Association of Ddx5/p68 protein with the upstream erythroid enhancer element (EHS1) of the Klf1 gene

Xiaoyong Chen[1], Sanjana Pillay[1], Felix Lohmann[1], and James J Bieker[1,2,3,4]

From the [1]Department of Cell, Developmental, and Regenerative Biology, Mount Sinai School of Medicine, New York, NY, [2]Black Familly Stem Cell Institute, [3]Tisch Cancer Institute, [4]Mindich Child Health and Development Institute, Mount Sinai School of Medicine, New York, NY

Author for correspondence:
James J Bieker, PhD
Mount Sinai School of Medicine
Department of Cell, Developmental, and Regenerative Biology
One Gustave L Levy Pl – Box 1020
New York, NY. 10029
USA
212-241-5067
james.bieker@mssm.edu

Present addresses: Xiaoyong Chen, Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut, USA; Felix Lohmann, Novartis Institutes for BioMedical Research, Basel, Switzerland

Keywords: transcription factor, biochemical affinity isolation, erythropoiesis, gene expression, transcription

Short title: Ddx5 involvement in Klf1 expression

Abstract

EKLF/KLF1 is an essential transcription factor that plays a global role in erythroid transcriptional activation. It's own regulation is of interest, as it displays a highly restricted expression pattern, limited to erythroid cells and its progenitors. Here we use biochemical affinity purification to identify the Ddx5/p68 protein as an activator of KLF1 by virtue of its interaction with the erythroid-specific DNAse hypersensitive site upstream enhancer element (EHS1). We postulate that its range of interactions with other proteins known to interact with this element render it part of the enhanseosome complex critical for optimal expression of KLF1. These individual interactions provide quantitative contributions that, in sum, establish high level activity of the KLF1 promoter and suggest they can be selectively manipulated for clinical benefit.

Introduction

The mechanisms by which intracellular transcriptional regulators interact to direct hematopoietic stem cells towards a particular lineage and exert control in target expression remains a major question in cellular and molecular studies (1-4). Analysis of the erythroid lineage has led to the successful characterization of regulators that act as transcription factors and establish the proper network to generate erythroid-specific expression. However, in many cases it remains unresolved how these factors are themselves regulated.

One way to address these issues is to isolate genomic clones that code for the gene-specific regulators and determine their own cis-acting regulatory elements and the trans-acting factors that bind to them. Based on this notion, we have been studying the regulation of erythroid Krüppel-like factor (EKLF;KLF1 (5)), an erythroid-enriched transcription factor that is intimately involved in global regulation of downstream erythroid-specific gene expression by binding to cognate CCMCRCCCN elements (6-9). A number of functional properties and expression characteristics of KLF1 make it of interest to study its regulation.

First, KLF1 expression remains tissue-restricted throughout early development and in the adult. Its onset is strictly limited to the mesodermal, primitive erythroid cells that populate the blood islands of the extraembryonic yolk sac at the early headfold stage (E7.5), switching by E9.5 to definitive cells within the hepatic primordia, and then to the red pulp of the adult spleen (10). It is expressed in the erythroblastic island macrophage, an in vivo niche that supplies a supportive environment for developing red cells (11-13). Early in hematopoietic differentiation KLF1 is expressed at low levels in the MPP and it retains an expression pattern restricted to the CMP and MEP as its transcript levels increase prior to eventual segregation to erythroid progeny (14-16).

Second, KLF1's activation target repertoire has expanded beyond the ß-like globin gene locus to include protein-stabilizing, heme biosynthetic pathway, red cell membrane protein, cell cycle, and transcription factor genes in both primitive and definitive erythroid cells (6-8,17). Relatedly, links have been established between mutant or haploinsufficient levels of KLF1 and altered human hematology and anemia (6,18,19), as some genes are uniquely sensitive to haploinsufficiency (20-22).

We have shown that a 950 base pair region, located just upstream of the KLF1 transcription initiation site, is sufficient to generate erythroid-specific expression in transient assays (23). This region exhibits the most significant homology upon a seven species alignment of 30 kB of surrounding genomic DNA (24), and harbors erythroid-restricted DNAse hypersensitive sites. One of these sites (erythroid hypersensitive site 1; EHS1) behaves as a very strong enhancer, which in conjunction with the proximal promoter (25-27) accounts for KLF1's tissue-specific expression (23). The importance of this short region has been verified *in vivo* (28-30). We have suggested a two-tiered mechanism for transcriptional regulation of KLF1, with GATA2 and SMAD5 proteins initially generating low transcript

levels, followed by high quantities of KLF1 transcript after GATA1 protein is produced (24). Coupled to inclusion of the DEK protein into this complex along with erythroid histone marks (31) and the RIOK2 protein (32), EHS1 can now be considered a well-characterized erythroid enhancer.

Our identification of DEK relied on a biochemical affinity purification to a region of EHS1 denoted as 'oligo2' region of EHS1 (31). In the present study, we have identified a second protein that also binds to that region in vitro and contributes to optimal enhancement activity by EHS1 in vivo.

Results

We had utilized a magnetic bead approach based on DNA affinity technology, coupled with mass spec analysis of the selected, associated proteins, to isolate proteins that bind the "oligo2" sequence within EHS1. This enabled identification of the "35kD" protein as DEK (31). We next focused our attention on the oligo2-enriched protein at ~60kD. MALDI/TOF isolation of peptides and their sequence analysis yielded a particularly significant hit with the p68/Ddx5 protein (Figure 1).

