The multifunctional role of EKLF/KLF1 during erythropoiesis

Miroslawa Siatecka¹ and James J. Bieker¹²³*

¹Department of Developmental and Regenerative Biology, ²Black Family Stem Cell Institute, ³Tisch Cancer Institute, Mount Sinai School of Medicine, New York, NY 10029, USA

Running title: EKLF/KLF1 in erythropoiesis
Key Words: red blood cells, transcription, anemia, globin switching, cell cycle
Category: review article

*Correspondence:
Mount Sinai School of Medicine
Department of Developmental and Regenerative Biology
One Gustave Levy Place
Box 1020
New York, NY 10029, USA.
E-mail: james.bieker@mssm.edu
Abstract

The cellular events that lead to terminal erythroid differentiation rely on the controlled interplay of extra- and intra-cellular regulatory factors. Their downstream effects are highly coordinated and result in the structural/morphologic and metabolic changes that uniquely characterize a maturing red blood cell. EKLF (Erythroid Krüppel-like factor; KLF1) is one of a very small number of intrinsic transcription factors that play a major role in regulating these events. This review covers three major aspects of erythropoiesis in which EKLF plays crucial functions. First, at the MEP (megakaryocyte-erythroid progenitor) stage, where EKLF is involved in erythroid lineage commitment; second, during the global expansion of erythroid gene expression in primitive and definitive lineages, where it plays a direct role in globin switching; and third, during the terminal maturation of red cells where it helps control exit from the cell cycle. We finish by describing recent studies of mammalian EKLF/KLF1 mutations that lead to altered red cell phenotypes and disease.
Introduction

EKLF (Erythroid Krüppel-like factor; KLF1) is the founding member of the Krüppel-like factor family that comprises 17 different transcription factors that play critical roles in diverse biological processes. KLFs are characterized by a highly conserved DNA-binding domain that belongs to the well-studied class of C2H2 zinc fingers.

EKLF is essential for erythropoiesis. Its expression is restricted to the hematopoietic organs throughout all developmental stages, initiating at embryonic day E7.5 in the yolk sac, then switching to the fetal liver, and finally to the adult bone marrow and the red pulp of the spleen. EKLF protein is functionally present in both primitive and definitive erythroid populations.

EKLF consists of two domains, an N-terminal proline-rich transactivation domain and a C-terminal DNA binding domain that recognizes the consensus binding site 5’CCN CNC CCN. Recently, this motif has been more precisely defined and comprises the sequence 5’ CCM CRC CCN (where R represents A or G and M represents A or C), thus providing additional specificity to the “N” positions previously predicted from in vitro studies. Such a motif is found in the regulatory regions of many erythroid genes. EKLF can also bind to non-canonical consensus sites as found in the AHSP promoter; however, it is not clear whether the surrounding nucleotides or the overlapping nature of the CACC sites facilitate binding.

An extensive description of EKLF biochemical and molecular studies are beyond the scope of this review. In brief, EKLF has been primarily characterized as a transcriptional activator, but can also exert transcriptional repression. EKLF activity is modulated by post-translational modifications that promote protein-protein interactions (Figure 1), some of which follow from its interaction with cofactors p300/CBP. Acetylation of lysine 288 and the phosphorylation status of threonine 41 are critical for optimal transactivation activity. Acetylation of K288 in particular leads to increased association of EKLF with the chromatin remodeling SWI/SNF-related complex in vitro, yielding an open chromatin domain and transcription of the adult β-globin promoter. Integration of transcription and chromatin modifying activities enables EKLF to play a central role in formation of the active chromatin hub across the β-like globin locus. Given its prominent role as an activator, it is intriguing that EKLF can also act as a repressor. For instance, sumoylation of K74 is implicated in EKLF repression activity, likely by tethering components of NuRD complex to specific DNA targets.

Genetic ablation of EKLF leads to profound β-thalassemia and embryonic lethality at day E14-15 of gestation due to defective definitive erythropoiesis partly caused by a major deficit in β-globin expression. However, restoration of β-like globin alone is not sufficient to rescue the lethal phenotype, suggesting additional deficits in expression of other EKLF target genes. Recent analysis of differential gene expression between wild-type and EKLF-deficient fetal liver cells together with EKLF chromatin immunoprecipitation coupled to high-throughput DNA sequencing (ChIP-seq) studies have expanded the target repertoire of EKLF beyond that of adult β-globin, implicating a more global regulatory role for its activity during erythropoiesis.

Dual role of EKLF in establishing lineage fate

Hematopoiesis is a relatively linear and hierarchical process in which hematopoietic stem cells undergo successive symmetric and asymmetric divisions to yield committed progenitor cells. Stem cell potency is progressively restricted as they differentiate towards mature lymphoid, myeloid, and erythroid lineages. Differentiation at each stage depends on the interplay and
equilibrium of different sets of transcription factors. EKLF is not present in hematopoietic stem cells and barely detectable in multipotent myeloid progenitors. However, starting with CMP (common myeloid progenitor) cells and their MEP progeny, expression of EKLF selectively increases as cells mature towards the erythroid lineage. In sharp contrast, EKLF levels are down-regulated in those MEPs that differentiate towards megakaryopoiesis, such that bipotential differentiation of MEPs results in erythroid progenitors exhibiting an 80-fold greater level of expression of EKLF compared with that seen in megakaryocyte progenitors.

Mouse embryonic stem (ES) cell lines with doxycycline-inducible EKLF have been used to study its role during MEP differentiation, particularly to test the effect of EKLF misregulation (gain-of-function; GOF) on the normal hematopoietic progression as the ES cells differentiate to embryoid bodies (EBs). Forced expression of EKLF negatively affects megakaryopoiesis as shown by significant decreases in megakaryocytic markers (double positive CD41/CD42b and CD41/CD42d) in EKLF-induced cultures as compared to uninduced. Intriguingly, this coincides with higher levels of the mature CD71/Ter119 red cell markers. Morphologically, the induced erythroid population contains more differentiated basophilic and polychromatophilic erythroblasts than uninduced cultures. The timing of EKLF induction during EB differentiation is important for both megakaryocyte inhibition and erythroid lineage promotion (Figure 2).

