid cells and preferentially activates  $\epsilon$ - and  $\gamma$ -, but not  $\beta$ -globin gene promoters in primitive erythroid cells. It is therefore possible that either BKLF or FKLF is responsible for the activation of the E + CAC promoter in primitive cells of animals lacking EKLK, and we can address this question once mice are available which lack these transcription factors. *Sp1*-null mutant mice express endogenous embryonic globin genes at slightly reduced levels (Marin et al. 1997), suggesting that the newly created proximal CAC site on the E + CAC promoter might also recruit Sp1 to replace EKLK missing in the EKLK-null mutant background (Fig. 5A).

It is formally possible that the introduction of the E + CAC site into the  $\epsilon$ -globin promoter simultaneously or independently disrupted a negative regulatory element. Filipe et al. (1999) recently noted the two sets of DR sequences in the  $\epsilon$ -proximal promoter region (AGGTCANAGGTCA, see Fig. 2B) and that these sequences are consensus binding sites for nuclear hormone receptors (Glass 1994). These authors showed that the orphan nuclear receptor COUP-TFII binds to the DR sequences *in vitro*. Furthermore, COUP-TFII has been shown to act both as an active or passive repressor (Cooney et al. 1992; Cooney et al. 1993; Achatz et al. 1997). In competitive EMSA experiments (Fig. 8B and data not shown), we found a decrease in binding affinity of the E + CAC fragment for COUP-TFII compared with the wild-type  $\epsilon$  fragment (Fig. 8B, lanes 7,8), although COUP-TFII can still bind to the E + CAC fragment through a second DR element (Fig. 8B, lanes 3,4). Thus, it seems possible that the activation of the E + CAC promoter mutant may be caused by preventing the binding of a repressor. However, when Filipe et al. (1999) removed both DR sites from the  $\epsilon$  promoter and examined the mutation in transgenic mice, no increase in expression was observed in primitive erythroid cells, and thus these sites are not functioning as silencers at the primitive stage. Furthermore, COUP-TFII binding cannot explain E + CAC-mediated  $\epsilon$ -globin gene increases during definitive erythropoiesis, as COUP-TFII is not expressed in definitive cells

(Filipe et al. 1999). In EMSA, we identified a novel factor, DRED, that is expressed in definitive erythroid cells and found that the E + CAC mutation also partially disrupted the binding of this factor to the promoter region (Fig. 8D, lane 4). Thus, it is possible that DRED or a related protein is an erythroid stage-specific repressor and may also contribute to E + CAC promoter activation uniquely during definitive erythropoiesis.

Because EKLF cannot be recruited to the E + CAC promoter and the level of E + CAC mutant expression was still low (in comparison to  $\beta$ ) in definitive cells, we introduced a second mutation into the  $\epsilon$ -globin promoter that more closely mimics the context of transcription factor binding sites in the adult  $\beta$ -globin gene promoter. We mutated four additional nucleotides around the CAAT motif, while retaining those already introduced in the E + CAC mutant (to generate Bepsi; Fig. 2B). This second alteration completely eliminated the binding of both COUP-TFII and DRED to the Bepsi mutant  $\epsilon$  promoter in vitro (Fig. 8A,B, lanes 5, and D). The Bepsi fragment is also able to compete for factors that bind to the  $\beta$ -globin gene promoter (Fig. 7B,C, lanes 5, and D, lane 6). When analyzed in vivo, we found no difference in primitive stage  $\epsilon$ -globin transcription in single copy E + CAC or Bepsi mutant transgenic lines (Fig. 6A). Again, this result is inconsistent with the hypothesis that the relief of repression is responsible for elevated transcription of the two mutant promoters at the primitive stage, because the two mutants have dramatically different affinities for COUP-TFII (Fig. 8B, lanes 4,5).

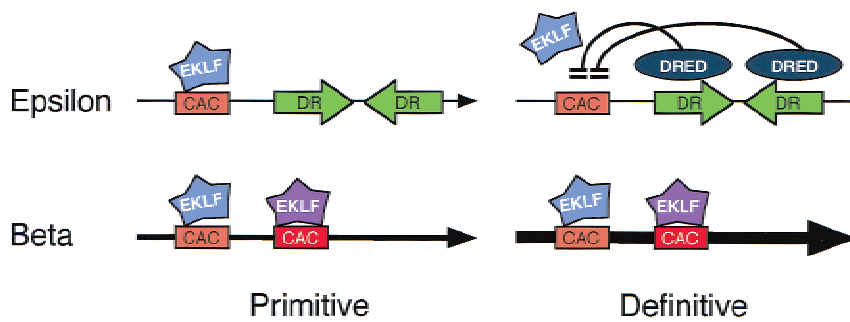
During definitive erythropoiesis, expression of  $\epsilon$ -globin in the Bepsi mutant was much higher than that in the E + CAC mutant, initially indicating that DNA sequences regulating the developmental stage-specific expression of the adult  $\beta$ -globin gene are located around the CAAT box. Because Bepsi, but not the E + CAC mutation, completely abolished DRED binding, we suggest that DRED is involved in definitive-stage-specific suppression of wild-type  $\epsilon$ -globin gene transcription. Surprisingly, when the Bepsi mutation was analyzed in an EKLF-null mutant background, its activity was severely attenuated, indicating that EKLF now activates this promoter in definitive cells. Importantly, this EKLF-dependent activation was only observed in definitive erythroid cells. We interpret these data to mean that a definitive-stage-specific repressor, DRED, prevents the binding of EKLF to the CAC sites in the  $\epsilon$ - and E + CAC promoters, and consequently this leads to suppression of  $\epsilon$ -globin transcription at definitive stages. In the E + CAC mutant, one DR site (around the CAAT box) appears to be enough to prevent EKLF recruitment to the neighboring (newly created) CAC site, although the presence of the single DR element is not sufficient to completely silence  $\epsilon$  transcription in definitive cells.

