

Probing the Onset and Regulation of Erythroid Cell-Specific Gene Expression

JAMES J. BIEKER, PH.D.

Abstract

Cell-restricted transcriptional modulators play a crucial role in the process of selective gene regulation during hematopoiesis. Erythroid Krüppel-like factor (EKLF; KLF1) is a transcription factor, originally identified in this laboratory, which plays a crucial role as a transcriptional activator at the adult β -globin locus. This review will describe a subset of experiments and approaches related to deciphering its mechanism of action and the means by which it itself is induced.

Key Words: Transcription, EKLF, erythroid cell, gene regulation.

Introduction

UNRAVELING THE MOLECULAR EVENTS that confer the ability to express lineage-specific genes upon an initially uncommitted, pluripotent hematopoietic stem cell remains a major unresolved issue in cell differentiation. This article will describe some of the experiments performed at my lab at the Mount Sinai School of Medicine that address this issue and will show how characterization of a cell-specific transcription factor has opened a window on the processes that lead to erythroid differentiation.

Erythroid Krüppel-like factor (EKLF) was originally isolated by a differential screening approach whereby mRNA that was specifically expressed in erythroid cells was identified after subtraction with common messages expressed in a myeloid cell line (1). Specifically, murine erythroleukemia (MEL) cells provided the erythroid mRNA sample, and the J774 monocyte-macrophage cell line provided the subtracting mRNA sample. This removed 99% of MEL mRNA from consideration. The remaining material was directly cloned. Most of these clones were related to Friend virus genes, metabolic clones (e.g.,

carbonic anhydrase), and transcription factors (e.g., c-myb and GATA1), but some clones were novel at the time. A Northern blot analysis of adult murine tissues directed the focus to one particular clone uniquely expressed in bone marrow and spleen. Further analysis of this clone (EKLF) indicated that it was expressed only in erythroid cell lines, but not in any lymphoid or myeloid lines, with the exception of a small amount of expression in mast cell lines.

A significant clue as to the function of this clone resulted from the completion of its primary sequence, which indicated that it contained three C2H2-type zinc fingers at the extreme carboxyl end, raising the possibility that it was a DNA-binding protein (1). More precisely, it was most closely homologous to the Krüppel family of transcription factors, which are named after the segmentation gene involved in *Drosophila* body patterning (2). This was fortuitous, as a variety of finger-swapping experiments had been performed with other members of this family (e.g., Zif 268, Sp1, Krox 20) with the aim of predicting its target DNA-binding site. The molecular analysis culminated in the crystallographic structure of Zif 268 bound to its target site (3). Of critical importance for predictive purposes were the interactions of basic residues, at specific amino acid locations, of Zif 268 with guanines within the DNA-binding sequence (4). In addition, each finger interacted with a triplet of nucleotides. This information enabled us to predict that EKLF would interact with 3'GGNGNGGGN5'.

Address all correspondence to James J. Bieker, Ph.D., Box 1020, Mount Sinai School of Medicine, One East 100th Street, New York, NY 10029.

Adapted from a Dean's Lecture presentation at the Mount Sinai School of Medicine, New York, NY on June 23, 2004, and updated as of March 2005.

Glossary

BAF = BRG associated factors
 BMP = bone morphogenetic protein
 BMPR = BMP receptor
 CBF = copper-binding protein
 CFU-E = colony-forming unit-erythroid
 DN = dominant negative
 EB = embryoid bodies
 E-RCI = EKLK remodeling complex-1

ES = embryonic stem
 EKLK = Erythroid Krüppel-like factor
 HS = hypersensitive site
 LCR = locus control region
 MEL = murine erythroleukemia
 P/CAF = p300/CBP associated factor
 SUMO = Small Ubiquitin-like Modifier

A subset of this sequence, known as the 5'CACCC3' element, was already recognized as one of a trio of elements that are repeated within erythroid enhancers and promoters. In particular, the CAC site within the β -globin gene (5'CCA-CACCCT3') precisely matched the predicted EKLK consensus binding sequence. An oligonucleotide containing this sequence was directly tested and found to bind EKLK *in vitro* with high affinity. In addition, transfection studies demonstrated that EKLK functioned as a transcriptional activator (not repressor) when bound to this site. These studies provided the initial isolation and characterization of this transcription factor (1).