Reciprocal loss-of-function analyses have been performed using primary EKLF-deficient mouse yolk sac or fetal liver cells (as sources of primitive or definitive erythropoietic cells), human cord blood progenitor cells, and mouse erythroleukemia (MEL) cells that stably express an anti-EKLF shRNA. In contrast to the GOF studies, EKLF knock-out embryos yield an increased number of circulating platelets, and E13.5 fetal liver cells from EKLF-null embryos cultured in the presence of thrombopoietin show more than a two-fold increase in double-positive CD41/42b and CD41/42d populations compared to wild type fetal liver cells. Moreover, megakaryocyte colony numbers are 3- to 4-fold greater from an unsorted or a CMP-sorted population when derived from EKLF-deficient cells. These colonies are more expanded in size compared with the relatively compact colonies derived from the wild-type cells. Thus, EKLF-null fetal liver cells produce a greater number of megakaryocyte progenitors and its progeny than do wild type cells.

However, at the genetic level there is considerable mis-regulation of megakaryocyte and erythroid expression patterns resulting from loss of EKLF. Both primitive (EryP) and definitive (EryD) erythrocytes (CD71+ population) from EKLF-null embryos show abnormal expression of a range of intracellular and cell surface proteins. For example, EKLF-null or -heterozygous EryP fail to express Ter119. At the same time, CD41 is aberrantly expressed on the surface of circulating E12.5-E13.5 EryP and E14.5 EryD EKLF-deficient cells, and expression of the gene for Itga2b/GpIIb is increased at least 80-fold over that measured in WT or EKLF-heterozygote EryP. mRNA levels for Fli-1, a transcription factor that plays a critical role in megakaryopoiesis, is increased 25-fold in EKLF-null EryP cells.

Generally then, the loss of EKLF leads to a failure to silence megakaryocyte-specific genes in erythroid progeny from the MEP, resulting in a partial identity crisis within maturing EryP and EryD. Specification to the erythroid lineage is compromised in favor of the megakaryocyte program, which is more pronounced in EryD. These observations suggest that EKLF regulates cell identity not only in MEP progenitors but also throughout EryP or EryD maturation. Consistent with this, the EKLF knock-down MEL cell line is associated with both significant reduction of EKLF and its target genes and increase of megakaryocytic cell surface markers and mRNA levels of Gp1b, GpIX, GpIIla and Fli1. Similarly, the expression level of erythroid marker GPA is reduced and in parallel the number of CD41/CD42 double positive cells are increased in human cord blood progenitor cells treated with shRNA against EKLF. However,
the timing is critical, as already-committed erythrocytes treated with shRNA do not show any reactivation of megakaryocytic genes, nor is there any bi-phenotypic readout.

Taken together, these studies suggest that EKLF plays a dual role in hematopoiesis (Figure 2) by both promoting erythroid lineage development while at the same time repressing or attenuating the megakaryocytic program, not only in the MEP population but also continuously throughout erythroid maturation. This distinguishes EKLF from other commonly expressed transcription factors (such as Gata1 and p45-NFE2) that play critical positive roles for both megakaryocytic and erythroid lineages.

While the mechanisms by which EKLF directs lineage fate decisions and differentiation programs remain incompletely understood, comparative analysis of expression arrays between wild-type and EKLF-deficient fetal liver cells reveal differential expression of transcription factor Fli-1, a critical megakaryocytic activator. Functional cross-antagonism between Fli1 and EKLF might be involved in the commitment toward erythroid versus megakaryocytic lineages (Figure 2). EKLF and Fli-1 can physically interact in vitro and thus may negatively affect each other’s activity. Fli1 can inhibit EKLF-dependent transcription of the ß-globin gene in cotransfected MEL cells, and EKLF can inhibit Fli-1 dependent transcription of the megakaryocyte GPIX gene in cotransfected COS7 cells. Moreover, EKLF overexpression in differentiating EBs leads to a dramatic down regulation of Fli1 mRNA levels within 24 hours. EKLF also activates the BKLF repressor, whose expression may be implicated in inhibition of megakaryocytic genes.

Additionally, Fli-1 and EKLF might compete for limiting amounts of common cofactors such as GATA-1 during formation of mutually exclusive multi-protein complexes that selectively activate either erythrocytic or megakaryocytic promoters. In support of this, EKLF knockdown is associated with a simultaneous increase of both Fli-1 and GATA-1 recruitment at megakaryocytic gene promoters.

Importantly, the sumoylation status of EKLF is critical for its effect on megakaryopoiesis. Transgenic mice designed to mis-express EKLF in the megakaryocytic progenitor compartment inhibit megakaryocyte colony formation and cellular output. However, a sumoylation-deficient mutant of EKLF, expressed in the same way, has no effect on megakaryocyte development. As a result, the post-translational modification status of EKLF and the identity of its protein partners may be relevant to any directive EKLF role in bipotential decisions from the MEP.

**EKLF in global expression of erythropoietic genes**

The differentiation status of MEP progeny committed to the erythroid lineage can be monitored by the relative expression of cell surface markers CD71 (transferrin receptor) and Ter119. Discrete populations of cells, enriched for a particular stage of erythropoiesis and characterized by distinct morphologies, can be separated by flow cytometry using these markers. In E13.5 wild-type liver cells five populations can be distinguished: R1 (CD71LO Ter119NEG, early erythroid progenitor; BFU-E), R2 (CD71HI Ter119LO, erythroid progenitor; CFU-E), R3 (CD71HI Ter119HI, proerythroblast, basophilic erythroblast), R4 (CD71MD Ter119HI, polychromatic erythroblast, orthochromatic erythroblast), and R5 (CD71LO Ter119HI, reticulocytes and erythrocytes).