In competitive EMSA experiments (data not shown) we found that DRED binds with different affinities to the various globin CAAT box regions (in the order:  $\epsilon$  CAAT >  $\gamma$ -globin CAAT > HPFH mutant  $\gamma$ -globin CAAT >> adult  $\beta$ -globin CAAT, which was undetectable). Filipe et al. (1999) reported identical binding char-

acteristics for COUP-TFII. Based on these results, we propose that DRED is also involved in the silencing of  $\gamma$ -globin transcription during definitive erythropoiesis, because DRED is expressed in definitive cells, whereas COUP-TFII is not. Both the E + CAC mutant and wild-type  $\gamma$ -globin genes bear single DR sequences around the CAAT box and their expression profiles during murine development are very similar (high at primitive and very low in the definitive stages). Taken together, we further suggest that the elevated levels of  $\gamma$ -globin mRNA in individuals bearing CAAT mutations (Gelinas et al. 1985) may be the result of reduced DRED affinity for the HPFH mutant promoters.

Lee et al. (2000) recently reported characterization of a protein from MEL cell extracts that interacts with the CCTTG motif located in the  $\gamma$ -globin proximal promoter region. This sequence, when introduced into the corresponding position in the  $\beta$ -globin promoter, prevented the recruitment of EKLF to the proximal CAC box, initially suggesting that this factor may be related to DRED, as the  $\epsilon$ -promoter also contains a CCTTG motif in the immediate vicinity of the CAAT box (Fig. 2B). The results presented here, however, suggest that the CCTTG motif does not represent the complete binding site for DRED, because even though this sequence exists in the  $\epsilon$ -promoter, two subfragments (designated Epsi proxi. and E + CAC proxi.; Fig. 8A) containing the CCTTG motifs interacted with DRED with significantly different affinities (Fig. 8D, lanes 9,10). Although we could not, by EMSA, detect a definitive-stage erythroid factor whose specificity fulfills that anticipated for a CCTTG-specific binding protein, we can not exclude the possibility that the CCTTG motif (and its binding protein) may indeed constitute some part of the silencing machinery as the Bepsi (but not E + CAC) mutation disrupts this motif (Fig. 2B).

Dillon et al. (1997) replaced the  $\epsilon$  gene with an additional copy of a marked  $\beta$ -globin gene ( $\beta$ m) in linked cosmids and analyzed gene expression from this construct in transgenic mice. They reported that the proximal  $\beta$ -globin gene was expressed at much higher levels than the distal  $\beta$ -globin gene (>90% of total  $\beta$ -globin mRNA was generated by the LCR-proximal gene), and they concluded that the position of a gene relative to the LCR is an important factor for selective gene expression in the globin locus. Because the E + CAC or Bepsi mutant  $\epsilon$ -globin genes were expressed from an identical LCR-proximal position at only low levels compared with a misplaced  $\beta$ -globin gene during definitive erythropoiesis, we assume that the two mutant  $\epsilon$ -promoters examined in this study are still missing important information required for their high level expression at the definitive stage. One possible candidate for the missing sequence information is a binding motif for transcription factor CP-1/NFY, which is present in the proximal promoter of the  $\beta$ -globin gene but not in the mutant or wild-type  $\epsilon$ -globin promoters. An alternative explanation is that because the adult  $\beta$ -globin gene is also regulated by a 3' enhancer element (Behringer et al. 1987; Kollias et al. 1987; Liu et al. 1997), the possibility exists



**Figure 9.** Schematic model of  $\beta$ -globin gene switching. At the primitive stage, the direct repeat (DR) sequences (shown as green inverted arrows) on the  $\epsilon$ -globin promoter are inactive as silencing elements, and therefore EKLK can interact with the distal CAC site and contribute to promoter activation (*top left*). The  $\beta$ -globin promoter is also potentially active at this stage and this activity is completely dependent on EKLK (*bottom left*). In a situation where multiple genes are potentially active, the genes that are closest to the LCR would be

activated because of gene competition (Tanimoto et al. 1999). In definitive erythroid cells, DRED repressor protein binding to the DR sites prevents EKLK (and other activating proteins) from binding to the distal CAC site of the  $\epsilon$ -globin promoter (*top right*) and this gene is autonomously silenced. However, the  $\beta$ -globin gene is active because it does not have DRED-binding sites.

that this additional stimulatory activity is simply missing in these experiments. It is also quite possible that the  $\epsilon$ -globin upstream promoter region harbors other negative regulatory elements in addition to the promoter proximal negative elements defined here (Raich et al. 1992, 1993; Peters et al. 1993; J. Li et al. 1998; Q. Li et al. 1998).