Functional Role of EKLK in Erythroid Expression

The β -globin gene CACCC element is the site of thalassemia-causing point mutations in humans (5). Given this fact, individual examples of these elements were directly tested for their ability to interact with EKLK (6). The results were dramatic: Not only was EKLK unable to transactivate through these mutated sites, but also the binding affinity was decreased 40- to 100-fold. Molecular modeling of these interactions indicated that both specific and non-specific hydrogen bond interactions were adversely affected by what appeared to be minor changes in the sequence of the CACCC element. Methylation interference assays further demonstrated that each of the guanines on the G-rich strand in the nine-base-pair β -globin CACCC element (3'GGTGTGGGA5') were critical for optimal binding by EKLK.

To understand the potential importance of these properties, a brief summary of β -globin regulation is needed. Expression of the β -like globin genes is regulated in a specific temporal sequence during mammalian development (5, 7, 8). In humans, the first blood cells formed within the yolk sac express the embryonic (epsilon) globin variant. After the site of hematopoiesis switches to the fetal liver, the fetal (gamma) globin genes are expressed. This is followed by yet another switch in

expression to the adult (β) globin gene after the bone marrow becomes the major site of hematopoiesis. In mice, there is a single switch in molecular expression from embryonic (γ , bh0, bh1) globin genes in the yolk sac to the adult (β) globin variant in the fetal liver and the adult bone marrow and spleen. Regulation of this large β -like globin cluster is thought to occur by competition of each globin promoter member (embryonic, fetal, and adult) for interaction with the locus control region (LCR), thus enabling high-level activity of only one member at its appropriate time during development. In reality, this control is more complex, as an endogenous silencing mechanism also plays a vital role in this process (7–10).

Hints that EKLK's role in β -globin regulation is complex followed from the point mutation analysis described above. Specifically, the murine embryonic (and human fetal) CACCC elements differ in one of the important guanines within the interaction sequence, 3'GAGGTGGGT5' (11). Given the deleterious effects of the thalassemia-causing point mutations, it was felt that this change would also adversely affect binding. This hypothesis was directly tested, and competition assays demonstrated that the embryonic CACCC element had an 8-fold lower affinity for EKLK compared to the adult CACCC element (12). The effects of such a mutation were tested *in vivo* by using K562 cells, a human erythroleukemic cell line that expresses gamma, but not β , globin. In this line, EKLK had minimal effects on the activity of an intact gamma globin promoter, but could specifically activate the β -globin promoter more than 1000-fold when it was linked in *cis* to the gamma promoter. These molecular data raised the possibility that EKLK may be important for the switch from fetal to adult β -like globin expression and led to the prediction that an EKLK-null mutation would lead to deficiency in adult, but not in embryonic or fetal, globin expression, a conjecture that was borne out by its genetic knock-out in mice (13, 14). This led to an embryonic-lethal phenotype by day 14, due to lack of onset of β -globin transcription, resulting in

extreme anemia and prenatal death, precisely at the time of the switch; embryonic globins were expressed at normal levels. These effects were unusually specific, since alpha-globins and other red-cell-specific genes were unaffected by the lack of EKLf expression. Additional studies that utilized EKLf-null embryonic stem cells to generate chimeric adult mice indicated that these cells did not contribute to the mature erythrocyte population, and that they were defective by virtue of a globin chain imbalance, leading to Heinz body formation, a shortened lifespan, and apparent clearance from the bloodstream (15).