There is a considerable literature on Ddx5 protein and its interactions with nucleic acids (33,34). Ddx5 is a generally expressed 'DEAD box' family protein that interacts with DNA via chromatin, and plays a role in a range of RNA processing complexes. Ddx5 contains an ATP-dependent RNA helicase activity that is an important part of the Drosha complex (35,36), although evidence suggests that this activity is dispensable for its transcriptional regulatory function (33,34). Disruption leads to mouse embryonic lethality by E11.5 (35).

Ddx5 is well expressed within erythroid cells at all stages of differentiation, and in progenitors (Figure 2A). One interesting aspect in this case (compared to DEK) is that there is no evidence that p68/Ddx5 interacts with a specific DNA sequence. Regardless, we performed a chromatin immunoprecipitation (ChIP) test for Ddx5 interaction with the EHS1 region in MEL cells, and find that it binds to a high level (relative to IgG). Use of a variant MEL line that has deleted the endogenous 'oligo2' site shows that Ddx5 binding decreases ~2.5 fold (Figure 2B). These data are consistent with prior published studies on the importance of the 'oligo2' site, which showed that it is required for optimal binding of transcriptional activators such as P300 and Tal1, and that its removal leads to a quantitative (2-fold) reduction in production of KLF1 RNA (31).

Of relevance to this point, Ddx5 interacts with P300 and CBP (37), and is part of a hematopoietic cell complex that forms with Scl/Tal1 (38), a protein already known to interact with the KLF1 EHS1 region (39,40). Ddx5 also interacts with Smad5 (36,41), a protein we and others have shown is critical for KLF1 induction (24,42). As a result, it is relatively straightforward to envision the Ddx5 protein's role in KLF1 transcriptional activation via EHS1 (Figure 3).

Discussion

We have identified p68/Ddx5 as a novel protein that interacts with the 'oligo2' region of EHS1. Interestingly, Ddx5 performs much of its function via the long noncoding (lnc) RNA molecule SRA (43,44), an interaction that is important for insulator control of gene expression (45). Ddx5 plays a necessary role in skeletal muscle differentiation, in part by complexing with and stabilizing the activating myoD/Brg1 interaction and aiding the recruitment of TBP and RNA pol II to muscle gene promoters (46). SRA has been shown to regulate global erythroid gene expression in primary human cells and cell lines (47).

Based on these properties, we hypothesize that Ddx5, in a complex with other proteins at EHS1 that likely form an enhanseosome (Figure 3), plays a directive role in establishing a 3-dimensional structure at the KLF1 genomic locus. This is consistent with the chromatin architecture and long-range KLF1 genetic region interaction points observed in erythroid cells (48). Related to this point, Ddx5 also interacts with CTCF (45), a protein that plays a critical role in higher order chromatin structure and

enhancer insulation (49-51). RNA pull-down of SRA recovers Ddx5 and CTCF in a complex (47,52). CTCF is already a well-known player in erythroid gene regulation, as it is required for proper looping at the ß-globin, SCL/TAL1, and myb transcription units (53-57).

A critical aspect from the extended studies of KLF1 gene regulation is that although removal of the complete EHS1 site leads to a dramatic 50-fold reduction in KLF1 expression, removal of the 18-bp 'oligo2' site leads to a more nuanced effect (Figure 3). These properties demonstrate it is possible to identify genetic modifications that quantitatively alter KLF1 levels, thus mimicking haploinsufficiency and potentially providing a clinically useful approach to relieve hemoglobinopathies by its effect of altering ß-like globin gene regulation (20-22).

Experimental Procedures

32DEpo1 and MEL cells were grown as previously described (23,58). The MEL line containing a genetic deletion of the oligo2 sequence has been described ("indel 8"; (31)).

EHS1-binding proteins were purified as previously described (31), focused on 'oligo2' (23). Briefly, M280 magnetic Dynabeads coupled with streptavidin (Life Technologies) were incubated with biotinylated oligo2 to create the oligo-Dynabeads. The 32DEpo1 cell lysate (23) from >10 liters of cells was added and incubation continued at room temperature with rotation. Using a magnetic apparatus, the beads were washed to remove most of the non-specific binding proteins before elution in high salt. Gel shift assays were used throughout the procedure to estimate enrichment (31). Eluted protein was dialyzed and concentrated by acetone precipitation prior to SDS-PAGE electrophoresis. Proteins were visualized by staining with colloidal blue and bands of interest were excised. Mass spec 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 (59).