In addition to transcription factor activity per se, differential stage-specific chromatin condensation could also contribute to impairment of  $\epsilon$ -globin transcription during the definitive stage (Engel and Tanimoto 2000; Gribnau et al. 2000). Because the activity of the Bepsi promoter is far less EKLK-dependent than is the adult  $\beta$ -globin promoter, there may be an additional binding site in the  $\beta$ -globin gene for EKLK (Guy et al. 2000) and this factor may exert an effect as a modifier of local chromatin structure rather than, or in addition to, its role as a conventional transcriptional regulatory molecule. The observation that EKLK associates with the chromatin remodeling complex E-RC1 is consistent with this possibility (Armstrong et al. 1998). Lee et al. (2000) recently showed that EKLK is recruited to the proximal CAC region of the  $\beta$ -globin promoter in transfected tissue culture cells. From these and other experiments, it is widely accepted that the primary target of EKLK is the proximal CAC box in the  $\beta$ -globin promoter. However, the reduction in  $\beta$ -globin mRNA accumulation in thalassemia patients bearing proximal CAC sequence point mutations (Orkin et al. 1982, 1984) is far less severe than the consequences of EKLK loss (Fig. 1A; Perkins et al. 1996; Wijgerde et al. 1996). Therefore decreased transcription in thalassemic patients may reflect diminished EKLK affinity for the point-mutant CAC site, rather than the complete loss of its activity as in the murine EKLK germ-line mutant.

In summary (Fig. 9), the wild-type  $\epsilon$ -globin gene promoter bears two intact DR sequences that can potentially bind to COUP-TFII in primitive erythroid cells (Fig. 4B). However, these sites are not functional as silencer elements at the primitive stage (Filipe et al. 1999, no. 2976). Hence, EKLK binding (to the natural distal CAC site) and other proteins activate this promoter and confer EKLK-dependent expression to this gene (Fig. 1D,E). In definitive erythroid cells, a putative repressor protein,

DRED, can bind to the DR sequences and interfere with EKLK binding. Consequently, the  $\epsilon$ -globin gene is silenced in definitive-stage cells. The  $\beta$ -globin gene promoter, in contrast, does not bear DR sites and is therefore not subject to silencing by any definitive stage-specific repressor. It is therefore apparent that the context of specific transcription factor binding sites in their promoters determine the stage-specific expression of the  $\epsilon$ - and  $\beta$ -globin genes. In the  $\epsilon$ -globin promoter, this process normally requires the participation of both positive and negative factors. We continue to focus efforts on identifying key switching regulators, foremost among them the DRED protein, that appears to act in concert with EKLK as a principal mediator of globin gene switching.

## Materials and methods

### Targeting vectors and YAC modification

For the introduction of the CAC site into the proximal position of the  $\epsilon$ -globin promoter, an  $\epsilon$ -gene restriction fragment (*SphI*-*BglIII*, nucleotide position from 18,050 to 20,064; <http://globin.cse.psu.edu>) was cloned into the yeast integrative plasmid pRS306 (Sikorski and Hieter 1989) and PCR-directed mutagenesis was performed using the following oligonucleotides (mutated nucleotides are underlined).

EP-S, 5'-CTGAGGCCACACCCTAGCCTTGACCAATGAC TTTTA-3'; EP-A, 5'-TAAAAGTCATTGGTCAAGGCTAGGG TGTGGCCTCAG-3'. For the introduction of the  $\beta$ -type CAAT box into the  $\epsilon$ -promoter region, PCR-directed mutagenesis was performed using the following oligonucleotides: EPBT-S, 5'-CTGAGGCCACACCCTAGCGTTGGCCAATCTCTTTTA-3'; EPBT-A, 5'-TAAAAGAGATTGGCCAACGCTAGGGTGTGG CCTCAG-3'. The plasmids were then linearized with *EcoNI* (at position 19,623) and used for homologous recombination in yeast to modify the  $\beta$ -globin YAC as described (Bungert et al. 1995).

### Transgenic mice

Purified YAC DNA was injected into fertilized mouse oocytes (CD1; Charles River Breeding Laboratory) before transfer to foster mothers (CD1) as described (Hogan et al. 1986). Tail DNA from offspring was initially analyzed for the presence of the L and R YAC vector arms by PCR and then for copy number and integrity of the transgenic  $\beta$ -globin locus by Southern blots of DNA prepared from transgenic tails or thymi.