Comparison of the CD71/Ter119 populations in fetal liver cells (E13.5) reveals that only the R1 and R2 categories are present in EKLF-deficient embryos, demonstrating that lack of EKLF prevents the progression of erythropoiesis between the R2 and R3 stages (Figure 2).
Alternatively, the Ter119 epitope might be a direct target of EKLF. In this case the R3-R5 populations might be present but not recognizable by FACS for Ter119. Gene expression profiling of sorted R1+R2 populations from WT and EKLF-null fetal liver cells shows dysregulation of ~3000 genes (2534 are downregulated and 665 are upregulated in EKLF-null embryos). Perturbation of genes involved in cell cycle and DNA replication pathways predominate. Consistent with a role for EKLF in erythroid progression, the level for E2F2, a transcription factor that controls S-phase entry and DNA synthesis, is significantly decreased in the EKLF-null cell to <5% of wild type. Cell cycle profiles of sorted R1+R2 E13.5 fetal liver cells confirm impaired progression from G1 to S phase in erythroid progenitor and precursor cells lacking EKLF. E2F2 has been shown to be a direct target of EKLF, and EKLF binding to E2F2 regulatory elements maintains the locus in an active chromatin state. In turn, E2F2 protein transactivates other cell cycle pathway members, allowing progression of early erythroid progenitor cells, including BFU-E and/or CFU-E, through the G1/S transition prior to terminal erythroid differentiation. Insufficient expression of E2F2 might be a rate-limiting factor in erythroid precursor progression and could explain at least in part the lack of terminally differentiated erythrocytes in EKLF-deficient fetal liver cells.

Thus, during erythropoiesis, EKLF’s function is essential for the completion of terminal differentiation beginning at the proerythroblast stage.

Initially, the lethal EKLF-null phenotype was thought to be primarily due to the loss of β-globin gene expression, but the deficit is more severe than that seen in mice with an inactivated adult β-globin gene, and correction of globin chain imbalance fails to completely rescue EKLF-null mice. These data argue for the existence of additional non-globin defects in EKLF-null red cells. Several approaches have been undertaken to address this issue and generate a more complete list of EKLF target genes, including comparisons between transcriptional profiles of wild-type versus EKLF-deficient EryD cells or of the B1.6 EKLF-null cell line (now called K1-ER cells) with or without EKLF-dependent rescue. These studies share a remarkably similar set of potential direct EKLF target genes. Many more genes are downregulated in EKLF-null cells than upregulated, confirming EKLF’s primary role as a transcriptional activator in vivo. In addition to the β-globin gene – the best studied EKLF target – putative targets comprise genes essential for all categories of red blood cell function, including red cell membrane, cytoskeletal, heme biosynthetic, transmembrane proteins, transcription and cell cycle factors, and blood group antigens. Representative examples from these categories are detailed below.

**EKLF targets in definitive cells**

Expression of dematin (Epb4.9), an important cytoskeletal protein required for membrane integrity and stability in erythrocytes, is highly dependent on EKLF, as wild type fetal liver cells express a ~20-fold greater level compared to EKLF-null cells. In addition, mRNA levels for several other membrane associated proteins (ankyrin, β-spectrin, band 3, aquaporin, Rh30, p55, protein 4.2, and ERMAP) show significant differences in expression in EKLF-deficient versus wild-type fetal liver cells.

AHSP (α hemoglobin stabilizing protein), a small protein that binds free α-globin chains, is another EKLF target gene whose expression is virtually absent in EKLF-null fetal liver cells and is dramatically upregulated in K1-ER cells after EKLF rescue. AHSP inhibits production of reactive oxygen species from α-hemoglobin and prevents the precipitation of highly unstable, cytotoxic free α-globin chains in red cells.

Heme biosynthetic pathway enzymes, such as Alas2, Pdbg, and Bzrp, are reduced in the EKLF-null red cells. This may account for the defective heme synthesis and siderotic (mitochondrial) iron accumulation phenotype.
As noted, EKLF-null fetal liver cells show little Ter119 expression\textsuperscript{25,30}. In addition, expression of the transferrin receptor 1 (\textit{TfR1}), another trans-membrane gene commonly used as an erythroid marker, is reduced\textsuperscript{25}.

The Gardos channel, encoded by \textit{Kcnn4}, has been implicated in cell size reduction during erythroid maturation. EKLF-null progenitors differentiated in culture do not reduce their cell size, which implicates \textit{Kcnn4} gene as a potential EKLF target\textsuperscript{26}.

Expression-profiling experiments suggest that many blood group antigens are regulated by EKLF; for example, \textit{Rh-cde}, \textit{Duffy}, \textit{ICAM4/LW}, \textit{Ermap/Scianna}, glycoporphin A (\textit{GPA}), \textit{CD47} (an Rh antigen–associated protein), and \textit{CD59} (a GPI-anchored protein lost in paroxysmal nocturnal hemoglobinuria [PNH]) are differentially expressed.\textsuperscript{25} Interestingly, most of these proteins coexist in a mega-complex that links the membrane to the underlying cytoskeleton\textsuperscript{42}. Importantly, mouse \textit{Rh-cde}, \textit{GPA}, \textit{CD59}, \textit{CD47}, and \textit{Ermap/Scianna} genes all contain CACC sites in their proximal promoters that are likely to bind EKLF\textsuperscript{43,44}.

Expression levels of transcription factors such as BKLF/KLF3, TIEG/KLF10, Eto/Mtg8, Lmo2, and \textit{RAR\gamma} are markedly reduced in EKLF-null cells\textsuperscript{25,33}. Low expression of such a combination of transcription factors may amplify erythroid-specific effects of EKLF deficiency.