These data revealed that EKLf is critical for the onset of adult β -globin expression. To address the effects on the other linked globin-like genes in the cluster, EKLf-null mice were crossed with a murine line that carries the complete human β -like globin cluster (16, 17). As seen with the murine locus, adult β -globin expression was completely dependent on EKLf. However, gamma-globin levels were 5-fold higher, and although they were eventually silenced (as the mice still died from lack of adult β -globin expression), this shut-off occurred significantly later than that observed in the presence of EKLf. In addition, most EKLf-null, colony forming unit-erythroid (CFU-E) colonies were reprogrammed to high gamma-globin expression. The mode of EKLf action thus appears to involve consolidating the switch from fetal to adult globin, possibly by enabling the β -globin promoter to form a structure that is able to compete for interaction with the LCR to a greater extent than the embryonic- or fetal-globin promoters.

Mechanism of EKLf Transcriptional Activation

Two aspects of EKLf function were studied. One was the possible role that post-translational modification plays in EKLf activity, since the minimal activation and inhibitory domains each contain conserved sites for serine/threonine phosphorylation (18). In fact, one of the double-amino acid mutations (at aa 43–44) that disrupted the *in vivo* competitive ability of EKLf overlapped with a casein kinase II site (threonine at aa 41). As expected, immunoprecipitation of EKLf from erythroid cells revealed that EKLf is a phosphoprotein *in vivo*, exclusively at serine and threonine (no evidence was seen for modification of tyrosine). In addition, EKLf is a suitable substrate for casein kinase II *in vitro*. Further, the double mutant, or a direct mutant of thr41, was not able to transactivate a reporter *in vivo*. This loss was not observed when the charge at aa 41 was retained (i.e., after conversion to aspartic acid). The location of this

important phosphorylation site within the interaction domain of the minimal activation region raises the possibility that EKLf interactions and activity may be modulated by phosphorylation.

The second aspect of EKLf function that was considered in the deletional studies was the importance of protein-protein interactions. The possibility that EKLf may also be involved in chromatin structure at the β -globin locus led to directly testing whether it interacted with the class of molecules known as histone acetyltransferases (19). These are known to modify histones via acetylation of the amino terminal tails, leading to a more open chromatin structure at that region, thus enabling the transcription machinery to be fully recruited and active (20, 21). The finding that some transcriptional co-activators (such as CBP, p300, and P/CAF) harbor intrinsic histone acetyltransferase activity suggests a mechanism to directly link transcriptional activation and alteration of chromatin structure (22). In the present case (19), EKLf was found to associate with p300, CREB-binding protein (CBP), and p300/CBP associated factor (P/CAF) *in vivo*. Interestingly, EKLf is itself acetylated by p300 and CBP *in vitro*, but not by P/CAF. These sites have been mapped to two lysines (288 and 302). Mutation of lys-288 (but not lys-302) decreases EKLf transactivation of the β -globin promoter and disables its ability to be superactivated by p300 and CBP (23). EKLf interactions with proteins, rather than with DNA, are affected by its acetylation. In addition to these post-translational modifications (Fig. 1), EKLf is also SUMOylated (SUMO: Small Ubiquitin-like Modifier) (Siatecka and Bieker, unpublished) and ubiquitinated (Quadrini and Bieker, unpublished).

An additional area in which EKLf-protein interactions may be playing an important role is in chromatin remodeling (24, 25). Using an *in vitro* chromatin reconstitution/transcription system, EKLf and a MEL-derived cell extract (EKLf remodeling complex-1; E-RC1) were both

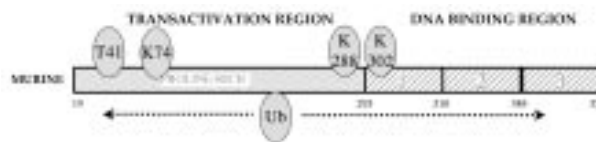


Fig 1. Post-translational modification of EKLf protein.