Chromatin immunoprecipitation (ChIP) was performed using the anti-Ddx5 antibody (Bethyl A300-523A; (60)). 10 X 10^6 cells (per replicate) were crosslinked in 1% formaldehyde at room temperature for 10 min and quenched with 0.125M glycine for 5 mins at room temperature followed by 15 mins on ice. The cell pellet was washed twice with PBS to remove excess formaldehyde and resuspended in 1mL of MNase lysis buffer (0.2% NP-40, 10mM NaCl & 10mM Tris pH 8.0) and kept on ice for 15 mins. Cells were spun and resuspended in 500uL of MNase lysis buffer with 0.1M $CaCl_2$ and 2uL of MNase (20U/uL) and incubated for 10 min at 37^0C. 10uL of 0.1M EDTA was added to stop the reaction and the cells were spun down (4^0C) and resuspended in 400uL of RIPA buffer (Cell Signaling Tech) with 1X protease inhibitor cocktail (Roche) and sonicated in a Bioruptor device (22 pulses of 40s on and 60s off at 21 amplitude). The cells were diluted with 1X RIPA buffer to a total volume of 1000uL and centrifuged at 13000 rpm, and 50uL of the was saved as input (20%). The rest of the supernatant was incubated with the anti-Ddx5 antibody (5ug per replicate) and rabbit IgG (Life Technologies) overnight at 4^0C on a rotating stand. Protein A Dynabeads (Thermo Scientific) were added the following day and incubated for another 2hrs at 4^0C on a rotating stand. The beads were collected using a magnetic stand and the supernatant was discarded. The beads were resuspended and first washed with 1X RIPA buffer (3 times, 5 min each) and then once in TSEII (0.1% SDS, 10% Triton X-100, 0.5M EDTA, 1M Tris pH 8 and 5M NaCl) and twice in TSEIII buffer (1% NP-40, 1% Na-deoxycholate, 0.5M EDTA, 1M Tris pH 8 and 0.25M LiCl). Final washes were done in TE buffer (0.5M EDTA and 1M Tris pH 8) and the DNA was then eluted from the beads at 65^0C for 20 min in 200uL of elution buffer (0.2M NaCl, 1% SDS and 1X TE). After the incubation the beads were separated using a magnet and the eluate was transferred to a fresh tube and crosslinked using proteinase K overnight at 65^0C. Next day, DNA was purified by phenol:chloroform extraction and the resulting pellet was resuspended in 30uL (IP) and 50uL (input). For qPCR, both the ChIP and input was diluted 1:3, and 3uL of this was used for amplification per reaction (10uL total). The primer pair for quantitation was designed using IDT software

and is as follows: 5'GCTCAGACCTCAACACAACA (forward) and 5'TGTCTGATGATGCCTGAAAGG (reverse).

Databases used as part of this study include transcription factor (61,62), and erythroid expression (63) sources.

Data availability: all data are contained within the manuscript.

Acknowledgements

This work was supported by NIH grants DK048721 and DK046865 to JJB.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Figure legends

Figure 1. Analysis of affinity-purified proteins.

DNA binding proteins that interact with oligo2 were monitored by SDS/PAGE analysis of ^{35}S methionine-labeled proteins after oligo1 (negative control (23)) or oligo2 affinity purification. Two non-adjacent lanes from the same gel are shown; a vertical line marks the boundary between the two. Purified material derived from ~5X10E6 cell equivalents are shown. Arrowhead on right marks the protein unique to oligo 2 preparation that is the focus of this study, along with the mass spec data (Sonar MS/MS analysis: a:b:y ratio is of the fragmentation ions; $^{z}m/z^{m-a}$ is the ^{charge}mass/charge$^{measured\ minus\ calculated}$ mass; the vertical bar between amino acid pairs indicates the ion intensity within the peptide fragment (59)) and peptide sequences used in its identification; lines on left indicate molecular weight as indicated (kD).

Figure 2. Erythroid expression of Ddx5/p68.

A) *Top*, data from the Erythron database (63) shows expression levels of Ddx5/p68 during mouse primitive and definitive (fetal liver or bone marrow) differentiation. P=proerythroblasts, B=basophilic erythroblasts, O=orthochromatiphilic erythroblasts, R=reticulocytes.

Bottom, browser data of RNA-seq expression in erythroid progenitors and erythroblast cells.

B) In vivo chromatin immunoprecipitation of Ddx5 protein or isotype IgG binding to the EHS1 site in WT MEL cells and MEL cells containing a genomic deletion centered on oligo2 (oligo2 sequence (23) is underlined, and the extent of deletion on ∆oligo2 MEL line (31) is as indicated). Results from two experiments each performed in triplicate are shown.

Figure 3. Sequence homology of KLF1 EHS1.

Detailed layout of conserved *Klf1* cis-regulatory elements at the upstream enhancer (EHS1) (24,31). Blocks of conserved sequence homology between seven mammalian species are color-coded for their transcription factor binding sites, highlighting Smad, DEK, Gata-Ebox-Gata sites; location of oligo2 used in this study (based on (23)) is as indicated.

Also drawn in are schematics of potential enhanceosome components at EHS1: *Above* includes the known multicomponent Gata/Tal1/Lmo2/Ldb1 multi-protein complex (based on (3,39,64-66)); *below* brings in known Ddx5 protein interactions as discussed in the text, including the recently-described interaction of Riok2 with this region (32). Also indicated are the extent of deletions that lead to altered levels of Klf1 expression: the 49 bp deletion of -715 to -666 (23) that leads to a 50-fold drop (30), and the 18 bp deletion of oligo2 (23) that leads to a 2-fold drop (31).