*Preparation of agarose plugs for PFG analysis*

Cells ( $2 \times 10^7$ ) from the thymus (from <2-month-old animals) were suspended in 0.8 mL of cell suspension buffer (CSB; 10 mM Tris-HCl at pH 7.4, 20 mM NaCl, 50 mM EDTA). This solution was then mixed with 2% low-melting-point (lmp) agarose solution in CSB and poured into clamped homogeneous electric field (CHEF) plug molds. Solidified plugs were treated at 50°C in Proteinase K buffer (100 mM EDTA at pH 8.0, 0.2% sodium deoxycholate, 1% Sarkosyl, 10 mg/mL Proteinase K) for 48 h, followed by four washes with SE solution (75 mM NaCl, 25 mM EDTA) for 1 h and two washes with TE buffer (10 mM Tris-HCl at pH 7.4, 1 mM EDTA) for 1 h.

*Semiquantitative RT-PCR analysis of globin gene expression*

Total RNA from a minimum of two animals from each of the transgenic lines was extracted from yolk sac (9.5 dpc), fetal liver (14.5 dpc), or anemic adult spleens using ISOGEN (Nippon Gene). First-strand cDNA was synthesized with Moloney murine reverse transcriptase (GIBCO BRL) using 2.5 µg of total RNA in a 20-µL reaction volume. PCR was performed with 0.4 µL of cDNA in a 10-µL reaction volume (200 µM each dNTP, 40 ng of each gene-specific primer set, 0.2 units of AmpliTaq, 0.5 µCi of [ $\alpha$ - $^{32}$ P]dCTP, 1× Perkin-Elmer PCR buffer (2.5 mM MgCl<sub>2</sub>) at 94°C for 30 sec, 58°C for 1 min, and 72°C for 1 min. An aliquot of each PCR reaction was electrophoresed on 8% polyacrylamide gels, dried, and subjected to autoradiography. The PCR primers used have been described (Tanimoto et al. 1999b).

*EKLF-null mutant mice and genotyping of the EKLF locus*

Generation of EKLF-null mutant mice was described previously (Nuez et al. 1995). Genotyping of the endogenous EKLF locus was performed by PCR analysis of genomic DNA from embryos. The sequences of the primers used are: EKLF-5S, 5'-AGGCCTGAGGTCCAGGT-3'; EKLF-3A, 5'-GAGAAATAGGCCTCA CCTTG-3'; NEO-3', 5'-TTCGACGCGCATCGCCTTCTATC GC-3'. For detecting wild-type and mutant EKLF loci, a combination of EKLF-5S and 3A or EKLF-3A and Neo-3', was used, respectively.

*EMSA assay*

Nuclear extracts were prepared from mouse erythroleukemia (MEL) or K562 cells. Cells ( $5 \times 10^7$ ) were washed with PBS (-), pelleted (1500g, 3 min, 4°C), and resuspended in cell lysis buffer (20 mM HEPES at pH 7.9, 20% glycerol, 10 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 0.1% Triton-X, 1 mM DTT, 1 mM PMSF). After incubation on ice for 3 min, cells were pelleted (300g, 5 min, 4°C), resuspended in nuclear extraction buffer (cell lysis buffer minus Triton-X with 400 mM NaCl), and gently rocked for 1 h at 4°C. The supernatant was separated from debris by centrifugation (7000g, 10 min, 4°C) and dialyzed against dialysis buffer (20 mM HEPES at pH 7.9, 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 1 mM DTT, 1 mM PMSF) for 6 h to overnight at 4°C.

For specific experiments (those shown in Figs. 7B and 8B,C) we used the following protocol to avoid loss of COUP-TFII-activity (which seemed to diminish after dialysis). Cells ( $2.5 \times 10^7$ ) were washed with PBS(-), pelleted, and resuspended in 5× cell volume of Buffer A (10 mM HEPES at pH 7.9, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.5 mM DTT, 0.5 mM PMSF). After 4°C incubation for 10 min, cells were pelleted, resuspended in 3× cell volume of Buffer A containing 0.05% (v/v final) of NP-40, and dounced 20 times with a ho-

mogenizer (2 mL, Wheaton). The supernatant was removed after centrifugation (800g, 5 min, 4°C) and nuclei were resuspended in 5× cell volume of Buffer C (5 mM HEPES at pH 7.9, 26% glycerol, 1.5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, 0.5 mM DTT, 0.5 mM PMSF). NaCl was slowly added to a final concentration of 0.3 M (0.06 volume of 5 M NaCl) and gently rocked for 1 h at 4°C. The supernatant was separated from debris by centrifugation (9000g, 10 min, 4°C) and used for binding reactions.

Recombinant EKLF protein was expressed in and purified from *E. coli* BL21 cells as a GST-fusion protein as described (Bieker and Southwood 1995). The fusion protein was subsequently affinity-purified using glutathione beads (Sigma).

Binding reactions and electrophoresis were performed as described (Tanimoto et al. 1996). For the reaction with nuclear extracts prepared by the second protocol, the binding buffer (20 mM HEPES at pH 7.9, 1 mM MgCl<sub>2</sub>, 4% Ficoll, 0.5 mM DTT) was supplemented with KCl so that the final monovalent cation concentration was 50 mM.

