The DEK oncoprotein is a critical component of the EKLF/KLF1 enhancer in erythroid cells

Felix Lohmann  
Mount Sinai School of Medicine

Mohan Dangeti  
Mount Sinai School of Medicine

Shefali Soni  
Mount Sinai School of Medicine

Xiaoyong Chen  
Mount Sinai School of Medicine

Antanas Planutis  
Mount Sinai School of Medicine

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
Lohmann, Felix; Dangeti, Mohan; Soni, Shefali; Chen, Xiaoyong; Planutis, Antanas; Baron, Margaret H.; Choi, Kyunghee; and Bieker, James J., "The DEK oncoprotein is a critical component of the EKLF/KLF1 enhancer in erythroid cells." Molecular and Cellular Biology.35,21. 3726 –3738. (2015).  
https://digitalcommons.wustl.edu/open_access_pubs/4833

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Authors
Felix Lohmann, Mohan Dangeti, Shefali Soni, Xiaoyong Chen, Antanas Planutis, Margaret H. Baron, Kyunghee Choi, and James J. Bieker
The DEK Oncoprotein Is a Critical Component of the EKLF/KLF1 Enhancer in Erythroid Cells

Felix Lohmann,a* Mohan Dangeti,a Shefali Soni,a Xiaoyong Chen,a* Antonas Planutis,a Margaret H. Baron,a,b,c,d Kyunghee Choi,e James J. Bieker,a,b,c

Department of Developmental and Regenerative Biology,a Black Family Stem Cell Institute,b Tisch Cancer Institute,c and Department of Medicine,d Mount Sinai School of Medicine, New York, New York, USA; Department of Pathology, Washington University School of Medicine, St. Louis, Missouri, USA*e

Understanding how transcriptional regulators are themselves controlled is important in attaining a complete picture of the intracellular effects that follow signaling cascades during early development and cell-restricted differentiation. We have addressed this issue by focusing on the regulation of EKLF/KLF1, a zinc finger transcription factor that plays a necessary role in the global regulation of erythroid gene expression. Using biochemical affinity purification, we have identified the DEK oncoprotein as a critical factor that interacts with an essential upstream enhancer element of the EKLF promoter and exerts a positive effect on EKLF levels. This element also binds a core set of erythroid transcription factors, suggesting that DEK is part of a tissue-restricted enhanceosome that contains BMP4-dependent and -independent components. Together with local enrichment of properly coded histones and an open chromatin domain, optimal transcriptional activation of the EKLF locus can be established.

Studies of the erythroid lineage have led to the successful characterization of intracellular regulators that act as transcription factors to generate red-cell-specific expression (1). However, in many cases it remains unresolved how these factors are themselves regulated. Based on this notion, we have been studying the regulation of erythroid Krüppel-like factor (EKLF or KLF1 [2]), a zinc finger hematopoietic transcription factor that plays a global role in activation of genes critical for genetic control within the erythroid lineage (3–5). It performs this function by binding to its cognate DNA 5′CCMRCCCCN3′ element and recruiting chromatin-remodeling proteins and histone modifiers.

Regulation of EKLF itself is of interest because of a number of functional properties and expression characteristics. EKLF expression remains tissue specific throughout early development and in the adult. Its onset in the yolk sac is strictly limited to the mesodermal, primitive erythroid cells that populate the blood islands at the early headfold stage (embryonic day 7.5 [E7.5]), switching by E9.5 to definitive cells within the hepatic primordia and then to the red pulp of the adult spleen and the bone marrow (6). During definitive hematopoietic differentiation, EKLF is expressed at low levels in multipotent progenitors (MPP) and retains an expression pattern restricted to the common myeloid progenitor (CMP) and megakaryocyte erythroid progenitor (MEP) prior to eventual segregation to erythroid progeny (7–9).

We have demonstrated that a 950-bp region adjacent to the EKLF start site of transcription, encompassing two critical erythroid hypersensitive sites (EHS1 and -2) and the proximal promoter (10), contains all the information needed for developmentally regulated, blood-cell-specific expression of a linked reporter gene that faithfully recapitulates the onset of endogenous KLF1 expression after GATA1 protein is produced (15). Chromatin immunoprecipitation (ChiP) of GATA proteins reveals a switch from GATA2 to GATA1 in Gata site occupancy (15). In addition, the use of a doxycycline-inducible small hairpin RNA (shRNA) line directed against Sma5 verified its critical importance for KLF1 expression. Together with the promoter analyses, these data led to the proposal of a two-tiered mechanism for transcriptional regulation of KLF1, with GATA2 and SMAD5 proteins initially generating low transcript levels, followed by upregulation of KLF1 expression after GATA1 protein is produced (15).