References

1. Novershtern, N., Subramanian, A., Lawton, L. N., Mak, R. H., Haining, W. N., McConkey, M. E., Habib, N., Yosef, N., Chang, C. Y., Shay, T., Frampton, G. M., Drake, A. C., Leskov, I., Nilsson, B., Preffer, F., Dombkowski, D., Evans, J. W., Liefeld, T., Smutko, J. S., Chen, J., Friedman, N., Young, R. A., Golub, T. R., Regev, A., and Ebert, B. L. (2011) Densely interconnected transcriptional circuits control cell states in human hematopoiesis. *Cell* **144**, 296-309
2. Hewitt, K. J., Sanalkumar, R., Johnson, K. D., Keles, S., and Bresnick, E. H. (2014) Epigenetic and genetic mechanisms in red cell biology. *Curr Opin Hematol* **21**, 155-164
3. Xu, J., Shao, Z., Glass, K., Bauer, D. E., Pinello, L., Van Handel, B., Hou, S., Stamatoyannopoulos, J. A., Mikkola, H. K., Yuan, G. C., and Orkin, S. H. (2012) Combinatorial assembly of developmental stage-specific enhancers controls gene expression programs during human erythropoiesis. *Dev Cell* **23**, 796-811
4. Laurenti, E., and Gottgens, B. (2018) From haematopoietic stem cells to complex differentiation landscapes. *Nature (London)* **553**, 418-426
5. Miller, I. J., and Bieker, J. J. (1993) A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Krüppel family of nuclear proteins. *Mol. Cell Biol.* **13**, 2776-2786
6. Siatecka, M., and Bieker, J. J. (2011) The multifunctional role of EKLF/KLF1 during erythropoiesis. *Blood* **118**, 2044-2054
7. Tallack, M. R., and Perkins, A. C. (2010) KLF1 directly coordinates almost all aspects of terminal erythroid differentiation. *IUBMB Life* **62**, 886-890
8. Yien, Y. Y., and Bieker, J. J. (2013) EKLF/KLF1, a tissue-restricted integrator of transcriptional control, chromatin remodeling, and lineage determination. *Mol Cell Biol* **33**, 4-13
9. Mukherjee, K., and Bieker, J. J. (2022) EKLF/Klf1 regulates erythroid transcription by its pioneering activity and selective control of RNA Pol II pause-release. *Cell Rep* **41**, 111830
10. Southwood, C. M., Downs, K. M., and Bieker, J. J. (1996) Erythroid Kruppel-like Factor (EKLF) exhibits an early and sequentially localized pattern of expression during mammalian erythroid ontogeny. *Devel. Dyn.* **206**, 248-259
11. Xue, L., Galdass, M., Gnanapragasam, M. N., Manwani, D., and Bieker, J. J. (2014) Extrinsic and intrinsic control by EKLF (KLF1) within a specialized erythroid niche. *Development* **141**, 2245-2254
12. Li, W., Wang, Y., Zhao, H., Zhang, H., Xu, Y., Wang, S., Guo, X., Huang, Y., Zhang, S., Han, Y., Wu, X., Rice, C. M., Huang, G., Gallagher, P. G., Mendelson, A., Yazdanbakhsh, K., Liu, J., Chen, L., and An, X. (2019) Identification and transcriptome analysis of erythroblastic island macrophages. *Blood*
13. Mukherjee, K., Xue, L., Planutis, A., Gnanapragasam, M. N., Chess, A., and Bieker, J. J. (2021) EKLF/KLF1 expression defines a unique macrophage subset during mouse erythropoiesis. *Elife* **10**
14. Frontelo, P., Manwani, D., Galdass, M., Karsunky, H., Lohmann, F., Gallagher, P. G., and Bieker, J. J. (2007) Novel role for EKLF in megakaryocyte lineage commitment. *Blood* **110**, 3871-3880
15. Cui, K., Zang, C., Roh, T. Y., Schones, D. E., Childs, R. W., Peng, W., and Zhao, K. (2009) Chromatin signatures in multipotent human hematopoietic stem cells indicate the fate of bivalent genes during differentiation. *Cell Stem Cell* **4**, 80-93
16. Li, B., Ding, L., Li, W., Story, M. D., and Pace, B. S. (2012) Characterization of the transcriptome profiles related to globin gene switching during in vitro erythroid maturation. *BMC Genomics* **13**, 153
17. Gnanapragasam, M. N., and Bieker, J. J. (2017) Orchestration of late events in erythropoiesis by KLF1/EKLF. *Curr Opin Hematol* **24**, 183-190
18. Singleton, B. K., Frayne, J., and Anstee, D. J. (2012) Blood group phenotypes resulting from mutations in erythroid transcription factors. *Curr Opin Hematol* **19**, 486-493
19. Tallack, M. R., and Perkins, A. C. (2013) Three fingers on the switch: Kruppel-like factor 1 regulation of gamma-globin to beta-globin gene switching. *Curr Opin Hematol* **20**, 193-200
20. Perkins, A., Xu, X., Higgs, D. R., Patrinos, G. P., Arnaud, L., Bieker, J. J., and Philipsen, S. (2016) Kruppeling erythropoiesis: an unexpected broad spectrum of human red blood cell disorders due to KLF1 variants. *Blood* **127**, 1856-1862
21. Waye, J. S., and Eng, B. (2015) Kruppel-like factor 1: hematologic phenotypes associated with KLF1 gene mutations. *Int J Lab Hematol* **37 Suppl 1**, 78-84
22. Caria, C. A., Faa, V., and Ristaldi, M. S. (2022) Kruppel-Like Factor 1: A Pivotal Gene Regulator in Erythropoiesis. *Cells* **11**
23. Chen, X., Reitman, M., and Bieker, J. J. (1998) Chromatin structure and transcriptional control elements of the erythroid Kruppel-like factor (EKLF) gene. *J Biol Chem* **273**, 25031-25040
24. Lohmann, F., and Bieker, J. J. (2008) Activation of Eklf expression during hematopoiesis by Gata2 and Smad5 prior to erythroid commitment. *Development* **135**, 2071-2082