**EKLF targets in primitive cells**

The expression of the novel EKLF target genes has also been examined in E11.5 embryonic blood during primitive erythroid cell differentiation\textsuperscript{25,26,30}. There is a dramatic reduction in expression of \textit{AHSP}, dematin, \textit{Mgst3}, \textit{Ermap/Scianna}, \textit{Acp3}, the peripheral benzodiazepine receptor (\textit{Bzrp}), \textit{Rh-cde}, and \textit{ICAM4/LW} in EKLF-null embryonic red cells. Primitive EKLF-null red cells show defects in gene expression and surface exposure of Ter119. Loss of dematin and Ter119 might partly explain the abnormally ruffled red cell membrane. Despite this abnormal phenotype, the embryonic EKLF-deficient red cells are not significantly different in gross morphology from wild-type littermates. Given that EKLF-null primitive erythrocytes appear to have a relatively normal (albeit transient) lifespan and that embryos survive until later in gestation, it is likely that other CACC-binding proteins, such as LKLF (\textit{KLF2}), may partly compensate for the absence of EKLF at certain promoters of genes such as the zeta-, \textit{ey}-, and \textit{bh1}-globins\textsuperscript{45} as well as non-globin genes\textsuperscript{25}.

Together these data implicate EKLF as an essential regulator of global erythroid gene expression in both primitive and definitive red cell populations.

**Genomic EKLF occupancy of regulatory regions**

To ascertain whether genes identified from the genome-wide expression profiling experiments are indeed direct EKLF transcriptional targets, EKLF ChIP-seq in primary erythroid tissue has been performed\textsuperscript{13}. At least 945 sites in the genome of E14.5 fetal liver erythroid cells are occupied by endogenous EKLF. 16\% of these in vivo EKLF binding sites occur within 1 kb of transcription start sites (TSSs), but the majority (52\%) are located at distances >10 kb from any known TSS\textsuperscript{13}. These observations confirm that EKLF is a direct regulator of many differentially regulated genes, particularly those involved in the production of the globins, heme, blood group antigens, and in the maintenance and integrity of the erythroid cell membrane and cytoskeleton. Intriguingly, EKLF binding is implicated in regulation of apoptosis in red cells, being detected within the anti-apoptotic, pro-survival \textit{Bcl2l1} (\textit{Bcl-X}), \textit{Xiap}, \textit{Rad23a} and \textit{Pim1} genes, and also paradoxically within the pro-apoptotic \textit{Siva1} and \textit{Ctsb} genes\textsuperscript{13,46}. EKLF also occupies the \textit{Fn3k}, \textit{Slc2a4}, and \textit{Pigq} genes, suggesting a direct role in sugar metabolism and protein modifications such as glycosylation and GPI biosynthesis. EKLF occupancy within the alpha-globin gene cluster is detected at HS-31, HS-26, HS-21, HS-8, and \textit{Hba-a2} and \textit{Hba-a1} gene promoters\textsuperscript{13}. 
consistent with earlier ChIP analyses at that locus\textsuperscript{47}.

EKLF ChIP-seq peaks are found in the vicinity of the transferrin receptor (Tfrc1; CD71), the mitochondrial iron transporter mitoferrin (Slc25a37), and the Slc11a2 and Steap3 genes that encode proteins involved in iron transport and its reduction in the endosome\textsuperscript{13}. EKLF directly regulates the Abcb10 (ABC-me) and Abcg2 genes that produce heme exporter proteins involved in the final step in the assembly of the hemoglobin molecules\textsuperscript{13}.

The ChIP-seq analysis reveals that EKLF occupancy lies within the first intron or at intragenic regions of a large majority of the activated target genes and not at the promoter region. This finding supports the notion that in many cases EKLF might be activating target genes at alternative erythroid-specific promoters, or via intronic enhancer elements, as has been shown at the E2f2, E2f4, and p21 (Cdkn1a) genes\textsuperscript{13,36,48}. However, the global binding pattern is not static, as the categories of EKLF occupancy in erythroid progenitors are not equivalent to that seen in mature erythroblasts\textsuperscript{49}.

The degree to which gene activation occurs depends on many additional factors besides the binding of this single transcription factor. EKLF and GATA1 co-occupy many sites in the erythroid genome in a distinct nonrandom configuration\textsuperscript{13}. For example, 48\% of EKLF ChIP-seq peaks have a corresponding GATA1 ChIP-seq peak within 1 kb, suggesting that the two transcription factors cooperate at each of these cis regulatory elements in vivo\textsuperscript{13}.

Coordination of global erythroid gene expression
Recent studies establish a new gene expression paradigm by showing that co-regulated genes together with their regulatory factors preferentially cluster at specialized transcription factories\textsuperscript{50}. Such factories are associated with the active, hyperphosphorylated forms of RNA polymerase II (RNAPII S\textsuperscript{5P}) to optimize efficient and coordinated transcriptional control. EKLF mediates preferential (at very high frequencies (63–79\%)) co-associations of EKLF-regulated genes at a limited number of these specialized transcription factories. Nearly all EKLF foci overlap with RNAPII-S\textsuperscript{5P} foci, indicating that 10–20\% of transcription factories in mouse erythroid cells contain high levels of EKLF\textsuperscript{50}.

These interactions are specifically disrupted in the absence of EKLF. Hbb, Epb4.9 and Ahsp all show markedly reduced association with transcription factories in EKLF-null erythroid cells, consistent with their dependence on EKLF for expression. Both intrachromosomal (Hbb-Kcnn4) and interchromosomal (Hbb-Hba; Hbb-Epb4.9) interaction frequencies were reduced in EKLF-null cells. These data indicate that EKLF is required for preferential colocalization of EKLF-regulated genes at shared transcription factories\textsuperscript{50}.

It is worth noting that the majority of EKLF-regulated genes are not absolutely dependent on EKLF for basal expression but instead require EKLF for increased expression to attain optimal levels in the definitive erythroid lineage. The individual genes may indirectly benefit from cooperative associations in these specialized microenvironments. The preferential associations in transcription factories substantially affect higher order chromosomal conformations and are a major driving force in tissue-specific chromosomal positioning\textsuperscript{50,51}.

EKLF in globin switching
One of the main genetic signatures distinguishing primitive and definitive erythropoiesis is the expression of stage-specific hemoglobin variants. The transitions from embryonic to fetal and to adult \(\beta\)-like globin expression accompany a shift in the primary site of erythropoiesis from the yolk sac to the fetal liver and ultimately the bone marrow\textsuperscript{52}. Historically, the role of EKLF in
enabling the final switch to adult b-globin has been the major focus of genetic, biochemical, and molecular studies since its discovery\textsuperscript{53,54}.