The slow kinetics of BMP4 induction of EKLF (20) raised the possibility that other factors are corequired prior to EKLF onset. Therefore, we focused on isolating proteins that interact with the particularly strong enhancer element located within EHS1. Using a biochemical affinity purification approach, we have identified dependent on and regulated by growth factor signals. In this context, BMP4 has been shown to play a necessary role in the induction of EKLF transcription, and the BMPR/Smad pathway is critical (20).

The importance of the Gata and Sma5 elements in EHS1 and the proximal promoter has been verified by mutational analyses using a 950-bp KLF1 promoter/green fluorescent protein (GFP) transgene that faithfully recapitulates the onset of endogenous KLF1 expression during mouse embryo body (EB) differentiation (15). Chromatin immunoprecipitation (ChiP) of GATA proteins reveals a switch from GATA2 to GATA1 in Gata site occupancy (21, 22) when comparing early and late times of EB differentiation (15). In addition, the use of a doxycycline-inducible small hairpin RNA (shRNA) line directed against Sma5 verified its critical importance for KLF1 expression. Together with the promoter analyses, these data led to the proposal of a two-tiered mechanism for transcriptional regulation of KLF1, with GATA2 and Sma5 proteins initially generating low transcript levels, followed by upregulation of KLF1 expression after GATA1 protein is produced (15).

The slow kinetics of BMP4 induction of EKLF (20) raised the possibility that other factors are corequired prior to EKLF onset. Therefore, we focused on isolating proteins that interact with the particularly strong enhancer element located within EHS1. Using a biochemical affinity purification approach, we have identified
one of these factors as the DEK oncoprotein (23). DEK was originally identified as a fusion with the CAN (now called Nup214 [24]) protein in a subset of patients with acute myeloid leukemia (AML) who carry a t(6;9)(p23;q34) translocation (25). We now show that DEK plays a critical role within the protein complex at EHS1 and enables optimal levels of EKLF transcription.

MATERIALS AND METHODS

Cell culture. Murine erythroleukemia (MEL) line 745A, 32DEpo1, and 293T cells were grown as previously described (10, 26).

Purification of EHS1-binding proteins. Competitive gel shifts were performed with oligonucleotide 2 (oligo2), oligo3, and oligo4 (10) to help us determine which one to select for purification of EHS1-binding proteins. Use of competitive gel shifts continued during our determination of the optimal purification procedure and in the subsequent follow-through (e.g., see Fig. 2C). M280 magnetic Dynabeads coupled with streptavidin (Life Technologies) were incubated with biotinylated oligonucleotides to create the oligonucleotide Dynabeads using buffers, washes, and conditions as recommended by the manufacturer's protocol. oligo1 and oligo2 sequences have been described previously (10). The 32DEpo1 cell lysate (prepared as described in reference 10) was added, and incubation was continued for 2 to 3 h at room temperature with rotation. A magnetic apparatus was used for all washes. The beads were sequentially washed in 0.2 M NaCl (which removes most of the nonspecific binding proteins) and 0.5 M NaCl before elution in 1 M NaCl. Eluted protein was dialyzed and concentrated by acetone precipitation prior to SDS-PAGE. Proteins were visualized by staining with colloidal blue, and bands of interest were excised. Mass spectrometry analyses of the gel slices were performed by the Rockefeller University Protein and DNA Technology Center using the Sonar MS/MS search engine coupled with statistical scoring methods.

Electrophoretic mobility shift assay. EHS1 DNA oligonucleotides were fill-in labeled with [32P]dCTP, incubated with purified His-tagged human recombinant DEK protein (pET-28a vector [28]) or with 32DEpo1 extracts, and analyzed on 5% or 8% native polyacrylamide gels as described previously (10, 26, 29). We also performed purified DEK DNA binding reactions in 20 mM HEPES (pH 7.9), 60 mM KCl, 3% Ficoll, 0.5 mM MgCl2, 0.06% Nonidet P-40, 1 mM dithiothreitol, and 20 ng of double-stranded poly(dl-dC). Mouse oligo1 and oligo2 sequences were as described previously (10); human oligo2 (5′-AGCTTCTAGCTGGCCTGGGCCCC) and scrambled (5′-AGCTGCGCTGGTGTACAGTAGTAG) were also used (the top strand only is shown; the EKLF sequence is underlined). The DEK10 sequence was 5′-AGCTTATGGATTATATAGCACCC (the italicized bases were changes from the mouse wild type).

