

# The GATA-E box-GATA motif in the EKLF promoter is required for in vivo expression

Kathleen P. Anderson, Scott C. Crable, and Jerry B. Lingrel

**The erythroid Krüppel-like factor (EKLF) is a key regulatory protein in globin gene expression. This zinc finger transcription factor is required for expression of the adult  $\beta$  globin gene, and it has been suggested that it plays an important role in the developmental switch from fetal  $\gamma$  to adult  $\beta$  globin gene expression. We**

**have previously described a sequence element in the distal promoter region of the mouse EKLF gene that is critical for the expression of this transcription factor. The element consists of an E box motif flanked by 2 GATA-1 binding sites. Here we demonstrate that mutation of the E box or the GATA-1 consensus se-**

**quences eliminates expression from the EKLF promoter in transgenic mice. These results confirm the importance of this activator element for in vivo expression of the EKLF gene. (Blood. 2000;95:1652-1655)**

© 2000 by The American Society of Hematology

## Introduction

The erythroid Krüppel-like factor (EKLF) has 2 zinc finger motifs in the C-terminal half and a proline-rich transactivation domain in the N-terminus of the protein.<sup>1</sup> It is the founding member of a family of transcription factors related to each other through a high degree of homology in the zinc finger region.<sup>2</sup> This class of proteins has been termed the KLF family for the conserved C2H2 finger regions, and it consists of 12 members expressed with varying degrees of tissue specificity.<sup>1-13</sup>

EKLF is first expressed in yolk sac blood islands and remains erythroid specific throughout development, with adult expression observed in the spleen.<sup>14</sup> Targeted disruption of the EKLF gene in mice has demonstrated that this factor is critical for hematopoiesis.<sup>15,16</sup> Indeed, EKLF is required for expression of the adult  $\beta$ -hemoglobin gene, and it plays an important role in the developmental switch from fetal to adult  $\beta$ -globin expression.<sup>17-19</sup>

We have been investigating the regulated expression of this important erythroid-specific protein and have recently identified an interesting activator motif in the distal promoter of the EKLF gene.<sup>20</sup> The site consists of 3 consensus-binding sequences, a GATA site, an E box, and another GATA site. The GATA-E box-GATA, or GEG, motif appears to organize a protein complex rather than supporting independent binding by individual factors. This conclusion arises from the observation that a mutation at any 1 of the 3 sites in the motif eliminates activation of the EKLF promoter. Consistent with this functional data, there have been reports describing a physical association in erythroid cells between GATA1, Tal1/E2A, Ldb1, and Lmo2, where Lmo2 acts as a bridging molecule between GATA1 and Tal1.<sup>21-25</sup> Tal1/E2A are a helix-loop-helix heterodimer pair recognizing an E-box sequence motif.<sup>24,26,27</sup> Lmo2 lacks the ability to bind DNA but does contain LIM domains for protein-protein interactions.<sup>28,29</sup> The Ldb1, LIM domain binding, protein is a recently characterized binding partner for Lmo2.<sup>23,30</sup> Significantly, targeted ablations of the gene for either GATA1, Tal1, or Lmo2 all result in a failure in development of the

hematopoietic system.<sup>31-35</sup> With both the Tal1 and the Lmo2 genes, this effect is manifested with embryonic lethality at approximately day 8.5 of gestation (E8.5); the GATA1 null embryos survive until E11. This further implicates the interaction of these transcription factors in the regulation of genes involved in the hematopoietic pathway.

Our previous study identifying the GEG motif in the distal EKLF promoter was based on stable and transient transfection experiments in MEL cells.<sup>20</sup> The current work extends these studies and demonstrates the importance of this combination of consensus sites for in vivo erythroid expression from the EKLF promoter.

## Material and methods

### Plasmids

The wild-type -1150 EKLFCAT and the -1150 mutant constructs have been previously described.<sup>20</sup> Briefly, these plasmids contain the mouse EKLF promoter region fused to a CAT reporter cassette in a modified pBKCMV vector (Stratagene, LaJolla, CA). The specific mutations in the GATA and E box binding sites for -1150(G<sup>MEG</sup>M)EKLFCAT, -1150(GE<sup>M</sup>G)EKLFCAT, and -1150(G<sup>M</sup>E<sup>M</sup>G<sup>M</sup>)EKLFCAT are listed in Figure 1.

