KRÜPPEL-LIKE ZINC FINGERS BIND TO NUCLEAR IMPORT PROTEINS AND ARE REQUIRED FOR EFFICIENT NUCLEAR LOCALIZATION OF ERYTHROID KRÜPPEL-LIKE FACTOR*

Karen J. Quadrini and James J. Bieker‡

From the Department of Molecular, Cell and Developmental Biology, Mount Sinai School of Medicine, New York, New York 10029

Erythroid Krüppel-like Factor (EKLF/KLF-1) is an erythroid-specific transcription factor that contains three C2H2 zinc fingers and is required for correct chromatin structure and expression of the β-globin locus. However, regions within the EKLF protein that serve as signals for its nuclear localization and the proteins that may enable it to become localized are unknown. Two approaches were used to address these issues. First, green fluorescent protein or pyruvate kinase was fused to EKLF domains, and localization was monitored and quantitated by confocal microscopy. Two necessary and sufficient nuclear localization signals (NLSs) were identified: one (NLS1) adjacent to the zinc finger DNA binding domain within a highly basic stretch of amino acids (275–296), and another more efficient signal (NLS2) within the zinc finger domain itself (amino acids 293–376). Interestingly, each zinc finger contributes to the overall effectiveness of NLS2 and requires an intact finger structure. Second, each NLS was tested in vitro for binding to importin proteins. Surprisingly, both EKLF NLSs, but principally the zinc finger domain, bind importin α and importin β. These findings demonstrate that two nuclear localization signals target EKLF to the nucleus and suggest this transport relies primarily on a novel zinc finger/importin protein interaction.

Proteins that function in the nucleus require access to that compartment of the cell. The aqueous channel of the nuclear pore has an effective diameter of 9 nm, and ions, molecules, and small proteins < 20 kDa can freely diffuse through it. However, proteins larger than 20 kDa are unable to freely diffuse through the nuclear pore and require an active transport process (reviewed in Refs. 1 and 2). The translocation of a protein can be divided into two steps. First, the protein is recognized by a soluble importin receptor, which then docks at the nuclear membrane. This step is energy-independent. Then these molecules are passed through the pore. Initially, it was suspected that energy played a role in this second step (reviewed in Ref. 3), but current models suggest that this step is also energy-independent; translocation per se is not directly coupled to energy expenditure (Refs. 1 and 2 and references therein).

The importin β superfamily is highly conserved across species and is the receptor necessary for nuclear pore recognition (reviewed in Ref. 4). Different subsets of molecules required to be imported into the nucleus are recognized by alternate importin β isoforms. These isoforms recognize distinct classes of signals on their respective cargo. The classical pathway is the most well characterized and involves the first isolated importin receptor, importin β1, which recognizes basic stretches of amino acids in many cell proteins and transcription factors (e.g. Smad3 (5)). Most proteins that utilize the classical pathway are not directly recognized by importin β, but require an adaptor protein, importin α. This protein recognizes basic stretches of amino acids present in a wide range of cellular proteins. The importin α adaptor family consists of various isoforms, which have distinct binding properties and exhibit tissue-specific expression (6–11).

Although there are a variety of signals recognized by different importin β isoforms, nuclear localization signals (NLSs) recognized by the classical pathway receptors are characterized by a stretch of basic amino acids that can fall loosely into one of two categories. The first type, the monopartite NLS, is a continuous stretch of basic amino acids, where four out of six are positively charged (arginine or lysine). The prototype of this class is the SV40 large T-antigen protein with its sequence PKK128KRKV (12, 13). The second category, the bipartite NLS, is best illustrated by nucleoplasmin, whereby a 10-amino acid linker region containing a few prolines separates two sets of two to three positively charged amino acids (14, 15). Either one of these signals is necessary and sufficient to localize a protein to the nucleus. A few proteins have unusual NLSs, which resemble neither canonical signal, because they lack clusters of arginine or lysine amino acids (16–19). It has been suggested that they are half NLSs that become juxtaposed to another positive cluster of amino acids within the three-dimensional structure of the protein (20).

Erythroid Krüppel-like factor (EKLF/KLF1) was isolated in an effort to identify genes that contribute to erythroid cell differentiation (21). Since the discovery of EKLF, a closely related family of Krüppel-like transcription factors has emerged. These proteins share extensive homology within their C-terminal C2H2 zinc finger DNA binding domains, yet contain distinct N-terminal transactivation or transrepression do-

* This work was supported by National Institutes of Health Shared Instrumentation Grant 1 S10 RR09145-01 and National Science Foundation Major Research Instrumentation Grant DBI-9724504 and by United States Public Health Services Grants T32 GM08553 (to K. J. Q.) and RO1 DK46856 (to J. J. B.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Molecular, Cell and Developmental Biology, Mount Sinai School of Medicine, Box 1020, One Gustave L. Levy Place, New York, NY 10029, Tel.: 212-241-4143; Fax: 212-860-9279; E-mail: james.bieker@mssm.edu.