shRNA knockdowns. A 21-nucleotide (nt) stretch within the mouse DEK mRNA sequence was identified according to the RNA interference (RNAi) target design rules put forth by reference 30 (bp 392 to 412 of mouse DEK mRNA, NCBI accession no. NM_025900) and incorporated into a double-stranded DNA 70-mer to encode an anti-DEK small hairpin RNA (shRNA) according to the design rules described in reference 31: 5′-AGATCCCCCTGGTTTACAACCGGGGCTTACAGGACCCCTGTTGGAAGCTTATTATTAGACCC. This 70-mer was cloned into the pSUPER vector via its BgIII and HindIII sites and verified by sequencing. The resulting plasmid, the empty vector (pSUPER), or a GFP-expressing control plasmid (carrying a neomycin resistance gene), or a GFP-expressing control plasmid (carrying a neomycin resistance gene) was transfected with pXM330/DEK plasmid in 5 bases of the DEK target sequence that altered the most conserved residues. MEL cells were cotransfected with the pXM330/DEK plasmid, the donor polynucleotide, and a red fluorescent protein (RFP)-expressing plasmid using the Neon transfection system (Life Technologies) set at 1,450 V and 3 pulses with a bandwidth of 10 ms. RFP-expressing cells were single cell sorted with a FACSAria II (Becton Dickinson) into 96-well dishes for growth. Genomic DNA from these clonal populations was isolated and directly sequenced across the region of interest. Because nonhomologous end joining is more efficient than homologous recombination, most of the clones contained deletions in one or both alleles at the DEK binding site even though we had included donor DNA; we selected one clone with an identical deletion in both alleles (inde8) and one with a perfectly recombined target (DEK10) for analysis.

Data sets analyzed. Databases used as part of this study include histone modification (49–51), transcription factor (49, 50, 52–54), and erythroid expression (55) sources and are indicated within the relevant figure legends.
RESULTS

The conserved EKLF upstream enhancer element is the site of epigenetic modulation. We have shown by cotransfection and in vivo analyses that the small 950-bp region adjacent to the EKLF transcription start site is sufficient for tissue-restricted expression (10, 11) and wished to address whether the epigenetic profile of this region supports its importance. Database perusal of modified histone H3 interactions within the critical 950-bp region of the EKLF promoter/enhancer during hematopoiesis (51, 56) shows that levels of the H3K4me1 active enhancer mark increase exactly in parallel with activation of EKLF expression during hematopoiesis (7), that is, low within the hematopoietic repopulating cells (long-term [LT] and short-term [ST] hematopoietic stem cells [HSC]) and in myeloid (CMP and MPP) progenitors but high in MEPs and erythroblasts (Fig. 1A). High levels of the active H3K4me2 and H3K27ac marks are also evident in this region of the EKLF promoter in erythroblasts (57, 58), as is a wider H3K4me3 mark (59) within the body of the gene (Fig. 1A). Other erythroid database analyses (49, 50) support this pattern and additionally show that the active H3K4me2 and H3K27ac marks are also evident in this region of the EKLF promoter in erythroblasts (57, 58), as is a wider H3K4me3 mark (59) within the body of the gene (Fig. 1A).

Isolation of proteins that bind the core fragment in EHS1. Within the evolutionarily conserved 950-bp region adjacent to the EKLF transcription initiation site, our in vitro analysis of the 49-bp EHS1 region (10) found unique DNA binding proteins when oligonucleotides spanning this region were used in gel shift assays with extracts from 32DEpo1 erythroid cells. We had originally used these cells because of their high transfection efficiency in reporter assays; as a result, we retained the same cells for the biochemical purification of binding proteins. We focused on the 18-bp oligo2 binding activity (Fig. 2A) because its DNA/protein complex was more resistant to dissociation than that formed with oligo3, there was not any detectable complex formed with oligo1, and oligo3/oligo4 span two Gata factor sites already known to bind the Gata1-Scl/Tal1 complex (14).

In our approach, double-stranded oligonucleotide 2 (or oligonucleotide 1 as a negative control) is 5’ biotinylated at one end and attached to a streptavidin/magnetic bead column, which provides the solid-state matrix to which the protein(s) is to be bound (60, 61). This column was sequentially washed in 0.2 M NaCl (which removes most of the nonspecific binding proteins) and 0.5 M NaCl before elution in 1 M NaCl, which contains almost all of the specific DNA binding responsible for the oligonucleotide 2 probe shift in the crude extract. Elution is efficient, as active DNA binding material is recovered from the oligonucleotide 2 column but not from the oligonucleotide 1 column (Fig. 2B and C), and a second elution does not recover any additional activity (not shown). Based on DNA binding activity and protein estimates, this step represents an ~5,000-fold purification.