### Transgenic mice

EKLFCAT expression plasmids were purified (Qiagen, Santa Clarita, CA) before injection, and transgenic animals were produced in the FVB/N mouse strain.<sup>36</sup> Putative transgenic animals were screened by a polymerase chain reaction assay. DNA from established lines was analyzed by Southern blotting to determine copy number and to ensure the transgene was intact and free of rearrangements.

### CAT assays

Blood was collected from F0 and F1 animals and placed in a phosphate-buffered saline/heparin solution. The samples were centrifuged, and the cell

From the Department of Molecular Genetics, Biochemistry, and Microbiology, University of Cincinnati, Cincinnati, OH.

Submitted June 16, 1999; accepted November 4, 1999.

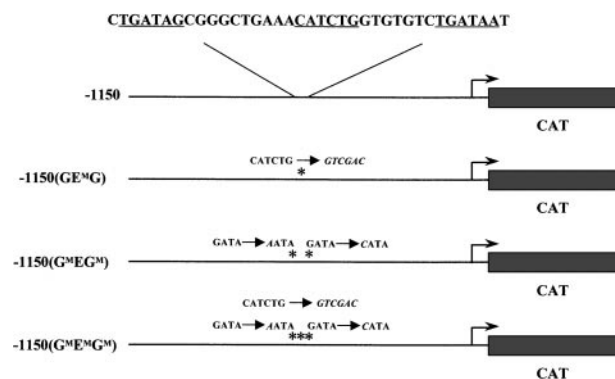
Supported by National Institutes of Health Sickle Cell Center grant HL58421.

**Reprints:** Kathleen P. Anderson, Department of Molecular Genetics, Biochem-

istry, and Microbiology, University of Cincinnati, 231 Bethesda Avenue, ML 524, Cincinnati, OH 45267; e-mail: kathleen.anderson@uc.edu.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 U.S.C. section 1734.

© 2000 by The American Society of Hematology



**Figure 1. Wild-type and mutant EKLf CAT transgenic constructs.** The mouse EKLf promoter from  $-1150$  to  $+60$  was fused to the CAT reporter gene. The nucleotide sequence for GATA-E box-GATA motif is shown above the wild-type construct. Each consensus-binding site is underlined. Asterisks mark the mutations in subsequent constructs; the nucleotide changes are shown in italics.

pellet was lysed by freeze-thaw in a 0.1 mol/L Tris, pH 7.5 solution. Tissue extracts were also prepared from adult transgenic animals by homogenization in the Tris solution, followed by freeze-thaw lysis. CAT assays were conducted as previously described.<sup>2</sup> Results were normalized for protein levels using a bicinchoninic acid assay (Pierce Chemicals, Rockford, IL).

## Results

The GATA-E box-GATA (GEG) element in the distal EKLf promoter was originally identified based on studies in mouse erythroleukemia (MEL) cells.<sup>20</sup> This cell line has been used extensively to analyze critical cis and trans regulatory elements required for the expression of erythroid genes. Because of the complexity of the hematopoietic system, however, complementary studies in transgenic mice can provide a more rigorous test of the importance of such elements or factors. With this view, we produced a series of transgenic animals to ascertain the *in vivo* relevance of the GEG motif in the EKLf promoter.

Constructs containing either a wild-type or a mutated form of the mouse EKLf promoter driving expression of a CAT reporter gene were introduced into fertilized oocytes. Three mutant constructs were used in which either the E box site was replaced ( $-1150(\text{GE}^{\text{M}}\text{G})$ ) or point mutations were introduced to disrupt each of the flanking GATA sites ( $-1150(\text{G}^{\text{M}}\text{EG}^{\text{M}})$ ), or mutations at all 3 sites in the GEG motif were incorporated into a single construct ( $-1150(\text{G}^{\text{M}}\text{E}^{\text{M}}\text{G}^{\text{M}})$ ). These transgenes and the specific GEG sequence alterations are shown in Figure 1.