1 The abbreviations used are: NLS, nuclear localization signal; EKLF, erythroid Krüppel-like factor; FBS, fetal bovine serum; GFP, green fluorescent protein; GST, glutathione S-transferase; PBS, phosphate-buffered saline; DAPI, 4′,6-diamidino-2-phenylindole; DIC, differential interference contrast; BSA, bovine serum albumin; PK, pyruvate kinase; STAT1, signal transducers and activators of transcription 1.
 mains and exhibit differential tissue expression patterns (22). EKLF, a red cell-specific protein, is 358 amino acids long and consists of two separable, functional domains (23). One serves as a DNA binding domain that is basic and is composed of three Krüppel-like C_{2}H_{2} zinc finger motifs. The second is a transcription domain that is proline-rich and contains an acidic region near its N terminus (amino acids 76–94) (21). The EKLF protein binds the consensus sequence 5’-CCN CCC CCN-3’ (CACCC element) with each finger making contacts with three bases in the recognition motif (24). Although CACCC elements are present in a multitude of gene promoters, not all CACCC sites are homogenous as illustrated by the ability of EKLF to differentiate between sites and preferentially bind the adult β-globin CACCC site (25). In vivo, EKLF –/– mice die at E14.5, at the time of the switch in expression to adult β-globin, due to a fatal anemia caused by an α- to β-globin chain imbalance (26, 27). These animals have little to no adult β-globin expression, whereas embryonic β-like globin expression levels are unaffected. In addition to these transcriptional effects, the absence of EKLF leads to loss of chromatin repressor activity (28). Notably, β-globin expression levels are unaffected. In addition to these transcriptional effects, the absence of EKLF leads to loss of chromatin structure at the adult β-globin promoter (28). A molecular explanation for this follows from the observations that EKLF interacts with the SWI/SNF chromatin remodeling complex and that this affinity is increased by the in vivo acetylation of EKLF by p300/CBP (29–32). Although most functional evidence depicts EKLF as a transcriptional activator, recently it has been demonstrated that EKLF can also interact with mSin3A and HDAC1 and thus may behave as a transcriptional repressor as well (33).

Although EKLF is a nuclear transcription factor, it is not known which regions of the protein serve as signals for its nuclear localization. Of particular interest is whether the zinc finger domain plays a role in this process because it is not known how individual fingers contribute to localization nor whether a zinc finger transcription factor could interact with importin proteins utilized by the classical pathway for nuclear import. The present studies elucidate the NLSs within the EKLF protein and the import machinery that it may utilize for its nuclear localization.

MATERIALS AND METHODS

Cell Culture—K562 cells were maintained in RPMI medium (Invitrogen) supplemented with 10% FBS. CV-1 cells were maintained in Dulbecco’s modified enriched medium (Invitrogen) supplemented with 10% FBS.

Plasmid Constructs—Synthetic oligonucleotide primers containing BamHI sites and sequences complimentary to the 5’- and 3’-ends of full-length EKLF (amino acids 20–376), zinc finger only (EKLF{376}), and proline domain only (EKLF{297–376}) were used for PCR amplification on the full-length EKLF cDNA template. The inserts were ligated into the BamHI site of the mammalian expression vector, pEGFP-C3 (CLONTECH). Similarly, primers were used to amplify and clone the 5’ basic region (amino acids 275–298), zinc finger 1 (amino acids 293–318), zinc finger 2 (amino acids 319–348), and zinc finger 3 (amino acids 349–376) into pEGFP-C1 (CLONTECH) to maintain the proper reading frame of the coding regions. Generation of site-directed mutants and internal deletions used the QuikChange kit (Stratagene) according to the manufacturer’s protocol after appropriate design of PCR primers.