We performed two analyses to ascertain whether any proteins had been sufficiently enriched for subsequent analysis by this procedure. First, we prepared extracts from [35S]methionine-labeled
cells and performed the affinity purification. Analysis of labeled proteins indicates that two proteins (molecular masses of ~60 and 33 kDa) are uniquely present in the oligonucleotide 2-purified fraction (Fig. 2D), demonstrating that we can well resolve the proteins that are specific to oligonucleotide 2. Second, we scaled up the protocol and monitored a portion of the eluted material on a higher-percentage, silver-stained SDS gel. This higher-resolution analysis confirmed that the two proteins seen previously are uniquely present in the oligonucleotide 2-derived material compared to that from oligonucleotide 1 (not shown).

We purified protein from ~10 liters of 32DEpo1 cells, performed matrix-assisted laser desorption ionization (MALDI-TOF) and tandem mass spectrometric analysis, and focused on the material that migrated at ~35 kDa and that is uniquely present in the oligonucleotide 2-derived preparation. This preparation (but not the negative control) yields the expected gel shift with a labeled

FIG 2 oligo2 binding proteins within EKLF EHS1. (A) (Top) Schematic (15) of conserved Eklf cis-regulatory elements (upstream enhancer and proximal promoter) together with erythroid hypersensitive sites (EHS1 and EHS2), the Gata-E box-Gata site (GEG) at EHS1, the Gata and Cp1 sites at the proximal promoter, and potential Smad binding motifs (SBM). (Bottom) Detail of EKLF upstream enhancer (EHS1) region. Blocks of conserved sequence homology between seven mammalian species are color coded for their transcription factor binding sites (if known); locations of oligo1 through 4 sequences as identified by reference 10 within EHS1 are as indicated, along with evidence of their in vitro ability to bind proteins in cell extracts. (B) Enrichment of oligo2 binding protein after affinity chromatography with an oligo1 or oligo2 (3 samples) probe as indicated was monitored by a gel shift assay with radioactively labeled oligonucleotide 2. Input is nonenriched 32DEpo1 cell extract. (C) Competitive gel shifts were performed during the purification to assess specificity of binding by extracts to oligonucleotide 2. Specific (S) or nonspecific (NS) competitor DNA was included at a 100-fold excess as indicated. (D) DNA binding proteins that interact with oligo2 were monitored by SDS-PAGE analysis of [35S]methionine-labeled proteins after oligo1 or oligo2 affinity purification. Purified material derived from ~5 × 10⁶ cell equivalents is shown. The arrow marks the protein unique to the oligonucleotide 2 preparation that is the focus of this study, along with the mass spectrometry data (Sonar MS/MS analysis; a:b:y ratio is of the fragmentation ions; m/z is the measured mass of the peptide fragment [27]) and peptide sequences used in its identification; molecular masses are indicated in kilodaltons on the left.
Using the mass spectrometry data, software tools enabled appropriate database searches and determination of which subset of peptides gives the specified ion fragments (62). Of particular interest are two peptides that match the mammalian DEK oncoprotein with high significance (Fig. 2D). DEK is a protein of particular relevance to the present studies as it is predicted to be involved in the induction of mesoderm during early embryonic development (63), and it plays a role in a subset of myeloid leukemia as a translocation protein fusion product with Nup214 (25). Database inspection shows that all DEK exons are strongly expressed in primitive and definitive (fetal liver or bone marrow) differentiation. P, proerythroblasts; B, basophilic erythroblasts; O, orthochromatophilic erythroblasts; R, reticulocytes. (B) Semiquantitative RT-PCR analysis of DEK expression was performed with RNA from staged whole embryos at E6.5 (S, streak) and early or late E7.5 (B, bud) as indicated; water was used as a negative (neg) control; 32DEpo1 RNA was used as a positive (pos) control. Results are representative of two experiments. (C) Levels of EKLF (left scale) or DEK (right scale) were measured in RNA from EBs harvested at the indicated day of differentiation using quantitative RT-PCR and normalized to glyceraldehyde-3-phosphate dehydrogenase. Values are presented relative to day 0, which is set to 1. Results are the averages of triplicates from a single experiment that is representative of two experiments.

oligonucleotide 2 probe. Using the mass spectrometry data, software tools enabled appropriate database searches and determination of which subset of peptides gives the specified ion fragments (62). Of particular interest are two peptides that match the mammalian DEK oncoprotein with high significance (Fig. 2D).