Four lines of mice carrying the wild-type  $-1150$  EKLfCAT construct were obtained and analyzed. We have shown that this promoter can direct correct tissue and developmentally specific expression of a  $\beta$ -galactosidase gene in transgenic mice.<sup>20</sup> This finding was confirmed in our current study using the CAT reporter gene. The  $-1150(\text{G}^{\text{M}}\text{EG}^{\text{M}})$  construct was used to generate 12 lines of transgenic mice. The F0 animals carrying this mutant transgene were analyzed for CAT expression in the blood. Additionally, 11 of these lines were bred and analyzed as F1. Similarly, we produced 14 F0 mice using the E box mutant construct,  $-1150(\text{GE}^{\text{M}}\text{G})$ . All F0 lines were screened for transgene expression in adult red cells, and 10 lines were bred for confirmation of these results in F1 animals. The promoter construct with alterations at all three binding sites was analyzed in two transgenic lines.

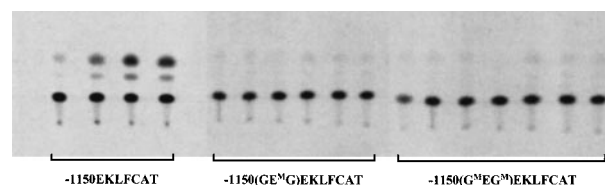
The results of transgene expression in adult F1 blood samples

are shown in Figure 2 and are quantitated in Table 1. All 4 wild-type EKLf promoter lines expressed the reporter gene to varying degrees, in agreement with our previous results.<sup>20</sup> No expression was detected in the 6 representative lines carrying point mutations in the 2 GATA sites of the GEG motif. Furthermore, all 7 of the F1 lines with the E box disruption failed to express the transgene. Finally, as expected from these results, no CAT activity was observed in the blood from the 2 lines in which all 3 sites were altered in the GATA-E box-GATA motif.

As indicated in Table 1, the initial characterization of the F0 mice revealed 1 of 12  $\text{G}^{\text{M}}\text{EG}^{\text{M}}$  animals and 2 of 14  $\text{GE}^{\text{M}}\text{G}$  mice with expression of the transgene in adult blood at a level significantly exceeding the wild-type values. We suspected this expression resulted from a position effect wherein the transgene integrated near a strong enhancer element. Subsequent analysis from 1 of these lines confirmed this, with high levels of CAT activity present in extracts from lung, heart, and spleen, and lower levels in liver and brain. Activity arising from blood contamination of the organs was ruled out by assaying similar tissue samples from a mouse carrying the wild-type EKLfCAT construct. All tissues were negative with this wild-type promoter, with the exception of a low level in the spleen as shown in Figure 3. These results are in agreement with our previous data demonstrating the tissue specificity of this wild-type EKLf promoter, and they support the premise that the infrequent instances of CAT activity arising from the mutant promoter are correlated with unregulated ectopic expression. In addition to reconfirming the tissue-specific expression of the wild-type EKLf promoter, we also assayed tissue extracts from developmental time points. No expression was observed from any samples, either yolk sac or fetal liver, obtained from mutant lines (data not shown). In contrast, the  $-1150$  EKLfCAT construct was appropriately expressed in fetal liver as expected. Therefore, it is not that the mutant transgenes are active early and subsequently shut down in the adult hematopoietic tissue, but, rather, they are never expressed if the GEG site is not intact.

## Discussion

Our previous studies identified a critical sequence element in the distal EKLf promoter comprised of a consensus E box flanked by 2 GATA sites.<sup>20</sup> Studies using various promoter constructs in erythroid and nonerythroid cells indicated that a mutation in any one of these three protein-binding sites would eliminate enhanced expression from the EKLf promoter. The current study extends those findings by demonstrating the *in vivo* importance of the GEG element. Generally only those transgenic animals carrying the wild-type EKLf promoter expressed the reporter gene at all.



**Figure 2. Mutations in either the E box or the GATA sites eliminate expression in adult blood.** Samples from the blood of F1 mice carrying wild-type or mutant EKLfCAT constructs were assayed for CAT activity. All 4 wild-type lines are included. Six representative lines for the E box mutant construct and 7 lines with the GATA site mutations are shown. Additional data are summarized in Table 1.