PCR primers containing a KpnI site in the 5′ primer and a stop codon and NofI site in the 3′ primer were used to amplify the full-length, 5′ basic region, zinc finger 1, zinc finger 2, and zinc finger 3 regions of EKLF as well as the regions containing both ZnF2 and ZnF3 (amino acids 318–376) and all three zinc fingers (ZnF1–3, amino acids 293–376). These primers were also utilized in various combinations to amplify EKLF RR/GL, EKLFΔpro(20–274), EKLFΔZnF1, and EKLFΔZnF3 variants of the EKLF protein. A 3′ primer complimentary to the 3′-end of the basic region, which also contained a stop codon and a NofI site, was generated and used in conjunction with the 5′ primer described above to amplify EKLFΔZnF1(297–376) and EKLFΔZnF3(297–376) RR/GL. PCR products were ligated into the C terminus of the Myc-tagged pyruvate kinase vector (kind gift of Gideon Dreyfuss (34)). Internal deletion of ZnF2 (EKLFΔZnF2) and point mutations within the zinc finger domain of EKLF (EKLF ZnF1,2,3 HA) were generated using the QuikChange kit (Stratagene) and Multimeric kit (Stratagene) according to the manufacturer’s protocols after the appropriate design of PCR primers.

In general, GFP alone and GFP-fused murine and human EKLF regions were amplified with specific PCR primers that contained a NdeI site in the 5′ primer and a stop codon and NofI site in the 3′ primer, and cloned into the pet30a vector (Novagen). The GFP-SMAD3–MH domain construct (kind gift of Harvey Lodish (35)) was similarly amplified and subcloned. A pcDNA3.1 vector containing HA-p53 was obtained as a kind gift from Jonathan Licht.

For the GST-pull-down assays, GST alone (21), GST-importin α (major isoforms NPI-1/hsSRPI and GST-RCH-1/hsSRPIα), and GST-importin β1 (kind gifts of Harvey Lodish) were as described previously (5). All PCR products were gel-purified (Qiagen), and integrity of all constructs was verified by restriction and/or sequence analysis. PCR primer sequences are available upon request.

Transfection—Transfection of K562 cells and CV-1 cells was performed using the DMRIE lipofection method (Invitrogen). K562 cells were resuspended at a density of 1 × 10^6 cells/ml in OPTI-MEM (Invitrogen), and 200 μl of cells were used per transfection. CV-1 cells were seeded at 75% ethanol-treated coverslips 1 day prior to transfection, so that coverslips would be 50–70% confluent on the day of transfection. 4 μg of test DNA and 8 μl of DMRIE reagent were used in each transfection. After 5 h at 37 °C, 2 ml of RPMI (K562) or DMEM (CV-1) supplemented with 15% FBS was added, and incubation continued at 37 °C.

Confocal Analysis of GFP-fused Constructs—At 25–27 h post-transfection, 750 μl of K562 cells was collected and spun onto frosted slides (Fisher) in a Shandon Cytospin 2 centrifuge at 600 rpm for 5 min. All cells were fixed with 2.5% formaldehyde/PBS at 12 min at room temperature. Fixative was added directly to CV-1 coverslips. Cells were washed with PBS, dried, and then mounted with DAPI containing Vectashield (Vector Laboratories). The DAPI was used to identify cell nuclei. DAPI, GFP, and DIC images were visualized on a Leica confocal microscope.

Indirect Immunofluorescence—MyC-PK-fused transfected K562 cells were attached to slides by cytocentrifugation as described above at 48 h post-transfection. All steps were carried out at room temperature except when indicated. Cells were fixed for 30 min with 3% formaldehyde/PBS and then permeabilized with 0.5% Triton X-100 for 10 min. Cells were blocked and incubated with PBS and blocked in 3% BSA/PBS for 1 h. The slides were incubated with 9E10 anti-myc monoclonal antibody (Mount Sinai Hybridoma Center) overnight at 4 °C at 2 μg/ml. Slides were washed 3× with 3% BSA/PBS and then incubated with goat anti-mouse secondary antibody conjugated to Alexa568 (Molecular Probes) for 1 h at 10 μg/ml. Cells were washed 3× with 3% BSA/PBS with a final rinse in PBS. Slides were then mounted with DAPI containing Vectashield. DAPI, Alexa568 (Texas Red), and DIC images were visualized on a Leica confocal microscope.

Calculations—The total cell area and GFP intensity, in addition to the nuclear area and its GFP intensity, was quantified per sample cell on a Leica confocal microscope. The nuclear GFP intensity was subtracted from the total cell GFP intensity for a given cell to obtain that cell’s cytoplasmic GFP intensity. The same cell’s nuclear area was subtracted from its total cell area to get its cytoplasmic area. Division of the cytoplasmic GFP intensity by its area yielded the cell’s cytoplasmic signal. Similarly, the nuclear GFP intensity was divided by its area to get the cell’s nuclear signal. Division of these two numbers gives the nuclear to cytoplasmic ratio of the cell. The sample size of cells chosen for each GFP-fused construct was as indicated in the figure legends. All the nuclear to cytoplasmic ratios calculated per sample construct were averaged. Although these calculations do not take into account possible variations in protein expression levels of individual constructs, they do allow assessment of the relative efficiency of a signal as compared with background (GFP alone).