A schematic diagram of EKLf shows its transactivation (amino terminal) and DNA binding (carboxyl terminal) regions. The DNA binding region contains three C2H2 zinc fingers. Also indicated are the locations of important phosphorylation (T41), sumoylation (K74), acetylation (K288 and K302), and ubiquitination (throughout) sites discussed in the text.

required to provide an open chromatin structure and transcription at the β -globin promoter (26). In the absence of either component, or by using a mutated EKLK binding site, the chromatin structure remained closed and refractory to transcription. Use of an EKLK deletion that had removed its amino terminal interaction domain still bound DNA and yielded an open chromatin structure, but no transcription was observed, thus uncoupling the two events. Intriguingly, purification of E-RC1 helped identify mammalian components of the SWI-SNF chromatin remodeling proteins, including BRG1, BAF (BRG1 associated factor) 170, BAF155, BAF47, and BAF57. Immunoprecipitation or immunoneutralization of this complex from the MEL fraction again led to an inability to activate the promoter. Interaction of EKLK with BRG1 appears critical for interaction with SWI-SNF (27).

These studies, taken together, implicate EKLK as a central player in the integration of a variety of signals that lead to protein modification, protein-protein interactions, opening of chromatin, and activation of transcription. There is ample precedent for each of the steps in this pathway from other systems, and a testable challenge for future studies will be to determine just how EKLK functions. The biological relevance of these studies has followed from studies that crossed the EKLK-null mice with mice that are transgenic for the complete human β -like globin locus (17). These studies demonstrated that the hypersensitive site (HS) at the β -globin promoter is absent when EKLK is not present. In addition, HS3 within the LCR was adversely affected. Thus, not only is EKLK required for transcription of β -globin *in vivo*, but it also plays a role in forming the correct higher-order structure at the globin locus (28). EKLK interactions with histone acetyltransferases and with chromatin remodeling proteins provide a mechanism that partially accounts for how this may occur.

Regulation of EKLK Expression

Analysis of EKLK biological expression has also provided critical information about its role (29). EKLK is not expressed in embryonic stem (ES) cells, but becomes apparent during embryoid body differentiation in culture. Analysis of developing murine embryos indicated that EKLK is first expressed at the neural plate stage (i.e., day 7.5) within the extraembryonic mesoderm of the visceral yolk sac, the earliest morphologically identifiable time of blood island formation, after which its level continues to expand. It is then expressed at day 9.0 within the hepatic primordia, and by day 14.5 the fetal liver is the only site of EKLK expression. In the adult mouse spleen, it is expressed solely within the red pulp.

Because of these erythroid cell-specific properties, EKLK's own regulation has been investigated (30–32). Both the murine and human genomic clones have been isolated, sequenced, and mapped (33, 34). A surprisingly short region of the promoter is sufficient to reconstitute erythroid specificity to an adjacent reporter (30). Although there is no TATA sequence, the proximal promoter region contains GATA and CCAAT sites. The importance of these sites was verified by transfection assays in erythroid cells (MEL and 32DEpo1 lines). More distal elements have been implicated by identifying erythroid-specific DNase hypersensitive sites at the EKLK genomic locus (31). Two sites within 1000 bp 5' of the transcription start site were identified by this approach. The sequence surrounding the more distal hypersensitive site not only is highly conserved with the human EKLK promoter (Lohmann and Bieker, unpublished), but also behaves as a strong enhancer element, even when placed adjacent to a non-erythroid promoter. However, erythroid specificity can be reestablished when this element is placed adjacent to the EKLK promoter, with the proximal GATA site proving critical for this enhancement (31). Experiments with transgenic mice that contain a β -galactosidase reporter whose expression is directed by EKLK promoter elements reveal that this short piece is sufficient to generate expression in the blood islands as well as in the fetal liver (35). In addition, both hypersensitive sites, along with the proximal promoter, are all required for this effect. Experiments with 32DEpo1 cells have directed attention to a 49 bp sequence within the most conserved, distal hypersensitive site, which contains the core sequence responsible for enhancement and contains three protein-binding sites *in vitro* (31).