Developmental stage-specific alterations in EKLF abundance may also help to mediate the competitive interactions of b-like globin gene family members with the far-upstream LCR (locus control region). The level of mouse EKLF increases 3-fold in definitive erythroid progenitors compared to primitive erythroid progenitors\textsuperscript{55}. Similarly, mice with a low-expressing hypomorphic EKLF allele demonstrate delayed silencing of murine embryonic globins (and transgenic human fetal globin)\textsuperscript{56}.

Recent studies show that EKLF plays a critical role in regulating the developmental switch between fetal and adult hemoglobin expression both by direct activation of b-globin and indirect repression of g-globin gene expression in adult erythroid progenitors via regulation of Bcl11a\textsuperscript{57} (Figure 3). For example, the HPFH (hereditary persistence of fetal hemoglobin) phenotype within a Maltese family has been mapped to a heterozygous nonsense mutation (K288X) in EKLF that leads to haploinsufficient expression. Genome-wide expression studies show that among known EKLF targets, the g-globin repressor Bcl11a\textsuperscript{58-60} is highly downregulated in these individuals while the embryonic and fetal globin are markedly upregulated\textsuperscript{57}. Manipulated EKLF levels are directly proportional to Bcl11a levels and inversely proportional to g-globin levels, and ChIP analyses confirm a direct EKLF occupancy within the promoter region of Bcl11a gene in both humans and mice\textsuperscript{56,57}. Consistent with these data, butyrate treatment of transgenic murine fetal liver definitive erythroid cells downregulates levels of EKLF and Bcl11a while increasing embryonic and fetal globin expression\textsuperscript{61}.

EKLF may contribute even more directly to b-like globin gene switching. Tagged-EKLF interacts in vivo with the mouse embryonic b-like globin or the transgenic human fetal g-globin promoter in primitive erythroid cells, and then switches to bind the mouse and human adult b-globin promoter in definitive cells\textsuperscript{55,56}. Although EKLF is not absolutely necessary for embryonic/fetal globin expression, it has quantitative effects on the levels of some of these genes\textsuperscript{62-64}. Its interaction with these promoters may help to establish their most optimal transcriptional activity prior to the final switch, at which point the lower affinity of EKLF for the g- vs b-CACCC element\textsuperscript{12} may enable a quick release and switch when the b-site becomes available. A complication is that EKLF levels are lower in the very cells (primitive) within which it is postulated to act at a suboptimal binding affinity target (embryonic/fetal CAC site). It remains to be demonstrated whether EKLF binds the g-globin promoter in human fetal liver cells. However, it is intriguing that treatment of primary human erythroid precursors with short-chain fatty acid derivatives leads to an increase in the occupancy of EKLF at the g-globin promoter\textsuperscript{65}.

Based on all these observations, EKLF knock-down could be used to increase g-globin expression in a therapeutic setting. However, such a design may compromise b-globin and other selected genes’ expression if it is even slightly too low (see below).

**EKLF during red cell terminal maturation**

The switch from proliferation to differentiation during the terminal stages of tissue maturation is a tightly controlled process that relies in part on transcription factor-mediated activation of cell cycle components. In general, differentiation is associated with cell cycle arrest. However, in the erythroid lineage, early stages of terminal differentiation are coupled to proliferation. Later stages of erythroid differentiation, from erythroid precursors to mature hemoglobinized
erythrocytes, require 3 to 4 cell divisions, during which the cell size is reduced, gene expression levels decrease, the nucleus condenses, and in the ultimate gene-silencing process, is lost by enucleation66,67.

**Divisions during erythroid maturation**

As mentioned before, EKLF, via activation of the E2f2 locus, is required for the cell cycle progression that precedes terminal erythroid differentiation34,36. Compared to wild type cells, EKLF-null E13.5 fetal liver definitive erythroid cells show a significant increase in the percentage of cells with G0/G1 (2N) DNA content and a parallel decrease in the percentage of cells containing S-phase (>2N) or G2/M (4N) DNA contents. A similar reduction in cells with S-phase DNA content is observed in EKLF-null cells during primitive (E10.5) erythropoiesis. Thus, expression of E2F2 is significantly reduced in both primitive and definitive erythropoiesis in EKLF-deficient embryos. However, expression of closely related transcription factor E2F4 is affected only at E14.5 in EKLF-null fetal liver when compared to wild type36. E2F2 and E2F4 are components of the Rb-E2F complex that controls S-phase entry and that are highly expressed during erythroid differentiation. A loss of either transcription factor in mice leads to macrocytic anemia, which is associated with defects in erythroid maturation36,68,69. Cell division and reduction in cell size are tightly linked, suggesting that the major role of E2F2 and E2F4 in erythropoiesis is the promotion of cell proliferation through the direct transactivation of downstream cell cycle genes. ChIP analysis in primary fetal liver cells show endogenous EKLF occupancy at E2f2 and E2f4 regulatory elements located in their promoter and intronic regions, some of which exhibit strong enhancer activity13,34,36. The appearance of enlarged red cells in the peripheral blood of E2F2-null mice is consistent with premature cell cycle exit.

Restoration of the cell cycle balance in EKLF-null red cells by the additional loss of Rb partially rescues S-phase entry during definitive erythropoiesis36. However, many other aspects of erythroid development remain perturbed; for example, there is no indication of a differentiation rescue, shown by the persistent absence of erythroid cell surface markers Ter119 and CD7136.