Calculations for Myc-PK-fused construct-expressing cells were analyzed and quantified as described above and in conjunction with the Macintosh IPLab analysis program. Sample sizes are as indicated in the figure legends.

In Vitro Transcription and Translation—All T7 promoter-containing constructs were in vitro transcribed and translated with the T7TP3-T7 kit as recommended by the supplier (Novagen). 2 μl of each translated sample was analyzed by 12% SDS-PAGE gel electrophoresis and autoradiography. Translated proteins were stored at −20 °C.
**GST Fusion Protein Preparations**—BL21 cells (Novagen) were transformed with all GST-containing constructs and used to generate glutathione beads (Amersham Biosciences) bound to the proteins of interest by established procedures (29). Bound proteins were eluted from beads by boiling in SDS loading buffer, electrophoresed on a 12% SDS-PAGE gel (Bio-Rad), and stained with Gelcode Blue reagent (Pierce) to estimate the concentration and monitor the integrity of each immobilized protein preparation.

**Fig. 1.** Nuclear accumulation of GFP-fused EKLF and EKLF functional domains in transfected K562 cells. A, schematic of the various constructs used to transfect K562 cells. The three zinc fingers are sequentially numbered in the hatched boxes. The 5′ basic region is denoted by the gray-shaded box. Amino acid boundaries are as indicated. All constructs were expressed from the pEGFP-C3 mammalian expression vector by means of the CMV promoter. B, representative cells from a transient transfection of K562 cells with the constructs shown in A. The arrow in panel c indicates an example of a cell with cytoplasmic staining. Cells were processed for microscopy 25–27 h post-transfection. From left to right, photographs are DAPI, GFP, and DIC images. Magnification: ×80. C, K562 cells transiently transfected with the constructs in A were randomly selected, quantitation of their individual signal intensities was performed, and their individual nuclear to cytoplasmic ratios were calculated as described under “Materials and Methods.” The averages ± S.E. (n = 16–21) are shown from one representative experiment.
GST-Pull-down Assay—4 μl of 35S-translation products, equal protein concentrations of GST-immobilized beads, and binding buffer (25 mM HEPES, pH 7.6, 100 mM NaCl, 10% glycerol, 2 mM β-mercaptoethanol, 0.1% Nonidet P-40, 3.4 μg/ml phenylmethylsulfonyl fluoride, 10 μg/ml leupeptin, and 10 μg/ml pepstatin) was added to a final volume of 100 μl and incubated for 2 h at 4°C. 400 μl of binding buffer was added after 2 h, and beads were centrifuged at 4°C for 2 min at 1000 rpm. Beads were washed three times with 500 μl of binding buffer, boiled in SDS loading buffer, and electrophoresed on a 12% SDS-PAGE gel (Bio-Rad) along with 0.5 μl of in vitro transcribed and translated proteins (representing 12.5% of input). All gels were treated with ENHANCE solution (PerkinElmer Life Sciences). Gels were exposed to film (Kodak X-AR) for 3–4 days at –80°C.

RESULTS

Two Regions of EKLF Enable Nuclear Accumulation of a Diffuse Protein to the Cell Nucleus—To assess the subcellular localization of the EKLF protein, full-length EKLF was fused in-frame to GFP, a freely diffusible protein (35), and monitored for localization by confocal microscopy after transfection into K562 or CV1 cells. K562 is an erythroleukemic cell line, whereas CV-1 is an epithelial cell line. While GFP alone was relatively diffuse, GFP-EKLF was accumulated in the nucleus in both K562 cells and CV-1 cells (Fig. 1A, panels a and b, and data not shown).

To identify regions sufficient for nuclear accumulation within the EKLF protein, its two main functional domains were separately fused to GFP (Fig. 1A): the N-terminal, proline-rich transactivation domain (EKLFΔZn, amino acids 20–291) and the C-terminal, three C-H2 zinc finger-containing DNA binding domain (EKLFΔpro, amino acids 287–376). Both protein domains caused GFP to accumulate in the nucleus (Fig. 1B). However, many cells transfected with EKLFΔZn showed cytoplasmic staining (Fig. 1B, panel c), whereas those transfected with EKLFΔpro did not (Fig. 1B, panel d). Signal quantitation followed by determination of the nuclear to cytoplasmic ratios of K562 cells transfected with these truncations verified that the zinc finger region can cause more effective nuclear accumulation of GFP (Fig. 1C).