Conditions to differentiate embryonic stem cells to embryoid bodies (EBs) in the absence of serum were established in order to identify cytokines responsible for induction of EKLK expression during early development (36). These studies, in conjunction with expression of dominant negative proteins, enabled us to conclude that the BMP (bone morphogenetic protein) 4/BMP receptor/Smad pathway plays a critical role in EKLK induction (Fig. 2).

Additional Unresolved Issues and Future Directions

A paradox follows from the observation that EKLK message is expressed in both primitive (yolk sac) and definitive (mouse fetal liver) erythroid cells, yet is functionally only required for definitive erythropoiesis. Since EKLK protein is present in both populations, it has been proposed

2. Schuh R, Aicher W, Gaul U, et al. A conserved family of nuclear proteins containing structural elements of the finger protein encoded by Krüppel, a Drosophila segmentation gene. *Cell* 1986; 47:1025–1032.
3. Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* 1991; 252:809–816.
4. Klevit RE. Recognition of DNA by Cys2,His2 zinc fingers. *Science* 1991; 253(5026):1367, 1393.
5. Stamatoyannopoulos G, Nienhuis AW. Hemoglobin switching. In: Stamatoyannopoulos G, Nienhuis AW, Majerus PW, Varmus H, editors. *The molecular bases of blood diseases*. 2nd ed. Philadelphia: W.B. Saunders Co.; 1994. pp. 107–155.
6. Feng WC, Southwood CM, Bieker JJ. Analyses of β -thalassaemia mutant DNA interactions with erythroid Krüppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J Biol Chem* 1994; 269:1493–1500.
7. Orkin SH. Regulation of globin gene expression in erythroid cells. *Eur J Biochem* 1995; 231:271–281.
8. Baron MH. Transcriptional control of globin gene switching during vertebrate development. *Biochim Biophys Acta* 1997; 1351: 51–72.
9. Wijgerde M, Grosveld F, Fraser P. Transcription complex stability and chromatin dynamics in vivo. *Nature* 1995; 377(6546):209–213.
10. Fraser P, Gribnau J, Trimborn T. Mechanisms of developmental regulation in globin loci. *Curr Opin Hematol* 1998; 5(2):139–144.
11. Bieker JJ. Role of erythroid Krüppel-like factor (EKLF) in erythroid-specific transcription. In: Stamatoyannopoulos G, editor. *Molecular biology of hemoglobin switching*. Andover (MA): Intercept, Ltd.; 1994. pp. 231–241.
12. Donze D, Townes TM, Bieker JJ. Role of erythroid Krüppel-like factor in human gamma- to beta-globin gene switching. *J Biol Chem* 1995; 270(4):1955–1959.
13. Nuez B, Michalovich D, Bygrave A, et al. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* 1995; 375:316–318.
14. Perkins AC, Sharpe AH, Orkin SH. Lethal β -thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* 1995; 375:318–322.
15. Lim SK, Bieker JJ, Lin CS, Costantini F. A shortened life span of EKLF^{-/-} adult erythrocytes, due to a deficiency of beta-globin chains, is ameliorated by human gamma-globin chains. *Blood* 1997; 90(3):1291–1299.
16. Perkins AC, Gaensler KML, Orkin SH. Silencing of human fetal globin expression is impaired in the absence of the adult β -globin gene activator protein EKLF. *Proc Natl Acad Sci* 1996; 93:12267–12271.
17. Wijgerde M, Gribnau J, Trimborn T, et al. The role of EKLF in human beta-globin gene competition. *Genes Dev* 1996; 10(22):2894–2902.
18. Ouyang L, Chen X, Bieker JJ. Regulation of erythroid Krüppel-like factor (EKLF) transcriptional activity by phosphorylation of a protein kinase casein kinase II site within its interaction domain. *J Biol Chem* 1998; 273(36):23019–23025.