**Cell cycle exit at late stages of terminal erythroid maturation**

The level of EKLF protein changes during erythroid development. An anti-proliferative effect of EKLF is suggested by an increased derivation frequency of immortal clones from EKLF-null compared with EKLF-heterozygous fetal liver cells after J2 virus infection, indicating that EKLF discourages immortalization70. The absence of EKLF also results in a markedly increased rate of transformation by cooperating oncogenes v-mil-raf and v-myc. It is perhaps then not surprising that increased levels of EKLF during terminal maturation correlate with the transition of the cell cycle status from proliferation to differentiation71,72. This change is mediated (at least in part) by EKLF’s direct activation of the cyclin dependent kinase (Cdk) inhibitors p18 and p2148,73. p18 is a member of Ink4 family that inhibits cyclin D-CDK4 complexes in early G1, whereas p21 belongs to the CIP/KIP family that regulates cell cycle transition through inhibition of cyclin E-CDK2 complexes at later stages of G1.

Analysis of terminal differentiation in G1E-ER cells after β-estradiol-induced nuclear translocation of GATA1 results in a concomitant increase in expression of EKLF and the cell cycle inhibitor p21; however, siRNA knock-down of EKLF under these conditions proportionately decreases expression of p2148. Similarly, forced over-expression of EKLF in MEL cells directly correlates with an increase in mRNA and protein level of p21, exit from the cell cycle, and G0/G1 arrest48. The cyclin dependent kinase inhibitor p18 was identified as a potential novel EKLF target gene from an expression profiling study73. ChIP analyses performed
using erythroid cell lines confirm EKLF occupancy at regulatory elements located in the proximal promoter and intron 1 of the p21 gene and at an upstream promoter region of the p18 gene. Together these data are consistent with the notion that EKLF levels directly activate p18 and p21 and help mediate cell cycle arrest. Neither the p18 or p21 genes are strongly occupied by EKLF in whole fetal liver nuclei. This may reflect the heterogeneity of erythroid cell differentiation states in fetal liver and possible dynamic occupancy of EKLF at different cell cycle gene regulators as differentiation progresses.

In summary, EKLF promotes the differentiation and maturation of red cells by direct regulation of several components of the cell cycle machinery. However, the complex set of events that occur at the final erythroid cell divisions may be influenced by its post-translational modification status (Figure 1) and is certainly accomplished in coordination with other transcription factors.

Although separate roles for EKLF during the proliferation (activation of E2F2) versus differentiation (activation of p18 and p21) stages of erythroid differentiation as described above may be operant, the role of cell cycle components in erythroid differentiation is complex and does not follow the course typically observed in other cell types for renewal divisions (e.g., E2F2 and E2F4). Indeed, the role of E2Fs may be heavily dependent on the cellular context in which they are expressed, effecting a dual functionality in progenitor cells compared to differentiating cells.

**EKLF in enucleation**

EKLF function remains crucial for the final steps of erythroid maturation. Among a number of relevant genes affected by absence of EKLF is ICAM4, which codes for an intercellular adhesion molecule that is critical for erythroblast/macrophage interactions within the erythroblastic island. Consistent with this requirement, peripheral blood from a Congenital Dyserythropoietic Anemia (CDA) patient containing a missense EKLF mutation (described further below) exhibits an abnormally high level of nucleated erythroblasts.

**EKLF mutations related to hematologic parameters and disease**

Given the importance of EKLF during erythropoiesis, it is not surprising that a number of mutations within the mouse and human EKLF/KLF1 locus have recently been shown to cause hematologic changes and disorders (summarized in Table 1). These vary in their severity depending on the nature of the mutation. We have divided these mutations into two major categories based on the mechanisms that lead to the altered phenotypes as shown in Figure 4.

**Dominant phenotype**

The most extreme phenotype is generated by mutations wherein a functionally important amino acid is replaced, leading to a protein with altered activation properties (Figure 4A). Two recent examples vividly illustrate this point. First, the mouse severe neonatal anemia (Nan) mutant exhibits a hemolytic anemia that displays many of the features of hereditary spherocytosis. Defects in structural membrane skeleton genes provide the molecular explanation for the associated red blood cell fragility. As a result, it came as a surprise to find not only that the primary defect in the Nan mouse maps to the EKLF gene, but also that it presents as a dominant phenotype despite the presence of the wild type EKLF allele. Second, Congenital Dyserythropoietic Anemia (CDA) is an inherited red blood cell disorder whose hallmarks are morphological abnormalities of erythroblasts in bone marrow, ineffective erythropoiesis and
hemolysis, that can be divided into three major categories based on the erythroblast nuclear morphology. The genetic explanation for the more prevalent CDA1 and 2 categories have been determined, but analysis of unclassified CDA patients enabled their causative mutation to be mapped to \textit{EKLF}, again displaying a dominant phenotype even in the presence of the wild type \textit{EKLF} allele. The remarkable aspect of these studies is that the mutations lead to a change in the same amino acid: E339 in mouse \textit{EKLF} is equivalent to E325 in human \textit{EKLF}. Glutamic acid at this site is absolutely conserved across \textit{EKLF} proteins from different species and across the entire KLF family. This amino acid is located within the second zinc finger of mouse \textit{EKLF}, precisely within the structural \(\alpha\)-helical motif that fits into the major groove, enabling a direct interaction with the middle nucleotide of the \textit{EKLF} binding motif. However, the amino acid substitution is not the same in the CDA patients as in Nan-\textit{EKLF} and is of opposite charge (aspartic acid in the mouse, lysine in the human) leading to intrinsic differences in DNA recognition and in the repertoire of affected target genes and resultant pathology.

Effects of the mouse E339D mutation have been characterized by biochemical tests and structural modeling of \textit{EKLF}. Although conservative, this change alters the DNA-binding specificity of Nan-\textit{EKLF}, limiting transcription to a select subset of \textit{EKLF} target genes based on a subtle change in its cognate DNA element in which a single nucleotide change makes the difference between recognition by both WT and Nan-\textit{EKLF} (at 5'NGG GCG NGG) versus recognition by WT \textit{EKLF} alone (at 5'NGG GTG NGG). An unusual observation is that even when expressed at an equivalent level to wild-type \textit{EKLF}, Nan-\textit{EKLF} selectively interferes with expression of target genes whose cognate elements are not recognized by the mutant. However, because some target genes remain unaffected by Nan-\textit{EKLF}, this mutant cannot act as a dominant-negative. The inability of the single WT copy to maintain normal expression at some but not all target sites in the \textit{Nan} mouse also differentiates this mechanism from haploinsufficiency (described below), particularly given the profound effects of Nan-\textit{EKLF} on select targets that are, at most, mildly sensitive to \textit{EKLF} dosage. Among the target genes most affected are \textit{E2f2}, \textit{BKLF/Klf3}, and the red cell membrane protein dematin, whose low expression level partly explains the spherocytosis and hemolysis seen in these mice. It is noteworthy that there is a high level of residual embryonic globin expression in \textit{Nan}. The net result is a distorted genetic output, anemia, and altered hemoglobin switching.