DEK is a protein of particular relevance to the present studies as it is predicted to be involved in the induction of mesoderm during early embryonic development (63), and it plays a role in a subset of myeloid leukemia as a translocation protein fusion product with Nup214 (25). Database inspection shows that all DEK exons are strongly expressed in primitive and definitive erythrocytoid cells at all stages, including erythrocytoid progenitors and the pro-erythroblast, with an expression pattern that parallels that of EKLF; although it dips in the reticulocyte, this remains at a relatively high level (Fig. 3A). In the gastrulating embryo, DEK expression increases dramatically between E6.5 and early E7.5 (Fig. 3B), just prior to or coincident with the onset of EKLF at E7.5 (6), at the time when the blood islands of the yolk sac are first apparent. DEK onset is also coincident with that of EKLF during EB differentiation at a time when hematopoiesis initiates (Fig. 3C). As a result, DEK is appropriately present to exert an effect on EKLF expression, both in the erythrocytoid cell and during early development and differentiation.

**DEK binding to its target site is important for optimal EKLF expression.** Sequence inspection indicates that oligonucleotide 2, although not absolutely conserved across its entire length, contains a core region that is identical across KLF1 promoters from seven vertebrate species (Fig. 4A). In addition, this core region matches nicely (only one mismatch) with the HIV-2 enhancer peri-ets site (pets site) bound specifically by DEK in vitro (28) (Fig. 4A). We used in vitro DNA binding assays to test whether recombinant DEK protein can bind oligonucleotide 2 sequences derived from mouse or human EKLF genomic DNA. These sequences are not identical (5/18 bp are different), but as shown in Fig. 4B, each interacts specifically with DEK in vitro as monitored by gel mobility shift experiments. We addressed by chromatin immunopre-
DEK Regulation of KLF1 Expression

FIG 4 DEK is critical for EKLF expression. (A) DEK consensus motif. Multispecies EKLF EHS1 oligo2 shares significant similarity with the peri-ets (pets) site bound by DEK in the HIV-2 enhancer. (B) DEK protein interacts with EHS1 in a sequence-specific manner in vitro. DEK binding to oligonucleotide 2 was monitored by gel shift analysis of purified recombinant DEK protein with labeled human (Hs) (lanes 1 to 3) or murine (Mm) (lanes 4 to 7) oligonucleotide 2. Incubations were neat (lanes 1 and 5) or included self (lanes 2 and 6) or scrambled (lanes 3 and 7) cold oligonucleotide 2 at a 100-fold excess. The arrow indicates the band of interest. (C) Quantitative analysis of chromatin immunoprecipitation of DEK protein at the EKLF EHS1 (upstream enhancer) region. Isotype IgG was used as a negative control. Results are the averages of triplicates from two experiments. (D) MEL cells were stably transfected (sh1) or infected (sh2 and sh3) with DEK shRNA-expressing constructs and their appropriate empty vector controls (V) or a nonspecific gene (NS). Untransfected MEL cells (745A) served as an additional control. Results of Western blotting of extracts probed with the indicated antibodies are shown; Hsp90 was used as a loading control. (E) RNA isolated from the same experiment as in panel D was semiquantitatively analyzed for expression of DEK, EKLF, and Gata1 expression; hypoxanthine phosphoribosyltransferase served as a control. Note that sh1 must be exerting its effect by translational repression (114), as DEK RNA levels are not affected in spite of effective protein knockdown.

DEK mRNA was targeted for degradation via RNAi. Using two different vectors (pSUPER [31] and pLKO.1 [32]) with either transfection/selection- or infection/selection-based approaches, we monitored the effects of stably expressing three DEK shRNAs in MEL (line 745A) cells (shDEK lines). Expression analyses from two of the three shDEK lines (but not the empty vector or the nonspecific control lines) show a strong decrease of expression of EKLF protein, which correlates with the loss of DEK protein (Fig. 4D). EKLF but not Gata1 RNA levels are decreased (Fig. 4E).

Collectively, these studies demonstrate that DEK is required for EKLF expression and directly interacts in vitro and in vivo with its cognate site in the EKLF EHS1 enhancer.