**Table 1. Expression data from transgenic mice carrying the wild-type and mutant forms of the EKLF promoter**

Construct/Line	Normalized CAT Values		Copy Number
	F0	F1	
1150 EKLF CAT	24	0.5	5
	26	10	7
	29	32	2
	33	20	1
1150 (G <sup>EM</sup> G)	7.1	<0.1	3
	7.2	<0.1	1
	7.4	98	6
	8.2	<0.1	2
	9.1	<0.1	1
	9.4	<0.1	5
	19.1	<0.1	3
	19.7	<0.1	5
	19.8	<0.1	1
	20.3	<0.1	7
	20.5	<0.1	3
	46	46	5
1150 (G <sup>MEG</sup> M)	58	<0.1	3
	69	<0.1	3
	11.1	<0.1	1
	11.4	<0.1	3
	11.5	<0.1	6
	12.1	<0.1	4
	12.5	<0.1	4
	15.4	<0.1	3
	15.5	<0.1	22
	16.1	<0.1	6
	17.5	<0.1	20
	17.7	96	18
1150 (G <sup>MEG</sup> M)	17.8	<0.1	20
	21.2	<0.1	2
	41	<0.1	14
	44	<0.1	1

CAT values are expressed as the percentage conversion of substrate to product, normalized for the amount of protein in each sample.

Any expression observed from a mutant promoter construct was high level and nonerythroid specific, suggesting a transgenic position effect.

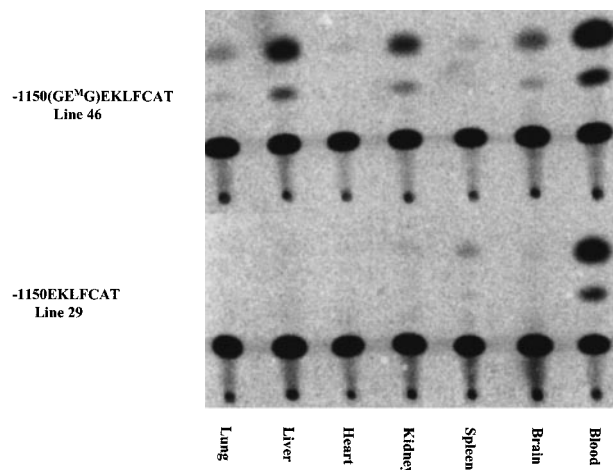
It is particularly interesting that the transgene with a mutation in the E box consensus site is silent. The identification of a multiprotein complex in which Lmo2 acts as a bridge between Tal1 and GATA1 has been identified in erythroid cells,<sup>21-25</sup> which might lead one to anticipate that Tal1 is the protein that binds this functional E box in the EKLF promoter. Our attempts to confirm this speculation have as yet been unsuccessful. Supershift assays using the GEG sequence and erythroid cell nuclear extracts with antibodies directed against Tal1 have yielded neither a complex with an altered mobility nor an inhibition of complex formation. In addition, we performed transfection experiments in an effort to up-regulate CAT expression from the EKLF promoter by overexpressing Tal1 in either erythroid or nonerythroid cells. No evidence for a direct effect by Tal1 was observed. Conversely, Vyas et al<sup>25</sup> have recently described an E box-GATA site upstream of the erythroid-specific promoter for GATA-1. In DNA-binding studies, this site binds a complex containing Tal1 in supershift assays. Our studies certainly do not rule out Tal1 as the E box-binding component of the complex associating with the GEG site in the EKLF promoter. The use of a different anti-Tal1 antibody, for

example, could explain the lack of an effect in supershift assays. Alternatively, however, the presence of 2 GATA sites may change the binding site configuration, and one might also then consider the possibility that another helix-loop-helix protein may be involved.

An erythroid-specific DNase-hypersensitive site has recently been mapped to a region overlapping the GEG motif.<sup>37</sup> These investigators localized the enhancer activity of this segment of the EKLF promoter to sequences from -715 to -666, whereas the element we have described spans the region -687 to -650. The discrepancy in these results may stem from a nucleotide change that eliminates the 3' GATA site in the constructs of the former group. Based on our previous analysis, this GATA site is critical for enhanced activation of the EKLF promoter and appears to bind the GATA1 protein with a higher affinity than the GATA sequence that resides 5' of the E box. We have verified our analysis of the EKLF promoter by determining the nucleotide sequence from the genomic DNA of 3 mouse strains, FVB/N, *Musculus spretus*, and C57BL/6, and from ES cells. Although 2 other single-base polymorphisms were present, all 3 samples possessed the GATA-E box-GATA motif (Anderson KP, unpublished data).