A different pathology is exhibited by the human E325K mutation, which leads to abolished expression of the water channel AQP1 and of the adhesion molecule CD44, and reduced expression of two other adhesion molecules, BCAM and ICAM4. The peripheral blood contains evidence of poikilocytosis, anisocytosis, fragmented erythrocytes, and a large number of nucleated red blood cells, mostly orthochromatic erythroblasts, which suggest a failure of terminal erythroid differentiation. Still, there is some similarity to the Nan mouse; for example, \(\beta\)-like globin expression is dysregulated, as these CDA patients exhibit high expression levels of fetal and embryonic hemoglobins.

**Haploinsufficiency**

Another category of \textit{EKLF} mutation is more directly related to haploinsufficiency, and follows from two generative mechanisms (Figure 4B). Historically, \textit{EKLF} wild type and heterozygote (+/-) embryos and adults were considered indistinguishable (although hints of dose-dependent effects had been noted), but recent studies conclusively show that both primitive and definitive erythropoiesis are sensitive to \textit{EKLF} heterozygosity, primarily because expression of a select set of \textit{EKLF} target genes are sensitive to the reduced cellular levels of \textit{EKLF}. For example, the rare blood group In(Lu) phenotype is caused by inheritance of a loss-of-function...
mutation on one allele of EKLF. Nine different mutations are found in the EKLF locus region. Some of them contain premature stop codons or frameshift mutations that lead to nonsense mediated decay of the transcript, leaving a single normal allele that results in a reduced level of EKLF protein (category B1 in Figure 4B). The presence of a mutated EKLF allele is easily detected by the almost complete absence of the Lutheran blood group antigen on the red cells. Other genes are significantly but not as dramatically affected, such as Alas2, AHSP, GPA, and Band3. Beyond this, the In(Lu) phenotype appears to be benign, indicating that one functional EKLF allele is sufficient to sustain human erythropoiesis.

As previously described, a subset of HPFH in a Maltese family is associated with K288X, a nonsense mutation in the EKLF gene that ablates its DNA-binding domain. Degradation of the transcript by nonsense-mediated decay leads to a haploinsufficient output (category B1). Expression profiling on primary erythroid progenitors shows that EKLF target genes are downregulated in samples from individuals with HPFH, among them the BCL11A gene that encodes a suppressor of HbF expression. These observations provide an explanation for the effects of EKLF haploinsufficiency on elevated fetal hemoglobin levels.

A subset of EKLF missense mutations leading to the In(Lu) phenotype yield an EKLF protein that is structurally compromised (category B2, Figure 4B). For example, the H299Y mutation is located at a critical amino acid required for zinc coordination and integrity of the finger structure. Practically speaking, this also generates a haploinsufficient condition, and dose-dependent target genes (such as Lutheran antigen) are most dramatically affected.

An interesting mix of the two subsets of category B is found in a Sardinian family. High levels of fetal hemoglobin and a high level of zinc protoporphyrin are present only in compound herozygotes that contain two EKLF mutations in trans: a S270X nonsense (category B1) and a K332Q missense (category B2) mutation. The K332Q mutation alters the efficiency of EKLF binding to the BCL11A but not the adult β-globin CAC element. Surprisingly, monoallelic, isolated mutations of S270X or K332Q are associated with normal (low) fetal hemoglobin levels.

It is not known what percent HPFH variability is accounted for by EKLF mutation. However, scanning the human genome browser SNP database (dbSNP) fails to find the Table 1 variants, indicating that they are not present in over 1200 alleles analyzed so far in the 1000 genomes project. The accessibility of unpublished EKLF gene variants deposited into the HbVar Database will expand this list and may lead to testable hypotheses related to EKLF structure/function at the β-like globin locus.

The global EKLF network suggests other ways in which EKLF may behave as a genetic modifier. For example, variable levels of AHSP play a quantitative role in the severity of β-thalassemia intermedia, particularly in patients with triplicated α-globin genes. Since AHSP is strongly regulated by EKLF, variation in EKLF levels might attenuate or accentuate the protective effect of AHSP.

**Conclusion**
EKLF/KLF1 plays an essential role in coordinating the intracellular events that occur during erythropoiesis, particularly in directing megakaryocyte/erythroid progenitor lineage decisions, in establishing proper levels of morphological and metabolic proteins, and in its continued requirement during the renewal and differentiative steps of terminal maturation. Coupled with the altered mammalian hematology that follows from mutated or haploinsufficient levels of EKLF/KLF1, these studies suggest that EKLF/KLF1 deficiency may be causative for a range of hematologic disorders.

Acknowledgments

We are grateful to Deepa Manwani, Jane Little, Shefali Soni, and Nithya Gnanapragasam for a critical reading of this manuscript. Referenced studies in the lab have been supported by NIH grants DK46865, DK48721, and DK77822 (JJB).

The authors declare no conflict of interest.

Authorship

MS and JJB wrote the manuscript.

References

10. Guy LG, Mei Q, Perkins AC, Orkin SH, Wall L. Erythroid Kruppel-like factor is essential for beta-globin gene expression even in absence of gene competition, but is not


84. Singleton BK, Burton NM, Green C, Brady RL, Anstee DJ. Mutations in EKLF/KLF1 form the molecular basis of the rare blood group In(Lu) phenotype. *Blood.* 2008;112(5):2081-2088.