Relevance of BMP4 pathway to transcription factor expression. Studies of the EKLF promoter have implicated the BMPR/BMP/Smad pathway in its regulation (15, 18, 20, 64). This conclusion was supported by the demonstration that BMP4 was necessary and sufficient to induce EKLF during serum-free ES cell differentiation and that interference with the pathway by dominant negative BMPR1B or constitutive Smad6 expression prevented EKLF expression even during serum-containing ES cell differentiation (20). At the same time, the Smad complex has been shown to play a direct role in EKLF induction via Smad5 (15,18). However, it takes 2 to 3 days for BMP4 to induce EKLF in differentiating EBs (20). The delayed nature of this effect led us to postulate the presence of a required protein that either is synthesized independently or is codependent on BMP4 induction.

To examine whether DEK might fit into such a scheme, we used the Flk1<sup>+/Cre</sup> (65) and the floxed BMPRIA (Alk3) and Smad4 (44) mouse lines to address whether DEK and EKLF expression is altered specifically within the Flk1<sup>+</sup> population of the developing embryo after conditional ablation. Flk1 is expressed by...
E7.0 in the extraembryonic and paraxial-lateral embryonic mesoderm (66, 67), and Flk1+ cells are critical for formation of blood vessels and hematopoietic cells (68, 69). Embryos that result from a cross between Flk1+Cre and floxed Alk3 or Smad4 mice die by E11 (44). RNA was isolated from either E10.5 or E9.5 yolk sacs and analyzed by quantitative RT-PCR. Levels are normalized to glyceraldehyde-3-phosphate dehydrogenase and then set to "1" for wild type. Results are averages from two to three independent biological replicates performed each in triplicate.

**FIG 5** Effects of yolk sac (YS)-specific ablation of Alk3 or Smad4 on selected targets. EKLF (left), Gata1 (middle), and DEK (right) transcript levels were monitored in yolk sac cells derived from Flk1-Cre (Flk1+Cre) × Smad4 (wild-type or fl/fl) E9.5 embryos (bottom) or Flk1-Cre (Flk1+Cre) × Alk3 (wild-type or fl/fl) E10.5 embryos (top). Samples were processed for RNA isolation and analyzed by quantitative RT-PCR. Levels are normalized to glyceraldehyde-3-phosphate dehydrogenase and then set to "1" for wild type. Results are averages from two to three independent biological replicates performed each in triplicate.

To test whether the 18-bp oligonucleotide 2 site plays a role in maintenance of this complex, we used a CRISPR/Cas9 gene-editing approach (47, 48) to modify the endogenous oligonucleotide 2 site in MEL cells. This resulted in a series of indels of various extents, of which one (indel8) contains a biallelic deletion that overlaps oligonucleotide 2 (Fig. 6C). We investigated the status of DEK and two of the core erythroid components in this line and found that the low level of DEK interaction with the oligonucleotide 2-edited EHS1 site was mirrored by a dramatic decrease in occupancy by P300 and Tal1 (Fig. 6D).

The DEK consensus target site within oligonucleotide 2 is fairly loose (Fig. 4A and B), and its DNA recognition properties are unusual (see Discussion), making a directed mutagenesis design less straightforward. Nevertheless, we used CRISPR/Cas9 under homologous recombination conditions to establish a MEL cell line with five substitutions within the most conserved nucleotide residues (DEK10 [Fig. 6C]). Interestingly, although DEK binding is not affected, binding by P300 and Tal1 is decreased by ~70% (Fig. 6D), suggesting that the core erythroid complex does not remain intact in the midst of this DEK-directed mutation. Direct testing by in vitro analysis supports the lack of effect of the mutated DEK10 site on DEK binding, which maintains specificity, as it is not competed by a scrambled oligonucleotide 2 variant (Fig. 6E).

We next addressed the importance of the oligonucleotide 2 sequence on EKLF expression level in the edited MEL lines, indel8 and DEK10, where we find a small (~40%) but significant effect on expression, at both the RNA and protein levels (Fig. 7A). This quantitative effect was further supported by analysis of the previously characterized ES cell line containing the 950-bp EKLF promoter linked to a GFP reporter that faithfully reproduces endogenous activation (15). We precisely deleted the oligonucleotide 2 sequence within the promoter and quantified expression levels during embryoid body formation. The data reveal a 2-fold drop in promoter-directed expression in the oligonucleotide 2-deleted cells (Fig. 7B).

Together, these results demonstrate that the oligonucleotide 2 sequence is vital for formation of the critical core erythroid enhancosome complex within EHS1 that minimally includes Tal1, P300, and DEK (Fig. 8). Although removal of the 49-bp EHS1 site has a dramatic effect in vivo, removal of the DEK/oligo2 element yields a more nuanced effect on adjacent gene expression and yet remains necessary to establish optimal levels of EKLF RNA and protein expression.