Because both domains contain sequences sufficient for nuclear accumulation of GFP, these were further subdivided. Studies on the highly homologous GKLF protein have indicated that it contains an NLS within the portion of the protein just N-terminal to the zinc finger domain, within its transactivation domain (36). Close analysis of the transactivation domain of EKLF indicated that this region could be a candidate for a classical NLS (because it contains a stretch of positive amino acids; Fig. 2B, panel b). This 5′ basic region alone directed nuclear accumulation of GFP (Fig. 2) to a level almost equal to that directed by the three zinc fingers together (compare quantitated data in Fig. 1C with Fig. 2D). Disruption of two arginines at amino acids 282 and 283 in this peptide greatly reduced its effect, suggesting that this 5′ basic region plays a role in nuclear accumulation (Fig. 2C, panels b and f, and Fig. 2D).

Each zinc finger was separately fused to GFP (Fig. 2A) to determine whether any one finger is sufficient to accumulate GFP in the nucleus. Of the three individual fingers, only Zinc finger 1 causes accumulation of GFP in the nucleus (Fig. 2C, panels c–e). Zinc finger 2 and Zinc finger 3 are very inefficient and cause only minimal nuclear accumulation when compared with control scrambled peptide (Fig. 2D). However, Zinc finger 1 did not fully reconstitute the nuclear accumulation effect seen by all three fingers together (Fig. 1C).

The previous data demonstrate that, in isolation, the EKLF 5′ basic region and Zinc finger 1 exhibit the most efficient nuclear accumulation effects upon GFP. To further examine their role, alterations were made in these regions within the context of full-length EKLF to determine if their removal was sufficient to disrupt EKLF accumulation. The RR to GL mutation, although decreasing the efficiency of the signal of the 5′
to GL mutation. Such a disruption of the C$_2$H$_2$ structure eliminates one of the amino acids necessary to coordinate the Zn$^{2+}$ ion and form the correct structure to bind DNA. However, this disruption does not influence the structure or DNA binding activity of nearby zinc finger domains (38). The results (Fig. 3B, panel d) show that this construct accumulates in the nucleus as well as the unaltered protein. The cumulative GFP data suggests that the individually weak effects of Zinc fingers 2 and 3 (seen in Fig. 2) might still be playing a significant role in the context of full-length EKLF, where they are together and adjacent.

The acetylation status of EKLF increases its transactivating ability (29) and influences its protein-protein interactions (32). Because one of the acetylated lysines (Lys-288) is located in the 5′ basic region, a site-directed mutation (to arginine) was made in GFP-fused full-length EKLF. This had no effect on EKLF localization because it remained accumulated in the nucleus. Abolishment of the positive charge altogether at Lys-288 through a mutation to alanine also had no effect on accumulation (data not shown). Therefore, the nuclear accumulation of EKLF is independent of charge or acetylation status of this lysine.

To test for any cell-specific effects of our observations, the subcellular localization of all GFP fusion constructs was simultaneously monitored in transfected CV-1 cells, a kidney epithelial cell line. Expression of the constructs in CV-1 cells shows localization patterns identical to those obtained as when expressed in K562 cells, indicating that there is no erythroid-specific effect that directs nuclear accumulation of EKLF in K562 cells (data not shown).

Two EKLF Domains Are Sufficient for Nuclear Localization—While GFP fusion is a common approach to visualize the nuclear or subcellular location of a protein, GFP can accumulate significantly to the nucleus on its own. Our quantitated data had yielded background (GFP-alone) nuclear/cytoplasmic ratios of ~2–3. To address nuclear localization directly, we fused regions of the EKLF protein to pyruvate kinase (PK), an exclusively cytoplasmic protein that has been used to test the active nuclear import of proteins (34, 39). This provides a very stringent test because significant redistribution of the PK protein to the nucleus after fusion to a test protein can only occur in the presence of an efficient bona fide nuclear localization signal. To easily identify transfected cells that express the PK-fused EKLF constructs, a myc-tag/PK fusion served as the epitope for monoclonal antibodies for indirect immunofluorescence.

All cells transfected with the empty pcDNA3 Myc-PK vector showed total nuclear exclusion of the PK protein (Fig. 4A, panel a) with a quantitated nuclear/cytoplasmic ratio of ~0.3–0.4, verifying our initial premise in using this reporter. In contrast, all full-length EKLF-expressing cells showed complete nuclear localization of the PK-EKLF protein (Fig. 4A, panel b). The 5′ basic region was sufficient to direct PK to the nucleus but with modest efficiency (Fig. 4A, panel c) as determined by signal quantification, determination of the nuclear to cytoplasmic ratio, and comparison to full-length (Fig. 4B). Thus, the 5′ basic region behaves as a nuclear localization signal, and we denote this as NLS1.