25. Crossley, M., Tsang, A. P., Bieker, J. J., and Orkin, S. H. (1994) Regulation of the erythroid Kruppel-like factor (EKLF) gene promoter by the erythroid transcription factor GATA-1. *J. Biol. Chem.* **269**, 15440-15444
26. Anderson, K. P., Crable, S. C., and Lingrel, J. B. (1998) Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Kruppel-like factor (EKLF) gene. *J Biol Chem* **273**, 14347-14354
27. Anderson, K. P., Crable, S. C., and Lingrel, J. B. (2000) The GATA-E box-GATA motif in the EKLF promoter is required for in vivo expression. *Blood* **95**, 1652-1655.
28. Adelman, C. A., Chattopadhyay, S., and Bieker, J. J. (2002) The BMP/BMPR/Smad pathway directs expression of the erythroid-specific EKLF and GATA1 transcription factors during embryoid body differentiation in serum-free media. *Development* **129**, 539-549.
29. Xue, L., Chen, X., Chang, Y., and Bieker, J. J. (2004) Regulatory elements of the EKLF gene that direct erythroid cell-specific expression during mammalian development. *Blood* **103**, 4078-4083
30. Zhou, D., Liu, K., Sun, C. W., Pawlik, K. M., and Townes, T. M. (2010) KLF1 regulates BCL11A expression and gamma- to beta-globin gene switching. *Nat Genet* **42**, 742-744
31. Lohmann, F., Dangeti, M., Soni, S., Chen, X., Planutis, A., Baron, M. H., Choi, K., and Bieker, J. J. (2015) The DEK Oncoprotein Is a Critical Component of the EKLF/KLF1 Enhancer in Erythroid Cells. *Mol Cell Biol* **35**, 3726-3738
32. Ghosh, S., Raundhal, M., Myers, S. A., Carr, S. A., Chen, X., Petsko, G. A., and Glimcher, L. H. (2022) Identification of RIOK2 as a master regulator of human blood cell development. *Nat Immunol* **23**, 109-121
33. Fuller-Pace, F. V. (2006) DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation. *Nucleic Acids Res* **34**, 4206-4215
34. Fuller-Pace, F. V. (2013) The DEAD box proteins DDX5 (p68) and DDX17 (p72): multi-tasking transcriptional regulators. *Biochim Biophys Acta* **1829**, 756-763
35. Fukuda, T., Yamagata, K., Fujiyama, S., Matsumoto, T., Koshida, I., Yoshimura, K., Mihara, M., Naitou, M., Endoh, H., Nakamura, T., Akimoto, C., Yamamoto, Y., Katagiri, T., Foulds, C., Takezawa, S., Kitagawa, H., Takeyama, K., O'Malley, B. W., and Kato, S. (2007) DEAD-box RNA helicase subunits of the Drosha complex are required for processing of rRNA and a subset of microRNAs. *Nat Cell Biol* **9**, 604-611
36. Davis, B. N., Hilyard, A. C., Lagna, G., and Hata, A. (2008) SMAD proteins control DROSHA-mediated microRNA maturation. *Nature (London)* **454**, 56-61
37. Rossow, K. L., and Janknecht, R. (2003) Synergism between p68 RNA helicase and the transcriptional coactivators CBP and p300. *Oncogene* **22**, 151-156
38. Schuh, A. H., Tipping, A. J., Clark, A. J., Hamlett, I., Guyot, B., Iborra, F. J., Rodriguez, P., Strouboulis, J., Enver, T., Vyas, P., and Porcher, C. (2005) ETO-2 associates with SCL in erythroid cells and megakaryocytes and provides repressor functions in erythropoiesis. *Mol Cell Biol* **25**, 10235-10250
39. Su, M. Y., Steiner, L. A., Bogardus, H., Mishra, T., Schulz, V. P., Hardison, R. C., and Gallagher, P. G. (2013) Identification of biologically relevant enhancers in human erythroid cells. *J Biol Chem* **288**, 8433-8444
40. Wontakal, S. N., Guo, X., Smith, C., MacCarthy, T., Bresnick, E. H., Bergman, A., Snyder, M. P., Weissman, S. M., Zheng, D., and Skoultchi, A. I. (2012) A core erythroid transcriptional network is repressed by a master regulator of myelo-lymphoid differentiation. *Proc Natl Acad Sci U S A* **109**, 3832-3837
41. Warner, D. R., Bhattacherjee, V., Yin, X., Singh, S., Mukhopadhyay, P., Pisano, M. M., and Greene, R. M. (2004) Functional interaction between Smad, CREB binding protein, and p68 RNA helicase. *Biochem Biophys Res Commun* **324**, 70-76
42. Perna, F., Vu, L. P., Themeli, M., Kriks, S., Hoya-Arias, R., Khanin, R., Hricik, T., Mansilla-Soto, J., Papapetrou, E. P., Levine, R. L., Studer, L., Sadelain, M., and Nimer, S. D. (2015) The Polycomb Group Protein L3MBTL1 Represses a SMAD5-Mediated Hematopoietic Transcriptional Program in Human Pluripotent Stem Cells. *Stem Cell Reports*
43. Watanabe, M., Yanagisawa, J., Kitagawa, H., Takeyama, K., Ogawa, S., Arao, Y., Suzawa, M., Kobayashi, Y., Yano, T., Yoshikawa, H., Masuhiro, Y., and Kato, S. (2001) A subfamily of RNA-binding DEAD-box proteins acts as an estrogen receptor alpha coactivator through the N-terminal activation domain (AF-1) with an RNA coactivator, SRA. *EMBO J* **20**, 1341-1352
44. Caretti, G., Lei, E. P., and Sartorelli, V. (2007) The DEAD-box p68/p72 proteins and the noncoding RNA steroid receptor activator SRA: eclectic regulators of disparate biological functions. *Cell Cycle* **6**, 1172-1176
45. Yao, H., Brick, K., Evrard, Y., Xiao, T., Camerini-Otero, R. D., and Felsenfeld, G. (2010) Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev* **24**, 2543-2555