19. Zhang W, Bieker JJ. Acetylation and modulation of erythroid Krüppel-like factor (EKLF) activity by interaction with histone acetyltransferases. *Proc Natl Acad Sci U S A* 1998; 95(17):9855–9860.
20. Wolffe AP, Pruss D. Targeting chromatin disruption: transcription regulators that acetylate histones. *Cell* 1996; 84:817–819.
21. Pazin MJ, Kadonaga JT. What's up and down with histone deacetylation and transcription? *Cell* 1997; 89:325–328.
22. Berger SL. Gene activation by histone and factor acetyltransferases. *Curr Opin Cell Biol* 1999; 11(3):336–341.
23. Zhang W, Kadam S, Emerson BM, Bieker JJ. Site-specific acetylation by p300 or CREB binding protein regulates erythroid Krüppel-like factor transcriptional activity via its interaction with the SWI-SNF complex. *Mol Cell Biol* 2001; 21(7):2413–2422.
24. Armstrong JA, Emerson BM. Transcription of chromatin: these are complex times. *Curr Opin Genet Dev* 1998; 8(2):165–172.
25. Kadonaga JT. Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell* 1998; 92(3):307–313.
26. Armstrong JA, Bieker JJ, Emerson BM. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* 1998; 95(1):93–104.
27. Kadam S, Emerson BM. Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes. *Mol Cell* 2003; 11(2):377–389.
28. Drissen R, Palstra RJ, Gillemans N, et al. The active spatial organization of the beta-globin locus requires the transcription factor EKLF. *Genes Dev* 2004; 18(20):2485–2490.
29. Southwood CM, Downs KM, Bieker JJ. Erythroid Krüppel-like factor exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. *Dev Dyn* 1996; 206(3):248–259.
30. Crossley M, Tsang AP, Bieker JJ, Orkin SH. Regulation of the erythroid Krüppel-like factor (EKLF) gene promoter by the erythroid transcription factor GATA-1. *J Biol Chem* 1994; 269:15440–444.
31. Chen X, Reitman M, Bieker JJ. Chromatin structure and transcriptional control elements of the erythroid Krüppel-like factor (EKLF) gene. *J Biol Chem* 1998; 273(39):25031–25040.
32. Anderson KP, Crable SC, Lingrel JB. Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Krüppel-like factor (EKLF) gene. *J Biol Chem* 1998; 273(23):14347–14354.
33. Jenkins NA, Gilbert DJ, Copeland NG, et al. Erythroid Krüppel-like transcription factor (EkLf) maps to a region of mouse chromosome 8 syntenic with human chromosome 19. *Mamm Genome* 1998; 9(2):174–176.
34. van Ree JH, Roskrow MA, Becher AM, et al. The human erythroid-specific transcription factor EKLF localizes to chromosome 19p13.12–p13.13. *Genomics* 1997; 39(3):393–395.
35. Xue L, Chen X, Chang Y, Bieker JJ. Regulatory elements of the EKLF gene that direct erythroid cell-specific expression during mammalian development. *Blood* 2004; 103(11):4078–4083.
36. Adelman CA, Chattopadhyay S, Bieker JJ. The BMP/BMPR/Smad pathway directs expression of the erythroid-specific EKLF and GATA1 transcription factors during embryoid body differentiation in serum-free media. *Development* 2002; 129(2):539–549.
37. Hu M, Krause D, Greaves M, et al. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev* 1997; 11(6):774–785.
38. Reese TT, Gregory RC, Sharlow ER, et al. Epo-induced hemoglobinization of SKT6 cells is mediated by minimal cytoplasmic domains of the Epo or prolactin receptors without modulation of GATA-1 or EKLF. *Growth Factors* 1997; 14(2–3):161–176.
39. Ziegler BL, Muller R, Valtieri M, et al. Unicellular-unilineage erythropoietic cultures: molecular analysis of regulatory gene expression at sibling cell level. *Blood* 1999; 93(10):3355–3568.