The GFP fusion data had suggested that the zinc fingers play a role in nuclear accumulation of EKLF. To test whether each zinc finger alone or in combination was sufficient for active nuclear transport, PK-zinc finger fusion proteins were generated. Each zinc finger, when expressed individually, had equally modest effects on relocating PK to the nucleus (Fig. 4A, panel d) at an even less efficient level than the 5′ basic region (compare Fig. 4A, panels c and d, and see Fig. 4B). A PK fusion
to two EKLF fingers together (ZnF2 and ZnF3) behaved similarly to that seen when each finger was tested individually (Fig. 4A, panel d).

Next, all three zinc fingers were fused to Myc-PK, and cells were analyzed for PK localization. PK was intensely localized to nuclei more efficiently than each zinc finger alone or than the 5' basic region but was insufficient to ensure wild type levels of localization (Fig. 4A, panel e, and B). Therefore, we conclude that the three EKLF zinc fingers collectively serve as a *bona fide* NLS that is even more efficient than its 5' basic region.
region. We denoted this region NLS2.

**Two EKLF Domains Are Necessary for Optimal EKLF Nuclear Localization**—Because the entire zinc finger domain (NLS2) and the 5' basic region (NLS1) can each direct accumulation (of GFP) and relocation (of PK) to the nucleus in isolation, we asked whether each region is necessary for localization when tested in the context of the full-length EKLF protein. Deletion of the zinc fingers significantly reduced the nuclear localization of EKLF (Fig. 5B, panel d, and C) but did not totally abolish it. Disruption of the 5'/H11032 basic region by the RR to GL mutation in the context of full-length EKLF also decreased localization but to a lesser extent (compare Fig. 5B, panels c and d, and see Fig. 5C). Disruption of the 5' basic region in combination with deletion of all three zinc fingers rendered PK in the cytoplasm at levels equal to empty PK vector alone (Fig. 5B, compare panel e to panel a, and see Fig. 5C). Conversely, deletion of the entire transactivation domain of EKLF that is N-terminal to both NLS1 and NLS2 resulted in wild type levels of nuclear PK protein. These data show that the N-terminal region is not necessary for localization and that the 5' basic region and the zinc finger domain together are sufficient for maximal EKLF nuclear localization (Fig. 5B, panel f, and C). Therefore we conclude that NLS1 and NLS2 are necessary and together sufficient for optimal nuclear localization of the EKLF protein.

**Each Zinc Finger and Its Structure Contributes to the Overall Effectiveness of NLS2**—Our studies demonstrated that the zinc finger domain plays a necessary role in EKLF nuclear localization. Consistent with this, our studies with the PK-zinc finger construct had indicated that all three fingers together were required for an efficient localization signal, similar to that seen for nuclear accumulation with the GFP-zinc finger fusion. However, each finger individually had yielded only modest effects on PK localization. To test the importance of each zinc finger in the context of full-length EKLF, individual fingers were deleted and monitored for their effect on PK-EKLF localization. Deletion of any single finger decreased, but did not abolish, the amount of EKLF localized to the nucleus, indicating that each finger individually contributes to the total zinc finger domain NLS2 function (Fig. 6).

We next determined if the structure of the finger domain also plays a role in localization. We disrupted one of the histidines that coordinate zinc for structure (as described above) in each finger. This decreased, but did not abolish, the amount of EKLF localized to the nucleus (Fig. 6). From the data of Fig. 6 we conclude that all three zinc fingers along with their proper
structure are necessary to ensure maximal EKLF targeting to the nucleus.

Each EKLF NLS Can Bind Either Importin $\alpha$ or Importin $\beta$—We next tested each NLS signal for their ability to interact with known importin proteins that recognize canonical NLSs in the classical import pathway. The importin $\beta_1$ docking receptor is ubiquitously expressed in mouse tissues (40). However, the importin $\alpha$ adaptor isoforms that associate with this receptor display a unique expression pattern in mouse tissue and culture cells (11, 40, 41). Therefore, we chose to test importin $\alpha_2$/Rch-1/hSRP1alpha/karyopherin $\alpha_2$, because it is a more abundant adaptor in tissues where EKLF expression is found (11, 41, 42). For simplicity, we will refer to importin $\alpha_2$/Rch-1/hSRP1alpha/karyopherin $\alpha_2$ as Rch-1. GST-importin $\beta_1$ and GST-Rch-1 were bound to glutathione-Sepharose beads. Selected fusions of GFP/EKLF were transcribed and translated in vitro, and pull-down assays were performed on the two populations of proteins. Although the Rch-1 and importin $\beta_1$ utilized in these studies are the human orthologues, they share 94 and 96% identity with mouse importin $\alpha_2$/mPendulin/mouse $\beta_2$ and mouse Imp $\beta_1$, respectively (10; NCBI accession numbers D67015 and NM_02265).