Various investigators have now provided evidence for the formation of a transcription factor complex in erythroid cells that apparently contacts the DNA through the GATA1 and Tal1 proteins.<sup>21-25</sup> The exact stoichiometry and complete list of components is still unclear, however. For example, gel mobility shift assays have been used to demonstrate the presence of GATA1 in these complexes, and our results provide evidence for the functional role of the GATA sites in the EKLF promoter. However, it is also known that EKLF and a number of other erythroid-specific genes are still expressed in GATA1 null embryos.<sup>38</sup> One explanation for this apparent inconsistency is compensation by GATA2 because the levels of this protein are increased 50 times more than those seen in wild-type embryos.<sup>38</sup> Whether GATA2 is normally also involved in complex formation and gene regulation is still an unanswered question.

The GEG site in the EKLF gene represents one of the first examples of a potential binding site for this complex in a hematopoietic promoter. Expression of the EKLF gene is a critical



**Figure 3. Ectopic expression infrequently occurred in some mutant constructs, whereas expression from the wild-type EKLF promoter remained tissue specific.** The tissue specificity was assayed for 1 of the E box mutant lines expressed in adult blood (see Table 1) and was compared against the expression from a wild-type construct. Tissues collected from adult F1 animals were assayed for CAT activity as described.

event in hematopoiesis, and thus the presence of a GATA-E box-GATA motif in the promoter of this gene is intriguing. The finding that both the GATA and the E box binding sites are required for expression of this gene *in vivo* further substantiates the view that multiple proteins must interface with the DNA and each other to achieve correct transcriptional activation.

## Acknowledgments

We thank Jon Neumann and Karen Yaeger for the generation of the transgenic mice. We also thank Maureen Lehmann for maintaining the mouse colonies.