46. Caretti, G., Schiltz, R. L., Dilworth, F. J., Di Padova, M., Zhao, P., Ogryzko, V., Fuller-Pace, F. V., Hoffman, E. P., Tapscott, S. J., and Sartorelli, V. (2006) The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. *Dev Cell* **11**, 547-560
47. Sawaengdee, W., Cui, K., Zhao, K., Hongeng, S., Fucharoen, S., and Wongtrakoongate, P. (2020) Genome-Wide Transcriptional Regulation of the Long Non-coding RNA Steroid Receptor RNA Activator in Human Erythroblasts. *Front Genet* **11**, 850
48. Hua, P., Badat, M., Hanssen, L. L. P., Hentges, L. D., Crump, N., Downes, D. J., Jeziorska, D. M., Oudelaar, A. M., Schwessinger, R., Taylor, S., Milne, T. A., Hughes, J. R., Higgs, D. R., and Davies, J. O. J. (2021) Defining genome architecture at base-pair resolution. *Nature (London)* **595**, 125-129
49. Phillips, J. E., and Corces, V. G. (2009) CTCF: master weaver of the genome. *Cell* **137**, 1194-1211
50. Wallace, J. A., and Felsenfeld, G. (2007) We gather together: insulators and genome organization. *Curr Opin Genet Dev* **17**, 400-407
51. Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., Sanborn, A. L., Machol, I., Omer, A. D., Lander, E. S., and Aiden, E. L. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680
52. Wongtrakoongate, P., Riddick, G., Fucharoen, S., and Felsenfeld, G. (2015) Association of the Long Non-coding RNA Steroid Receptor RNA Activator (SRA) with TrxG and PRC2 Complexes. *PLoS Genet* **11**, e1005615
53. Stadhouders, R., Thongjuea, S., Andrieu-Soler, C., Palstra, R. J., Bryne, J. C., van den Heuvel, A., Stevens, M., de Boer, E., Kockx, C., van der Sloot, A., van den Hout, M., van Ijcken, W., Eick, D., Lenhard, B., Grosveld, F., and Soler, E. (2012) Dynamic long-range chromatin interactions control Myb proto-oncogene transcription during erythroid development. *EMBO J* **31**, 986-999
54. Hou, C., Zhao, H., Tanimoto, K., and Dean, A. (2008) CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proc Natl Acad Sci U S A* **105**, 20398-20403
55. Zhou, Y., Kurukuti, S., Saffrey, P., Vukovic, M., Michie, A. M., Strogantsev, R., West, A. G., and Vetrie, D. (2013) Chromatin looping defines expression of TAL1, its flanking genes, and regulation in T-ALL. *Blood* **122**, 4199-4209
56. Splinter, E., Heath, H., Kooren, J., Palstra, R. J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* **20**, 2349-2354
57. Lee, J., Krivega, I., Dale, R. K., and Dean, A. (2017) The LDB1 Complex Co-opts CTCF for Erythroid Lineage-Specific Long-Range Enhancer Interactions. *Cell Rep* **19**, 2490-2502
58. Siatecka, M., Xue, L., and Bieker, J. J. (2007) Sumoylation of EKLF Promotes Transcriptional Repression and Is Involved in Inhibition of Megakaryopoiesis. *Mol Cell Biol* **27**, 8547-8560
59. Field, H. I., Fenyo, D., and Beavis, R. C. (2002) RADARS, a bioinformatics solution that automates proteome mass spectral analysis, optimises protein identification, and archives data in a relational database. *Proteomics* **2**, 36-47
60. Suthapot, P., Xiao, T., Felsenfeld, G., Hongeng, S., and Wongtrakoongate, P. (2022) The RNA helicases DDX5 and DDX17 facilitate neural differentiation of human pluripotent stem cells NTERA2. *Life Sci* **291**, 120298
61. Pilon, A. M., Ajay, S. S., Kumar, S. A., Steiner, L. A., Cherukuri, P. F., Wincovitch, S., Anderson, S. M., Mullikin, J. C., Gallagher, P. G., Hardison, R. C., Margulies, E. H., and Bodine, D. M. (2011) Genome-wide ChIP-Seq reveals a dramatic shift in the binding of the transcription factor erythroid Kruppel-like factor during erythrocyte differentiation. *Blood* **118**, e139-148
62. Stamatoyannopoulos, J. A., Snyder, M., Hardison, R., Ren, B., Gingeras, T., Gilbert, D. M., Groudine, M., Bender, M., Kaul, R., Canfield, T., Giste, E., Johnson, A., Zhang, M., Balasundaram, G., Byron, R., Roach, V., Sabo, P. J., Sandstrom, R., Stehling, A. S., Thurman, R. E., Weissman, S. M., Cayting, P., Hariharan, M., Lian, J., Cheng, Y., Landt, S. G., Ma, Z., Wold, B. J., Dekker, J., Crawford, G. E., Keller, C. A., Wu, W., Morrissey, C., Kumar, S. A., Mishra, T., Jain, D., Byrska-Bishop, M., Blankenberg, D., Lajoie, B. R., Jain, G., Sanyal, A., Chen, K. B., Denas, O., Taylor, J., Blobel, G. A., Weiss, M. J., Pimkin, M., Deng, W., Marinov, G. K., Williams, B. A., Fisher-Aylor, K. I., Desalvo, G., Kiralusha, A., Trout, D., Amrhein, H., Mortazavi, A., Edsall, L., McCleary, D., Kuan, S., Shen, Y., Yue, F., Ye, Z., Davis, C. A., Zaleski, C., Jha, S., Xue, C., Dobin, A., Lin, W., Fastuca, M., Wang, H., Guigo, R., Djebali, S., Lagarde, J., Ryba, T., Sasaki, T., Malladi, V. S., Cline, M. S., Kirkup, V. M., Learned, K., Rosenbloom, K. R., Kent, W. J., Feingold, E. A., Good, P. J., Pazin, M., Lowdon, R. F., and Adams, L. B. (2012) An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome Biol* **13**, 418
63. Kingsley, P. D., Greenfest-Allen, E., Frame, J. M., Bushnell, T. P., Malik, J., McGrath, K. E., Stoeckert, C. J., and Palis, J. (2013) Ontogeny of erythroid gene expression. *Blood* **121**, e5-e13

64. Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P. D., Dean, A., and Blobel, G. A. (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233-1244
65. Li, L., Freudenberg, J., Cui, K., Dale, R., Song, S. H., Dean, A., Zhao, K., Jothi, R., and Love, P. E. (2013) Ldb1-nucleated transcription complexes function as primary mediators of global erythroid gene activation. *Blood* **121**, 4575-4585
66. Ulirsch, J. C., Lacy, J. N., An, X., Mohandas, N., Mikkelsen, T. S., and Sankaran, V. G. (2014) Altered chromatin occupancy of master regulators underlies evolutionary divergence in the transcriptional landscape of erythroid differentiation. *PLoS Genet* **10**, e1004890

Figure 1

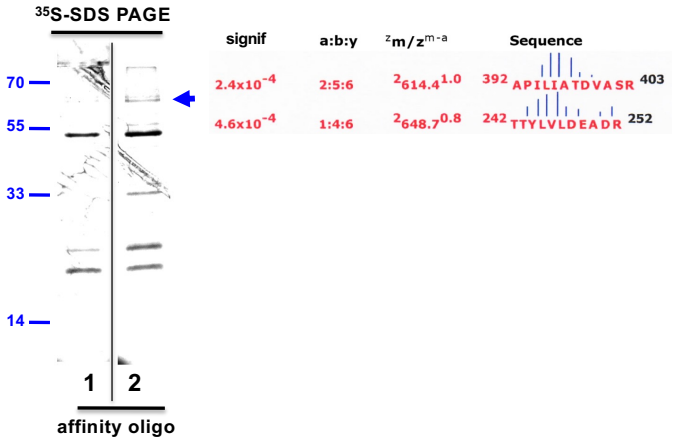

Figure 1. Analysis of affinity-purified proteins.
DNA binding proteins that interact with oligo2 were monitored by SDS/PAGE analysis of ^{35}S methionine-labeled proteins after oligo1 (negative control) or oligo2 affinity purification. Two non-adjacent lanes from the same gel are shown; a vertical line marks the boundary between the two. Purified material derived from ~5X10E6 cell equivalents are shown. Arrowhead on right marks the protein unique to oligo 2 preparation that is the focus of this study, along with the mass spec data (Sonar MS/MS analysis: a:b:y ratio is of the fragmentation ions; $^{z}m/z^{m-a}$ is the ^{charge}mass/charge$^{measured\ minus\ calculated}$ mass; the vertical bar between amino acid pairs indicates the ion intensity within the peptide fragment (59)) and peptide sequences used in its identification; lines on left indicate molecular weight as indicated (kD).