Approximately equal amounts of GST fusion proteins were utilized in the pull-down assays (Fig. 7A). p53 and SMAD3-MHI domain bound Rch-1 and importin $\beta_1$, respectively, with high affinity (Fig. 7, B and C, first panel) (5, 43, 44). Very little background was detected between the GST-bound beads and the in vitro transcribed and translated proteins (Fig. 7, B and C, second panel). EKLF$\Delta\alpha_1$ bound Rch-1 and importin $\beta_1$ with a greater intensity than either the 5' basic region or full-length EKLF (Fig. 7B, first panel). We conclude that the EKLF zinc finger region (NLS2) can efficiently bind both importin $\alpha_2$ and $\beta_1$ proteins with similar intensity, and we also note that this binding is always greater than that of full-length EKLF.

Binding by the 5' basic region was more complex. Murine EKLF 5' basic region bound Rch-1/importin $\alpha_2$/hSRP1alpha/karyopherin $\alpha_2$ but not importin $\beta_1$ (Fig. 7, B and C). In contrast, the human EKLF 5' basic region bound with a significantly greater affinity to both importin proteins (Fig. 7, B and C). As a result, we conclude that the 5' basic region (NLS1) can clearly bind importin $\alpha_1$, but that binding to importin $\beta_1$ may be influenced by species-specific effects (see “Discussion”).

A less abundant isoform of importin $\alpha_1$, importin $\alpha_1$/NPI-1/hSRP1/karyopherin $\alpha_1$, which is found in mouse tissues and culture cells where EKLF is present (11, 41, 42), was also used in pull-down assays in vitro with the above EKLF variants. Results for these were identical to those obtained for importin $\beta_1$ (data not shown).

DISCUSSION

EKLF is an erythroid cell-specific transcription factor that plays critical roles in the activation of $\beta$-globin transcription (25–27) and interacts with components of remodeling complexes to alter chromatin structure at the $\beta$-globin locus (30, 32). These roles ensure the completion of the $\gamma$- to $\beta$-globin switch (28, 45). Because EKLF performs these functions in the nucleus, we investigated what regions of the protein serve as NLSs to dictate its subcellular distribution as well as what potential proteins it may utilize for its nuclear localization.

Our studies indicate that there are two functional NLSs within the EKLF protein. The first, NLS1, is a “classic” NLS that contains a short series of positive amino acids that lie within amino acids 275–296, with the sequence $^{276}$PKR$^{278}$RKT-LAP$^{288}$. Whereas (using GFP-fused constructs) just the first one and a half zinc fingers are necessary for the nuclear localization of the closely related protein GKF/KLF4 (36), all three fingers serve as the second signal, NLS2, for the EKLF protein.

NLSs located nearby or within other types of zinc finger DNA binding domains have been identified in an array of transcription factors such as yeast Mata2 (18, 19), glucocorticoid receptor (46), yeast SW15 (47), and progesterone receptor (48, 49). However, the use of a complete zinc finger domain as a nuclear localization signal has only been observed in NGFI-A (37, 50). NGFI-A, an immediate early gene product, utilizes its $C_{\text{2n}}H_2$ zinc finger domain as its NLS, but its efficiency is abrogated once the overall structure of the domain is altered (37). Disruption of the overall structure of the EKLF zinc fingers similarly decreases the efficiency of EKLF nuclear localization. However, this level of nuclear localization remained higher than that obtained when the entire zinc finger domain was deleted. It is likely that the overall positive charge of the zinc finger domain, in addition to its secondary structure, are two important determinants that dictate the efficiency of NLS2.

Several subgroups exist within the larger family of Krüppel-like factors that share more extensive homology among themselves than with other members of the family. Because of this, the proteins within such a subgroup may share many conserved protein–protein interaction domains and protein modifications. EKLF/KLF1 shares a subgroup with both GKF/KLF4 and LKLF/KLF2 (22). It has been postulated that the 5' basic region within LKLF, GKF, and EKLF defines this subgroup of Krüppel-like factors and serves as the predominant NLS for each (22, 36). Because the entire Krüppel-like family contains additional conserved residues between and within each finger (22) and our studies demonstrate that the entire zinc finger domain serves as an NLS, we believe the highly conserved zinc finger domain in EKLF may be a common NLS that is used by all members of the KLF family. As the highly conserved 5' basic region is only shared by KLF1, -2, and -4 (22, 32), it will be interesting to determine whether the remaining family members contain a functionally analogous region, which serves as a second NLS, or whether they merely utilize the zinc finger domain for nuclear localization.