## References

- Miller IJ, Bieker JJ. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. *Mol Cell Biol.* 1993;13:2776.
- Anderson KP, Kern CB, Crable SC, Lingrel JB. Isolation of a gene encoding a functional zinc finger protein homologous to EKLF: identification of a new multigene family. *Mol Cell Biol.* 1995;15:5957.
- Crossley M, Whitelaw E, Perkins A, Williams G, Fujiwara Y, Orkin SH. Isolation and characterization of the cDNA encoding BKLF/TEF-2, a major CACCC-box-binding protein in erythroid cells and selected other cells. *Mol Cell Biol.* 1996;16:1695.
- Shields JM, Christy RJ, Yang VW. Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. *J Biol Chem.* 1996; 271:20,009.
- Garrett-Sinha LA, Eberspaecher H, Seldin MF, de Crombrugge B. A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. *J Biol Chem.* 1996;271:31,384.
- Conkright MD, Wani MA, Anderson KP, Lingrel JB. A gene encoding an intestinal-enriched member of the Kruppel-like factor family expressed in intestinal epithelial cells. *Nucl Acids Res.* 1999; 27:1263.
- Koritschoner NP, Bocco JL, Panzetta-Dutari GM, Dumur CI, Flury A, Patrino LC. A novel human zinc finger protein that interacts with the core promoter element of a TATA box-less gene. *J Biol Chem.* 1997;272:9573.
- Matsumoto N, Laub F, Aldabe R, et al. Cloning the cDNA for a new human zinc finger protein defines a group of closely related Kruppel-like transcription factors. *J Biol Chem.* 1998;273:28,229.
- Asano H, Li XS, Stamatoyannopoulos G. FKLf, a novel Kruppel-like factor that activates human embryonic and fetal beta-like globin genes. *Mol Cell Biol.* 1999;19:3571.
- Imataka H, Sogawa K, Yasumoto K, et al. Two regulatory proteins that bind to the basic transcription element (BTE), a GC box sequence in the promoter region of the rat P-4501A1 gene. *EMBO J.* 1992;11:3663.
- Fautsch MP, Vrabel A, Rickard D, Subramaniam M, Spelsberg TC, Wieben ED. Characterization of the mouse TGFbeta-inducible early gene (TIEG): conservation of exon and transcriptional regulatory sequences with evidence of additional transcripts. *Mamm Genome.* 1998;9:838.
- Cook T, Gebelein B, Mesa K, Mladek A, Urrutia R. Molecular cloning and characterization of TIEG2 reveals a new subfamily of transforming growth factor-beta-inducible Sp1-like zinc finger-encoding genes involved in the regulation of cell growth. *J Biol Chem.* 1998;273:25,929.
- Imhof A, Schuierer M, Werner O, et al. Transcriptional regulation of the AP-2alpha promoter by BTEB-1 and AP-2rep, a novel wt-1/egr-related zinc finger repressor. *Mol Cell Biol.* 1999;19:194.
- Southwood CM, Downs KM, Bieker JJ. Erythroid Kruppel-like factor exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. *Dev Dyn.* 1996;206:248.
- Perkins AC, Sharpe AH, Orkin SH. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature.* 1995;375:318.
- Nuez B, Michalovich D, Bygrave A, Ploemacher R, Grosveld F. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature.* 1995;375:316.
- Donze D, Townes TM, Bieker JJ. Role of erythroid Kruppel-like factor in human gamma-to-beta gene switching. *J Biol Chem.* 1995;270:1955.
- Bieker JJ, Southwood CM. The erythroid Kruppel-like factor transactivation domain is a critical component for cell-specific inducibility of a beta-globin promoter. *Mol Cell Biol.* 1995;15:852.
- Wijgerde M, Gribnau J, Nuez B, Philipsen S, Grosveld F, Fraser P. The role of EKLF in human beta-globin gene competition. *Genes Dev.* 1997;10:2894.
- Anderson KP, Crable SC, Lingrel JB. A multi-protein complex binding to a GATA-E-box-GATA motif regulates the erythroid Kruppel-like factor (EKLF) gene. *J Biol Chem.* 1998;273:14,347.
- Valge-Archer VE, Osada H, Warren A, et al. The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 are present in a complex in erythroid cells. *Proc Natl Acad Sci U S A.* 1994;91:8617.
- Osada H, Grutz G, Axelson H, Forster A, Rabbits TH. Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proc Natl Acad Sci U S A.* 1995;92:9585.
- Wadman IA, Osada H, Grutz GG, et al. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1, and Ldb1/NL1 proteins. *EMBO J.* 1997;16:3145.
- Begley CG, Green AR. The SCL gene: from case report to critical hematopoietic regulator. *Blood.* 1999;93:2760.
- Vyas P, McDevitt MA, Cantor AB, Katz SG, Fujiwara Y, Orkin SH. Different sequence requirements for expression in erythroid and megakaryocytic cells within a regulatory element upstream of the GATA-1 gene. *Development.* 1999;126:2799.
- Rabbits TH. Chromosomal translocations in human cancer. *Nature.* 1994;372:143.
- Baer R. TAL1, TAL2, and LYL1: a family of basic helix-loop-helix proteins implicated in T cell acute leukaemia. *Semin Cancer Biol.* 1993;4:341.
- Boehm T, Foroni L, Kennedy M, Rabbits TH. The rhombotin gene belongs to a class of transcriptional regulators with a potential novel protein dimerisation motif. *Oncogene.* 1990;5:1103.
- Rabbits TH, Boehm T. LIM domains. *Nature.* 1990;46:418.
- Visvader JE, Mao X, Fujiwara Y, Hahn K, Orkin SH. The LIM-domain binding protein Ldb1 and its partner LMO2 act as negative regulators of erythroid differentiation. *Proc Natl Acad Sci U S A.* 1997;94:13,707.
- Warren AJ, Colledge WH, Carlton MBL, Evans MJ, Smith AJH, Rabbits TH. The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell.* 1994;78:45.
- Shivdasani RA, Mayer E, Orkin SH. Absence of blood formation in mice lacking the T-cell leukemia oncoprotein tal-1/SCL. *Nature.* 1995;373:432.
- Robb L, Lyons I, Li R, et al. Absence of yolk sac hematopoiesis from mice with a targeted disruption of the *scl* gene. *Proc Natl Acad Sci U S A.* 1995;92:7075.
- Pevny L, Simon MC, Robertson E, et al. Erythroid differentiation in chimeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature.* 1991;349:257-260.
- Simon MC, Pevny L, Wiles M, Keller G, Constantini F, Orkin SH. Rescue of erythroid development in gene targeted GATA-1<sup>-</sup> mouse embryonic stem cells. *Nat Genet.* 1992;1:92-98.
- Hagen B, Constantini F, Lacy E. *Manipulating the Mouse Embryo.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory; 1986.
- Chen X, Reitman M, Bieker JJ. Chromatin structure and transcriptional control elements of the erythroid Kruppel-like factor (EKLF) gene. *J Biol Chem.* 1998;273:25,031.
- Weiss MJ, Yu C, Orkin SH. Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol Cell Biol.* 1997;17:1642.