*Oligonucleotide sequences and antibodies*

The following oligonucleotides (only sense strands are shown) were used as probes or competitors in EMSA experiments: Epsi, 5'-CTGAGGACACAGGTCAGCCTTGACCAATGACTTTTA-3'; E + CAC, 5'-CTGAGGCCACACCCTAGCCTTGACCAATGACTTTTA-3'; Bepsi, 5'-CTGAGGCCACACCCTAGCGTTG GCCAATCTCTTTTA-3'; Beta, 5'-GTGGAGCCACACCCTA GGGTTGGCCAATCTACTCCC-3'; Epsi distal, 5'-AGCTGAC TCCACCCTGAGCT-3'; Beta distal, 5'-AGCTGACCTACC CTGTAGCT-3'; Epsi proxi, 5'-CCCTGAGGACACAGGTCAG CCTTG-3'; E + CAC proxi, 5'-CCCTGAGGCCACACCCT AGCCTTG-3'; Beta proxi, 5'-CTGTGGAGCCACACCCTAG GGTTG-3'; Epsi CAAT, 5'-CAGCCTTGACCAATGACT TTTA-3'; Beta CAAT, 5'-TAGGGTTGGCCAATCTACTCCC-3'; Gamma CAAT, 5'-TTGCCTTGACCAATGACTCTGA-3'; EKLE, 5'-AGCTAGCCACACCCTAGAGCT-3'; SP-1, 5'-TGAA AAGGCGGTCTCCA-3'; RARE, 5'-TTGCTGTGACCTCT GCCCTTCTAGCCTCT-3'.

Double-stranded probe DNAs were end-labeled with [ $\gamma$ - $^{32}$ P]ATP, column-purified, and used for binding reactions.

Antibodies against EKLF (marked JB) were provided by James Bieker (Mt. Sinai, New York, NY). All other antibodies used were from SantaCruz Biotechnology; sc-7714x (CP-1/NF-Y), sc-6578x (COUP-TFII), sc-59x (SP-1; not cross-reactive with SP-2, SP-3, or SP-4), sc-1905x (EKLF; cross-reactive with multiple krüppel-like factors), sc-773x (RAR; cross-reactive with  $\alpha$ ,  $\beta$  and  $\gamma$  subunits), sc-774x (RXR; cross-reactive with  $\alpha$ ,  $\beta$  and  $\gamma$  subunits), sc-6062x (ROR), sc-1204x (FXR), sc-841x (DAX-1), sc-1201x (LXR) and sc-1985x, sc7197x, sc-1984x for PPAR  $\alpha$ ,  $\beta$ ,  $\gamma$ , respectively.

**Acknowledgments**

We are very grateful to J. Bieker for the EKLF plasmids, antisera, and advice, and to M.-J. Tsai for helpful discussions about COUP-TFII. We thank Y. Tanimoto and W. Song for outstanding technical assistance, and our colleagues K.-C. Lim, A. Ruiz-Rivera, and Y. Zhou for discussions. We also thank the NIH for research support (HL24415; J.D.E.) and shared facilities support through a CCSG to the Robert H. Lurie Comprehensive Cancer Center of Northwestern University (CA60553). K.T. was supported by fellowships from the Cooley's Anemia Foundation.