Other transcription factors have been described as containing more than one NLS (18, 19, 46, 51). Some of these play an essential role in cell growth and differentiation (52–54). Moreover, it has been observed that several signals within a protein may exert cooperative effects and increase the efficiency of uptake into the nucleus (55). This could be due to altering the
kinetics of nuclear uptake (55, 56) and influencing the rate-limiting interaction between cargo and the import machinery (reviewed in Ref. 4). Additionally, one signal may serve a more regulatory role. The presence of both NLS1 and NLS2 promotes maximal targeting of EKLF to the cell nucleus and ensures that it is properly localized in cells where it is to execute its many functions. The biological significance of properly localized EKLF is underscored by the severe phenotype resulting from the deletion of the EKLF protein in mice (26, 27).

The variety of import signals makes it difficult to predict from sequence alone whether a protein is actively imported, and if it is, which importin receptor is utilized. In these studies we find that both the EKLF 5 basic region and the zinc finger domains can bind importin proteins utilized in the classical pathway of nuclear import in vitro. It is not surprising that the 5 basic region (NLS1) can bind importin α2/Rch-1, because it most resembles a canonical monopartite NLS of the SV40 NLS type (PKKKRKV), which recognizes both importin α1 and importin α2 (6). On the other hand, the zinc finger domain does not exhibit a distinct distribution or pattern of positive amino acids similar to those present in a canonical NLS that would be recognized by a classical receptor for import, although it possesses a large number of positively charged amino acids. It is significant that the zinc finger domain can interact with importin proteins. The human 5 basic region bound both importin proteins with a higher affinity than the murine 5 basic region. This may not be so surprising because human and murine EKLF 5 basic region only share 73% identity (m5BR sequence 273TAPPKRMRRTLPKRQAAHTCG286 versus h5BR sequence 273TAPSKRGRRTSPKRQAAHTCA286). The percent homology of this region of EKLF is much less than the 93% homology between the human and mouse zinc finger domains (21, 57). In addition to the extra arginine present in the human 5 basic region, the sequences within and surrounding this region may have a positive effect on its affinity for binding the human importins. Another possibility is that NLS1 may indirectly regulate EKLF localization by means other than direct interaction with importin proteins. Functional testing by nuclear import assays may help resolve this issue.
Even more intriguing is the ability of both signals (5′ basic region and zinc finger domain) to bind either the common receptor itself, importin β, or its adapters, importin α1/NPI-1/ hSRP1/karyopherin α1 and importin α2/Rch-1/hSRP1α/karyopherin α2. These two importins are present in tissues and culture cells where EKLF is functional (11). It is believed that the number of distinct importin α isoforms in higher mammals implies a specialization in their cellular role and that different isoforms bind particular target proteins (4). For example, NPI-1 but not Rch-1 recognizes STAT1 transcription factor (7).

Although one NLS can bind multiple importin α adapters, the ability of a single NLS to bind multiple importin receptors occurs less often, for example in the case of ribosomal proteins (58, 59). The ability of a protein to be recognized by its import receptor can recognize the EKLF protein in vitro. If both the importin α/β heterodimer and importin β receptor alone are functional receptors for EKLF in vitro, then the presence of either would ensure that EKLF is properly localized. Whether separate binding sites on importin α and importin β are responsible for the recognition of either EKLF NLS signal and whether the importin α/β heterodimer or just the importin β receptor itself actually functions in the cell to mediate nuclear transport of EKLF remain to be determined.

Acknowledgments—We thank Zhan Xiao, Harvey Lodish, Gideon Dreyfuss, Glenn Fishman, Heidi Stublhammer, and Jonathan Licht for plasmids. We also thank Debbie French, Gillian Small, and the Bieker laboratory for discussion; Serafin Pinol-Roma for extensive discussion and comments on the manuscript; Scott Henderson for assistance with the Leica confocal microscope at the Mount Sinai School of Medicine Microscopy Center; and Victor Friedrich for antibodies and help with figure preparation.

Note Added in Proof—While this paper was in review, similar results for EKLF localization were published by Pandya and Townes (Pandya, K., and Townes, T. M. (2002) J. Biol. Chem. 277, 16504–16512).