Table 1. Hematologic disorders associated with mutations within the EKLF/KLF1 locus.

<table>
<thead>
<tr>
<th>Hematologic disorder</th>
<th>Monoallelic mutations in human EKLF</th>
<th>Phenotype</th>
<th>Ref</th>
</tr>
</thead>
</table>
| Congenital Dyserythropoietic Anemia (CDA) | E325K substitution in conserved residue in Znf2 | -hemolytic anemia  
-elevated HbF  
-morphological abnormalities of erythroblasts in bone marrow  
-nucleated orthochromatic erythroblasts in peripheral blood  
-abolished expression of water channel AQP1, CD44 | 77,82 |
| Blood type disorder In(Lu) phenotype      | L127X, K292X premature stop codons  
H299Y, R328I, R328H, R331G substitution in conserved residue in Znf1  
A base deletion in the codon for P190–L frameshift X47  
A base insertion in the codon for R319–G frameshift X34 (-124 T>C) in promoter abolished a GATA1 binding site | -lack of Lutheran blood group antigen  
-level of HbF not determined | 81   |
| Hereditary Persistence of Fetal Hemoglobin (HPFH) | M139L, K288X neutral substitution (?)  
premature stop codon  
S270X, K332Q substitution in conserved residue in Znf2 (in trans biallelic mutations - compound heterozygosity) | -elevated HbF (3-20%)  
microcytosis  
-elevated HbF (22-30.9%)  
elevated of zinc protoporphyrin | 97, 85 |
| Neonatal Anemia (Hereditary Spherocytosis) | E339D (Nan) substitution in conserved residue in Znf2 | -hemolytic anemia  
-spherocytosis  
elevated embryonic globins  
elevated zinc protoporphyrin  
-splenomegaly  
-iron deposition in the liver, kidney, spleen | 14, 88 |
Figure legends

Figure 1. A schematic diagram of EKLF protein. Shown are N-terminal proline-rich and C-terminal (C2H2 zinc fingers; ZnF) DNA-binding domains. EKLF activity is modulated by post-translational modifications that promote protein-protein interactions. The transactivation activity of EKLF (upper part of the figure) is associated with the phosphorylation status of threonine 41 and acetylation of lysine 288. Acetylated K288 is involved in interactions with the chromatin remodeling SWI/SNF-related complex (E-RC1) through the BRG1 subunit. The repression activity of EKLF (lower part of the figure) relies on an acetylation of K302, enabling its interaction with co-repressors Sin3a and HDAC1. Sumoylation of K74 serves to tether components of NuRD repression complex via interaction with Mi2B and HDAC1. EKLF is also a substrate for ubiquitination; however this modification is not specific to any particular internal lysine.

Figure 2. Role of EKLF in erythroid lineage development. At the MEP (megakaryocyte-erythroid progenitor) stage, EKLF represses megakaryopoiesis while at the same time it promotes erythropoiesis. This process likely relies, in part, on cross-antagonizing interactions between EKLF and Fli-1 (a crucial megakaryocyte transcription factor). Committed erythrocyte progenitors expand through renewal divisions that occur during fetal erythropoiesis or following hypoxic stress. Terminal maturation of erythrocytes occurs throughout differentiation divisions, where cells undergo many changes: size reduction, hemoglobin accumulation, decrease in overall gene expression level, nuclear condensation and ultimate extrusion. These events are particularly dependent on the presence and, in some cases, on the level of EKLF.

Inset – The progression of erythroid development in wild type and EKLF-deficient fetal liver cells as monitored by presence of the cell surface markers CD71 and Ter119. Discrete populations of cells (R1-R5) with distinct morphologies can be distinguished that correspond to particular stages of cells during erythropoiesis as described in the text (figure taken from).

Figure 3. EKLF regulates globin switching. During embryonic and fetal development or in EKLF haploinsufficient adults (left), EKLF levels are relatively low, resulting in low levels of adult β-globin and Bcl11a and high levels of γ-globin. In adults with two functional copies of EKLF (right), increased expression of EKLF in definitive red blood cells promotes high levels of adult β-globin and Bcl11a expression, the latter of which in turn represses γ-globin expression. Modified from.

Figure 4. Mechanisms linking mutations in the EKLF/KLF1 gene with the impact that they have on erythrocyte phenotypes. The flow of gene expression from each of the two EKLF/KLF1 alleles at the top is shown as it progresses through transcript and protein expression, ultimately leading to the resultant phenotype and disorder as described in the text. Selected examples from Table 1 are shown at the bottom that illustrate each mechanistic subtype. (A) A mutation (red) changes a functionally important amino acid (aa) and leads to an EKLF protein with altered properties that even in the presence of the WT allele causes a dominant mutant phenotype on a select subset of target genes. (B) Haploinsufficiency of EKLF follows as a consequence of: B(1) a mutation (green) that causes premature stop codon in one allele. The defective transcript undergoes nonsense mediated decay, leaving transcript from only a single WT allele to be translated into functional protein. B(2) a mutation (yellow) changes a structurally important amino acid and results in expression of a functionally impaired protein that decreases the pool of active EKLF to haploinsufficient levels.
Positions of the mutations in the schematic are symbolic and do not relate to any precise amino acids.
Figure 1
Figure 2
Figure 3
Figure 4

A

EKLFL/KLF1 gene

- substitution of a conserved, functionally important aa

transcripts

- mutation causes a premature stop codon

proteins

- substitution of a conserved, structurally important aa

Protein with altered properties

Dominant mutant

phenotype of erythrocyte

Haploinsufficiency

CDA
(EKLFL-E325K)

Nan
(mEKLFL-E339D)

Hematopoietic disorder (selected examples)

HPFH – Maltese family (EKLFL-K288X)

Rare blood group In(Lu) phenotype (EKLFL-H299Y)

HPFH – Sardinian family (EKLFL-S270X, K332Q, in trans)
The multifunctional role of EKLF/KLF1 during erythropoiesis

Miroslawa Siatecka and James J. Bieker