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

## References

- Achatz, G., Holzl, B., Speckmayer, R., Hauser, C., Sandhofer, F., and Paulweber, B. 1997. Functional domains of the human orphan receptor ARP-1/COUP-TFII involved in active repression and transrepression. *Mol. Cell. Biol.* **17**: 4914–4932.
- Anderson, K.P., Kern, C.B., Crable, S.C., and Lingrel, J.B. 1995. Isolation of a gene encoding a functional zinc finger protein homologous to erythroid Kruppel-like factor: Identification of a new multigene family. *Mol. Cell. Biol.* **15**: 5957–5965.
- Armstrong, J., Bieker, J.J., and Emerson, B.M. 1998. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* **95**: 93–104.
- Asano, H. and Stamatoyannopoulos, G. 1998. Activation of  $\beta$ -globin promoter by erythroid Kruppel-like factor. *Mol. Cell. Biol.* **18**: 102–109.
- Asano, H., Li, X.S., and Stamatoyannopoulos, G. 1999. FKLF, a novel Kruppel-like factor that activates human embryonic and fetal  $\beta$ -like globin genes. *Mol. Cell. Biol.* **19**: 3571–3579.
- Behringer, R.R., Hammer, R.E., Brinster, R.L., Palmiter, R.D., and Townes, T.M. 1987. Two 3' sequences direct adult erythroid-specific expression of human  $\beta$ -globin genes in transgenic mice. *Proc. Natl. Acad. Sci.* **84**: 7056–7060.
- Behringer, R.R., Ryan, T.M., Palmiter, R.D., Brinster, R.L., and Townes, T.M. 1990. Human  $\gamma$ - to  $\beta$ -globin gene switching in transgenic mice. *Genes & Dev.* **4**: 380–389.
- Bieker, J.J. and Southwood, C.M. 1995. The erythroid kruppel-like factor transactivation domain is a component for cell-specific inducibility of a  $\beta$ -globin promoter. *Mol. Cell. Biol.* **15**: 852–860.
- Bungert, J., Dave, U., Lim, K.C., Lieu, K.H., Shavit, J.A., Liu, Q., and Engel, J.D. 1995. Synergistic regulation of human  $\beta$ -globin gene switching by locus control region elements HS3 and HS4. *Genes & Dev.* **9**: 3083–3096.
- Chada, K., Magram, J., and Costantini, F. 1986. An embryonic pattern of expression of a human fetal globin gene in transgenic mice. *Nature* **319**: 685–686.
- Choi, O. and Engel, J.D. 1988. Developmental regulation of  $\beta$ -globin gene switching. *Cell* **55**: 17–26.
- Cooney, A.J., Tsai, S.Y., O'Malley, B.W., and Tsai, M.J. 1992. Chicken ovalbumin upstream promoter transcription factor (COUP-TF) dimers bind to different GGTC A response elements, allowing COUP-TF to repress hormonal induction of the vitamin D<sub>3</sub> thyroid hormone, and retinoic acid receptors. *Mol. Cell. Biol.* **12**: 4153–4163.
- Cooney, A.J., Leng, X., Tsai, S.Y., O'Malley, B.W., and Tsai, M.J. 1993. Multiple mechanisms of chicken ovalbumin upstream promoter transcription factor-dependent repression of transactivation by the vitamin D, thyroid hormone, and retinoic acid receptors. *J. Biol. Chem.* **268**: 4152–4160.
- Crossley, M., Whitelaw, E., Perkins, A., Williams, G., Fujiwara, Y., and Orkin, S.H. 1996. Isolation and characterization of the cDNA encoding BKLF/TEF-2, a major CACCC-box-binding protein in erythroid cells and selected other cells. *Mol. Cell. Biol.* **16**: 1695–1705.
- Dillon, N. and Grosveld, F. 1991. Human gamma-globin genes silenced independently of other genes in the  $\beta$ -globin locus. *Nature* **350**: 252–254.
- Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P., and Grosveld, F. 1997. The effect of distance on long-range chromatin interactions. *Mol. Cell* **1**: 131–139.
- Donze, D., Townes, T.M., and Bieker, J.J. 1995. Role of erythroid Kruppel-like factor in human  $\gamma$ - to  $\beta$ -globin gene switching. *J. Biol. Chem.* **270**: 1955–1959.
- Engel, J.D. and Tanimoto, K. 2000. Looping, linking, and chromatin activity: New insights into  $\beta$ -globin locus regulation. *Cell* **100**: 499–502.
- Enver, T., Ebens, A.J., Forrester, W.C., and Stamatoyannopoulos, G. 1989. The human  $\beta$ -globin locus activation region alters the developmental fate of a human fetal globin gene in transgenic mice. *Proc. Natl. Acad. Sci.* **86**: 7033–7037.
- Enver, T., Raich, N., Ebens, A.J., Papayannopoulou, T., Costantini, F., and Stamatoyannopoulos, G. 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* **344**: 309–313.
- Feng, W.C., Southwoods, C.W., and Bieker, J.J. 1994. Analyses of  $\beta$ -thalassemia DNA interactions with erythroid kruppel-like factor, an erythroid cell-specific transcription factor. *J. Biol. Chem.* **269**: 1–8.
- Filipe, A., Li, Q., Deveaux, S., Godin, I., Romeo, P.H., Stamatoyannopoulos, G., and Mignotte, V. 1999. Regulation of embryonic/fetal globin genes by nuclear hormone receptors: A novel perspective on hemoglobin switching. *EMBO J.* **18**: 687–697.
- Forrester, W.C., Takegawa, S., Papayannopoulou, T., Stamatoyannopoulos, G., and Groudine, M. 1987. Evidence for a locus activation region: The formation of developmentally stable hypersensitive sites in globin-expressing hybrids. *Nucleic Acids Res.* **15**: 10159–10177.
- Gaensler, K.M.L., Burmeister, M., Brownstein, B.H., Taillon-Miller, P., and Myers, R.M. 1991. Physical mapping of yeast artificial chromosomes containing sequences from the human  $\beta$ -globin gene region. *Genomics* **10**: 976–984.
- Gelinas, R., Endlich, B., Pfeiffer, C., Yagi, M., and Stamatoyannopoulos, G. 1985. G to A substitution in the distal CCAAT box of the A  $\gamma$ -globin gene in Greek hereditary persistence of fetal haemoglobin. *Nature* **313**: 323–325.
- Gillems, N., Tewari, R., Lindeboom, F., Rottier, R., de Wit, T., Wijgerde, M., Grosveld, F., and Philipsen, S. 1998. Altered DNA-binding specificity mutants of EKLF and Sp1 show that EKLF is an activator of the  $\beta$ -globin locus control region in vivo. *Genes & Dev.* **12**: 2863–2873.
- Glass, C.K. 1994. Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. *Endocr. Rev.* **15**: 391–407.
- Gribnau, J., Diderich, K., Pruzina, S., Calzolari, R., and Fraser, P. 2000. Intergenic transcription and developmental remodeling of chromatin subdomains in the human  $\beta$ -globin locus. *Mol. Cell* **5**: 377–388.
- Grosveld, F., van Assendelft, G.B., Greaves, D.R., and Kollias, G. 1987. Position-independent, high-level expression of the human  $\beta$ -globin gene in transgenic mice. *Cell* **51**: 975–985.
- Gumucio, D.L., Rood, K.L., Blanchard-McQuate, K.L., Gray, T.A., Saulino, A., and Collins, F.S. 1991. Interaction of Sp1 with the human gamma globin promoter: Binding and transactivation of normal and mutant promoters. *Blood* **78**: 1853–1863.
- Guy, L.G., Mei, Q., Perkins, A.C., Orkin, S.H., and Wall, L. 1998. Erythroid Kruppel-like factor is essential for  $\beta$ -globin gene expression even in absence of gene competition, but is not sufficient to induce the switch from  $\gamma$ -globin to  $\beta$ -globin gene expression. *Blood* **91**: 2259–2263.
- Guy, L.G., Delvoe, N., and Wall, L. 2000. Expression of a human  $\beta$ -globin transgene in mice with the CACC motif and upstream sequences deleted from the promoter still depends on erythroid Kruppel-like factor. *J. Biol. Chem.* **275**: 3675–3680.
- Hartzog, G.A. and Myers, R.M. 1993. Discrimination among potential activators of the  $\beta$ -globin CACC element by correlation of binding and transcriptional properties. *Mol. Cell.*