**DISCUSSION**

We have used a biochemical approach to identify the DEK oncoprotein as a critical factor for EKLF expression via its interaction with the 18-bp oligonucleotide 2 region of the 49-bp EHS1 enhancer element. Removal of the oligonucleotide 2 target site yields a significant disruption of the core erythroid complex at EHS1, and removal of DEK protein (via shRNA constructs) has a dramatic effect on EKLF expression. Our studies suggest that DEK forms a critical part of the core erythroid complex that is formed at EHS1.

**Unusual and relevant regulatory properties of DEK.** DEK was first identified as a fusion protein expressed from the (t(6; 9)(p23;q34) translocation found in a subset of patients with AML (25; see references 23 and 77 to 79 for recent reviews). The DEK gene encodes a 375-amino-acid nuclear protein whose only structural relatedness to known proteins is a well-conserved SAP do-
main (35, 80, 81), which forms one of its two DNA binding domains (82). DEK plays a role in gene control via its binding of nucleic acids, although the specific nature of these interactions remains complex. For example, while there is evidence that it may play a non-sequence-specific scaffolding function leading to induced alterations of chromatin topology (81, 83, 84), there is also evidence that it recognizes specific target DNA sequences, notably within the HIV-2 enhancer peri-ets sequence (28, 85) and in the class II major histocompatibility complex (MHC) gene promoter at its Y box sequence (86). These target sequences are activation elements and are both homologous to the oligonucleotide 2 sequence of the EKLF promoter (Fig. 4).

The association of DEK with Daxx and histone deacetylase (HDAC) suggests that it plays a role in gene repression (87) and in maintaining heterochromatin integrity via HP1/H9251 (36). However, and of most relevance to the present studies, DEK interacts with AP-2/H9251 and stimulates transactivation (88) and also interacts with C/EBPα, an important coactivator of myeloid target genes (89), where DEK knockdown disrupts granulocyte differentiation. DEK, accompanied by AP-2α recruitment, becomes enriched at a site proximal to the CD21 promoter upon its transcriptional activation in B lymphocytes (90). DEK plays a critical role in enabling activation of myogenic satellite cells and their exit from quiescence (91) and binds to highly expressed genes within open chromatin in U937 cells (92). Drosophila DEK is found at edcsy6s-induced puffs and is a coactivator for the edcsy6s receptor (93). Although the levels of hematopoietic stem cells (HSC) and progenitors are slightly but significantly enhanced in the absence of DEK, engraftment of HSC is curtailed (94).

Given the divergence in functions attributed to DEK, no direct connection to the regulation of the EKLF gene was apparent at first. However, the pets site motif that is occupied by DEK in the HIV-2 enhancer is almost identical to the oligo2 sequence of EHS1, which in turn is highly conserved between species. Furthermore, DEK is predicted to play a role in the induction of mesoderm during early embryonic development (63). Together, these properties provided us with a compelling reason to more fully investigate DEK as a functional component of the EKLF promoter, specifically via an interaction with the oligo2 region of EHS1. Based on its expression pattern during erythroid differentiation and development, on the activating role of its target DNA site, and on the negative effects of its knockdown, our findings are most consistent with a stimulatory function for DEK during erythropoiesis, similar to its role in myeloid gene activation.

As summarized in Fig. 8, DEK onset occurs independently of the BMP4 pathway, suggesting that its expression is separate but possibly coordinately required with BMP4 activation of Smad5 to fully induce EKLF. We have previously proposed that Gata2 is required for EKLF’s initial activation, which is then activated to higher levels once Gata1 is fully induced by Gata2 (15). An intact Gata1 protein is required for full activation of EKLF, as Gata1s is deficient (95). The requirement of the BMP4 pathway for Gata1
FIG 7 Importance of the oligonucleotide 2 sequence for optimal EKLF expression. (A) Total RNA was isolated or whole-cell extracts were prepared from wild-type (WT), indel8, or DEK10 MEL cells and analyzed for expression of EKLF RNA (quantitative RT-PCR; analysis of 11 analytical replicates each from two experiments) or protein (Western blotting) as indicated. “C” is an additional loading control to balance the EKLF protein signals across the gel. (B) (Top) Schematic of constructs containing the KLF1 promoter or one that has the oligonucleotide 2 sequence removed that were stably integrated into mouse ES cells adjacent to a GFP reporter, yielding two stable ES lines (“Peklf-GFP” and “Peklf-1xΔDek-GFP”). These are single-copy, unidirectional integrations into the same homing site, which thus avoids position effect variegation. Locations of the mapped upstream enhancer, proximal promoter, and intronic enhancer are as indicated. (Middle and bottom) ES cells from each were differentiated to EBs for the indicated number of days, and samples were quantitatively analyzed for expression of endogenous KLF1 (red) or exogenous reporter (green). The data show that the two cell lines differentiated and expressed endogenous KLF1 normally but that the Peklf-1xΔDek-GFP line generated a significantly lower relative level of reporter than did the Peklf-GFP line. Data are from three experiments analyzed in triplicate.