Figure 2

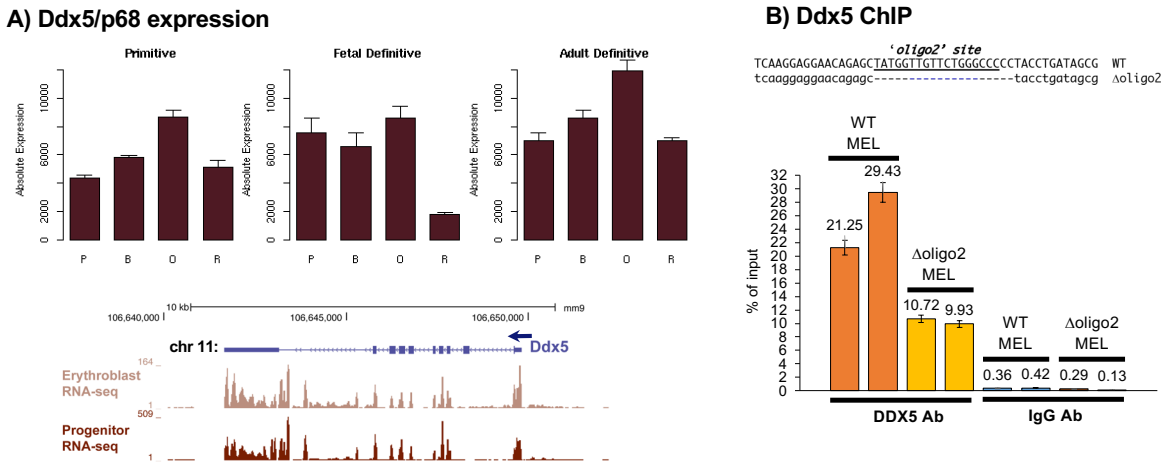

Figure 2. Erythroid expression of Ddx5/p68.
A) *Top*, data from the Erythron database (38) shows expression levels of Ddx5/p68 during mouse primitive and definitive (fetal liver or bone marrow) differentiation. P=proerythroblasts, B=basophilic erythroblasts, O=orthochromatiphilic erythroblasts, R=reticulocytes.
Bottom, browser data of RNA-seq expression in erythroid progenitors and erythroblast cells.
B) In vivo chromatin immunoprecipitation of Ddx5 protein or isotype IgG binding to the EHS1 site in WT MEL cells and MEL cells containing a genomic deletion centered on oligo2 (oligo2 sequence (23) is underlined, and the extent of deletion in the Δoligo2 MEL line (31) is as indicated) . Results from two experiments each performed in triplicate are shown.

Figure 3

EHS1=erythroid hypersensitive site 1

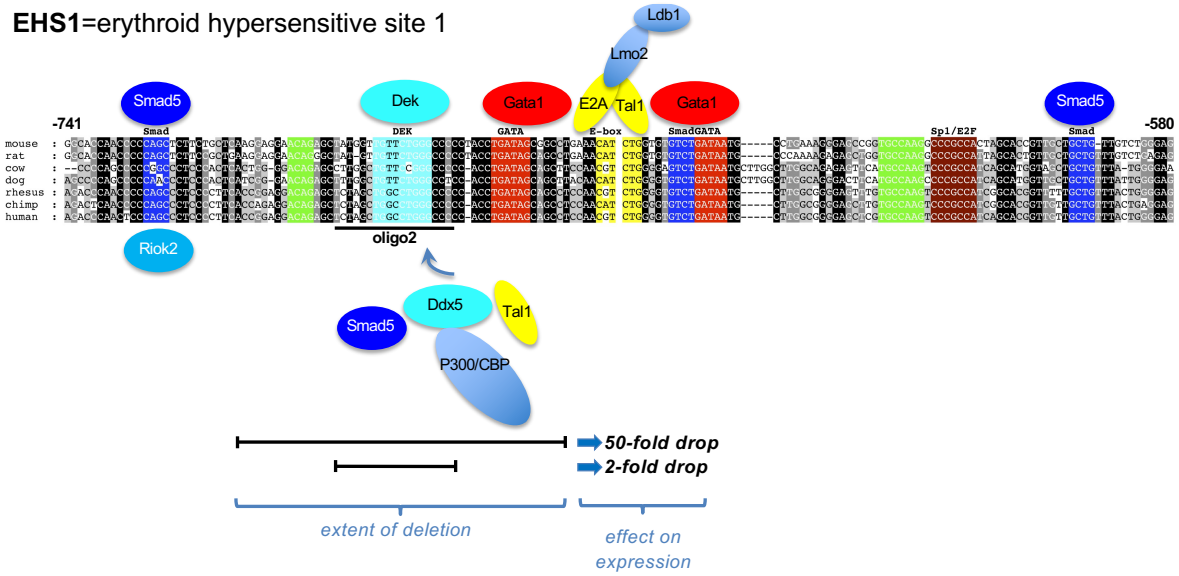

Figure 3. Sequence homology of KLF1 EHS1.
Detailed layout of conserved *Klf1* cis-regulatory elements at the upstream enhancer (EHS1) (24,31). Blocks of conserved sequence homology between seven mammalian species are color-coded for their transcription factor binding sites, highlighting Smad, DEK, Gata-Ebox-Gata sites; location of oligo2 used in this study (based on (23)) is as indicated.

Also drawn in are schematics of potential enhanceosome components at EHS1: *Above* includes the known multicomponent Gata/Tal1/Lmo2/Ldb1 multi-protein complex (based on (3,39,64-66)); *below* brings in known Ddx5 protein interactions as discussed in the text, including the recently-described interaction of Riok2 with this region (32). Also indicated are the extent of deletions that lead to altered levels of Klf1 expression: the 49 bp deletion of -715 to -666 (23) that leads to a 50-fold drop (30), and the 18 bp deletion of oligo2 (23) that leads to a 2-fold drop (31).