- Biol.* **13**: 44–56.
- Hogan, B., Costantini, F., and Lacy, E. 1986. *Manipulating the mouse embryo*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Kmita, M., van Der Hoeven, F., Zakany, J., Krumlauf, R., and Duboule, D. 2000. Mechanisms of Hox gene colinearity: Transposition of the anterior Hoxb1 gene into the posterior HoxD complex. *Genes & Dev.* **14**: 198–211.
- Kollias, G., Wrighton, N., Hurst, J., and Grosveld, F. 1986. Regulated expression of human  $\gamma$ -,  $\beta$ -, and hybrid  $\gamma\beta$ -globin genes in transgenic mice: Manipulation of the developmental expression patterns. *Cell* **46**: 89–94.
- Kollias, G., Hurst, J., deBoer, E., and Grosveld, F. 1987. The human  $\beta$ -globin gene contains a downstream developmental specific enhancer. *Nucleic Acids Res.* **15**: 5739–5747.
- Lee, J.-S., Ngo, H., Kim, D., and Chung, J.H. 2000. Erythroid Kruppel-like factor is recruited to the CACCC box in the  $\beta$ -globin promoter but not to the CACCC box in the  $\gamma$ -globin promoter: The role of the neighboring promoter elements. *Proc. Natl. Acad. Sci.* **97**: 2468–2473.
- Li, J., Noguchi, C.T., Miller, W., Hardison, R., and Schechter, A.N. 1998. Multiple regulatory elements in the 5'-flanking sequence of the human  $\epsilon$ -globin gene. *J. Biol. Chem.* **273**: 10202–10209.
- Li, Q., Blau, C.A., Clegg, C.H., Rohde, A., and Stamatoyannopoulos, G. 1998. Multiple  $\epsilon$ -promoter elements participate in the developmental control of  $\epsilon$ -globin genes in transgenic mice. *J. Biol. Chem.* **273**: 17361–17367.
- Liu, Q., Bungert, J., and Engel, J.D. 1997. Mutation of gene-proximal elements disrupts human  $\epsilon$ -,  $\gamma$ -, and  $\beta$ -globin expression in YAC transgenic mice. *Proc. Natl. Acad. Sci.* **94**: 169–174.
- Liu, Q., Bungert, J., Tanimoto, K., and Engel, J.D. 1998. The A $\gamma$ -globin 3' element provides no unique function(s) for human  $\beta$ -globin locus gene regulation. *Proc. Natl. Acad. Sci.* **95**: 9944–9949.
- Marin, M., Karis, A., Visser, P., Grosveld, F., and Philipsen, S. 1997. Transcription factor Sp1 is essential for early development but dispensable for cell growth and differentiation. *Cell* **89**: 619–628.
- Matsumoto, N., Laub, F., Aldabe, R., Zhang, W., Ramirez, F., Yoshida, T., and Terada, M. 1998. Cloning the cDNA for a new human zinc finger protein defines a group of closely related Kruppel-like transcription factors. *J. Biol. Chem.* **273**: 28229–28237.
- Miller, I.J. and Bieker, J.J. 1993. A novel, erythroid cell-specific transcription factor that binds to the CACC element related to the kruppel family of nuclear proteins. *Mol. Cell. Biol.* **13**: 2776–2786.
- Myers, R.M., Tilly, K., and Maniatis, T. 1986. Fine structure genetic analysis of a  $\beta$ -globin promoter. *Science* **232**: 613–618.
- 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.
- Ohtsuki, S. and Levine, M. 1998. GAGA mediates the enhancer blocking activity of the eve promoter in the *Drosophila* embryo. *Genes & Dev.* **12**: 3325–3330.
- Orkin, S.H., Kazazian, Jr., H.H., Antonarakis, S.E., Goff, S.C., Boehm, C.D., Sexton, J.P., Waber, P.G., and Giardina, P.J. 1982. Linkage of  $\beta$ -thalassaemia mutations and  $\beta$ -globin gene polymorphisms with DNA polymorphisms in human  $\beta$ -globin gene cluster. *Nature* **296**: 627–631.
- Orkin, S.H., Antonarakis, S.E., and Kazazian, Jr., H.H. 1984. Base substitution at position -88 in a  $\beta$ -thalassemic globin gene. Further evidence for the role of distal promoter element ACACCC. *J. Biol. Chem.* **259**: 8679–8681.