expression suggests a coherent type 1 feed-forward mechanism on EKLF induction between BMP4/Smad5 and Gata1 once Gata1 is induced by Gata2 (96, 97). Based on our earlier studies, this requires high levels of BMP4 (20).

Extended protein/DNA complex interactions. The oligonucleotide 2 target site is situated within the EKLF EHS1 enhancer region (Fig. 8) that is known to interact with “core” erythroid transcription components that consist of Gata1 or -2, Tal1, Ldb1, and P300 (56, 72–75, 98). Of relevance to the present study, a recent analysis shows that the accuracy of enhancer prediction is improved by including Smad binding to this core set (99). The sequence and protein layout are highly reminiscent of the beta interferon (IFN-β) enhansome (100, 101), a conserved 50-bp sequence with overlapping binding sites that associate with a specific set of transcriptional activators. Using DNA- and protein-dependent mechanisms, these cooperate to synergistically activate transcription. CBP/P300 is a strong coactivator that interacts with many of these regulators and stabilizes the complex. Supporting this analogy, global analyses of blood-cell-specific expression demonstrate that motifs with preferential spacing are prevalent in hematopoietic promoters and that they interact with multiple transcription factors (102). The spacing among these motifs is critical for maximal activity (102); for example, the Gata/E box motif pair retains an ~9-bp spacing, a property also seen within EHS1. Gata2 and Smad4 (the partner of Smad5) form part of a combinatorial interaction network (103), and we also find that DEK and Smad5 can interact after cotransfection (unpublished observations).

Our studies suggest that DEK is a critical non-tissue-specific component that supports formation of a complex structure of core erythroid activators at selected promoters/enhancers that contain Smad/Gata/E box/Gata DNA sequences (73–75, 98). DEK may or may not be needed for the integrity of the other components of the complex; the ChIP data (Fig. 6) suggest that it contributes to it but also that the DNA site plays a significant role in its coherence, such that the DEK10 mutant site is just not optimal for maximum transcription efficiency even though it binds DEK. In addition, other proteins may be involved (e.g., CTCF [104]). It will be of interest to query other erythroid sites that also interact with these core components to determine the extent of overlap with DEK.

Although we have focused our attention on the importance of EHS1, it is essential to remember that the proximal promoter is absolutely required for EKLF expression in cell culture and in vivo (10, 11, 15), and there is evidence that a conserved intron 1 element contributes to full activity as well (15). EHS1 and intron 1 may each provide quantitative contributions to expression that together with the crucial proximal promoter element are needed for full EKLF induction in vivo.

How might DEK function be so critical for EKLF transcriptional activation? DEK properties suggest a number of speculative scenarios. For example, DEK’s ability to interact with altered DNA and chromatin structures may come into play within any erythroid “activation hubs” (e.g., see references 105 to 107) that might form at the EKLF genetic region. Given that Scl/Tal1, a protein that is known to play a critical role in the formation of such 3-dimensional structures in erythroid cells via the Ldb complex (108), also binds at EHS1, it is not far-fetched to imagine such a scenario. In addition, DEK interacts with histone H3.3, particularly after its phosphorylation by CK2, and preferentially enriches
H3.3 localization to actively transcribed regions (93). Finally, DEK is found associated with splicing complexes (80, 109, 110), where it enables accurate discrimination of potential 3’ splice sites via its interaction with U2AF (110), thus potentially linking primary transcription to effective accumulation of the processed EKLF transcript (discussed in references 111 to 113).

ACKNOWLEDGMENTS

Chunhong Gong, Changwon Park, and Deanna Mohn were involved in different phases of this study. We thank Nithya Gnanapragasam for help with fluorescence-activated cell sorting. We thank David Markovitz and Ferdinand Kappes for DEK plasmids and antibodies. This work was supported by NIH grants R01 DK48721 and DK46865 to J.J.B., R01 HL55337 to K.C., and R01 HL62248 to M.H.B.

REFERENCES