- Perkins, A.C., Sharpe, A.H., and Orkin, S.H. 1995. Lethal  $\beta$ -thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* **375**: 318–322.
- Perkins, A.C., Gaensler, K.M., and Orkin, S.H. 1996. Silencing of human fetal globin expression is impaired in the absence of the adult  $\beta$ -globin gene activator protein EKLF. *Proc. Natl. Acad. Sci.* **93**: 12267–12271.
- Peters, B., Merezhinskaya, N., Diffley, J.F., and Noguchi, C.T. 1993. Protein-DNA interactions in the  $\epsilon$ -globin gene silencer. *J. Biol. Chem.* **268**: 3430–3437.
- Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papyannopoulou, T., and Stamatoyannopoulos, G. 1990. Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science* **250**: 1147–1149.
- Raich, N., Papyannopoulou, T., Stamatoyannopoulos, G., and Enver, T. 1992. Demonstration of a human  $\epsilon$ -globin gene silencer with studies in transgenic mice. *Blood* **79**: 861–864.
- Raich, N., Clegg, C.H., Grofti, J., Romeo, P.H., and Stamatoyannopoulos, G. 1995. GATA1 and YY1 are developmental repressors of the human  $\epsilon$ -globin gene. *EMBO J.* **14**: 801–809.
- Shields, J.M., Christy, R.J., and Yang, V.W. 1996. Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed in growth arrest. *J. Biol. Chem.* **271**: 20009–20017.
- Sikorski, R.S. and Hieter, P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**: 19–27.
- Southwood, C.M., Downs, K.M., and Bieker, J.J. 1996. Erythroid Kruppel-like factor exhibits an early and localized pattern of expression during erythroid ontogeny. *Dev. Dyn.* **206**: 248–259.
- Stamatoyannopoulos, G. and Neinhuis, A.W. 1994. Hemoglobin switching. In *The molecular basis of blood diseases*, 2nd ed. (ed. G. Stamatoyannopoulos, A.W. Nienhuis, P. Majerus, and H. Varmus), pp. 107–155. W.B. Saunders, New York.
- Tanimoto, K., Yoshida, E., Mita, S., Nibu, Y., Murakami, K., and Fukamizu, A. 1996. Human activin  $\beta$ A gene. Identification of novel 5' exon, functional promoter, and enhancers. *J. Biol. Chem.* **271**: 32760–32769.
- Tanimoto, K., Liu, Q., Bungert, J., and Engel, J.D. 1999a. Effects of altered gene order or orientation of the locus control region on human  $\beta$ -globin gene expression in mice. *Nature* **398**: 344–348.
- . 1999b. The polyoma virus enhancer cannot substitute for DNase I core HS sites 2–4 in the human  $\beta$ -globin LCR. *Nucleic Acids Res.* **27**: 3130–3137.
- Tewari, R., Gillemans, N., Wijgerde, M., Nuez, B., von Lindern, M., Grosveld, F., and Philipsen, S. 1998. Erythroid Kruppel-like factor (EKLF) is active in primitive and definitive erythroid cells and is required for the function of 5'HS3 of the  $\beta$ -globin locus control region. *EMBO J.* **17**: 2334–2341.
- Townes, T.M., Lingrel, J.B., Chen, H.Y., Brinster, R.L., and Palmiter, R.D. 1985. Erythroid-specific expression of human  $\beta$ -globin genes in transgenic mice. *EMBO J.* **4**: 1715–1723.
- Treisman, R., Orkin, S.H., and Maniatis, T. 1983. Specific transcription and RNA splicing defects in five cloned  $\beta$ -thalassaemia genes. *Nature* **302**: 591–596.
- Wijgerde, M., Gribnau, J., Trimborn, T., Nuez, B., Philipsen, S., Grosveld, F., and Fraser, P. 1996. The role of EKLF in human  $\beta$ -globin gene competition. *Genes & Dev.* **10**: 2894–2902.
- Yu, C.Y., Motamed, K., Chen, J., Bailey, A.D. and Shen, C.K. 1991. The CACC box upstream of human embryonic  $\epsilon$  globin gene binds Sp1 and is a functional promoter element in vitro and in vivo. *J. Biol. Chem.* **266**: 8907–8915.