Characterization of a Megakaryocyte-specific Enhancer of the Key Hemopoietic Transcription Factor GATA1

Received for publication, March 3, 2006, and in revised form, March 20, 2006. Published, JBC Papers in Press, March 20, 2006, DOI 10.1074/jbc.M602052200

Boris Guyot¹, Kasumi Murai¹, Yuko Fujiwara, Veronica Valverde-Garduno, Michele Hammett, Sara Wells, Neil Dear, Stuart H. Orkin, Catherine Porcher, and Paresh Vyas²

From the Weatherall Institute of Molecular Medicine, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom

Specification and differentiation of the megakaryocyte and erythroid lineages from a common bipotential progenitor provides a well studied model to dissect binary cell fate decisions. To understand how the distinct megakaryocyte- and erythroid-specific gene programs arise, we have examined the transcriptional regulation of the megakaryocyte erythroid transcription factor GATA1. Hemopoietic-specific mouse (m)GATA1 expression requires the mGata1 enhancer mHS-3.5. Within mHS-3.5, the 3′-179 bp of mHS-3.5 are required for megakaryocyte but not red cell expression. Here, we show mHS-3.5 binds key hemopoietic transcription factors in vitro and is required to maintain histone acetylation at the mGata1 locus in primary megakaryocytes. Analysis of GATA1-LacZ reporter gene expression in transgenic mice shows that a 25-bp element within the 3′-179 bp in mHS-3.5 is critical for megakaryocyte expression. In vitro three DNA binding activities, A, B, and C bind to the core of the 25-bp element, and these binding sites are conserved through evolution. Activity A is the zinc finger transcription factor ZBP89 that also binds to other cis elements in the mGata1 locus. Activity B is of particular interest as it is present in primary megakaryocytes but not red cells. Furthermore, mutation analysis in transgenic mice reveals activity B is required for megakaryocyte-specific enhancer function. Bioinformatic analysis shows sequence corresponding to the binding site for activity B is a previously unrecognized motif, present in the cis elements of the FlI1 gene, another important megakaryocyte-specific transcription factor. In summary, we have identified a motif and a DNA binding activity likely to be important in directing a megakaryocyte gene expression program that is distinct from that in red cells.

Understanding the molecular basis of lineage specification from multipotential progenitors is a central question in biology. The question in its simplest and most tractable form is to understand how two different lineages arise from a common progenitor. Hemopoiesis has arguably been one of the most informative and well studied model systems in furthering our understanding of lineage determination. In hemopoiesis, the megakaryocytic and erythroid lineages have distinctive phenotypes and gene expression profiles, and yet arise from a common progenitor (1). The mechanism by which they are differentially specified has been extensively studied but is not well understood.

Though lineage specification is regulated by external cues that are modulated by intracellular signaling pathways, it ultimately culminates in activation of unilineage programs of gene expression and repression of genes associated with alternative cell fate. Coordination of complex patterns of gene expression resulting in lineage specification is thought to be regulated, in part, by combinations of lineage-specific transcription factors. Therefore, in this model, in a common megakaryocyte-erythroid bipotential progenitor, lineage-specific (erythroid or megakaryocyte) combinations of transcription factors become expressed that direct differential specification of the erythroid and megakaryocyte lineages. The erythroid and megakaryocytic lineages share many critical hemopoietic transcription factors (e.g. GATA1, FOG-1, Gfi-1b, NF-E2p45, and SCL/TAL-1) but also express hemopoietic regulators unique to one or other lineage (such as RUNX-1, Meis1, and Fli1 in the megakaryocyte lineage and EKLF in erythroid cells). However, the detailed mechanisms by which these lineages utilize combinations of transcriptional regulators to direct lineage-specific programs of gene expression are unclear.

One way to uncover the combination of regulators required for differential specification is to identify the DNA sequences (cis elements), and through them the DNA binding transcriptional regulators, required to specifically express genes in either red cells or megakaryocytes. In this study, we examine the cis elements required to direct expression of a key erythroid and megakaryocyte transcription factor GATA1, in megakaryocytes but not red cells.

GATA1 is first expressed at low levels in the common myeloid progenitor (2), and its expression is maintained in the megakaryocyte and erythroid lineages (reviewed in Ref. 3). In both lineages, sustained GATA1 expression is required for terminal maturation (4–7). GATA1 expression is regulated by a complex set of cis-acting regulatory elements. In mice, the hemopoietic promotor (mIE), an upstream enhancer 3.5 kilobases from the GATA1 hemopoietic transcription start site, HS1/G1HE/mHS-3.5 (hereafter referred to as mHS-3.5), and an element in the first mGata1 intron (HS 4/5 or mHS +3.5) (see Fig. 1A) are required to direct reporter gene expression to both erythroid cells and megakaryocytes in transgenic mice (8–10). Deletion of mHS-3.5 from the reporter construct extinguishes reporter gene expression in both red cells and megakaryocytes, highlighting a non-redundant enhancer function for mHS-3.5 in this assay (11, 12). In contrast, germline deletion of a 7-kb region of genomic DNA including all of mHS-3.5 (ΔneoΔHS mice), virtually abrogates megakaryocyte GATA1 expression but GATA1 expression in red cells is unaffected (5, 10), suggesting that mHS-3.5 plays a unique non-redundant role in megakaryocyte GATA1 expression, and that other elements in the mGata1 locus must compensate for loss of mHS-3.5 function in red cells.

Furthermore, deletion analysis showed that within mHS-3.5, two distinct DNA sequences are important for enhancer activity in transgenic mice. First, a GATA site is absolutely required for mHS-3.5 enhancer activity in both red cells and megakaryocytes (11, 12). In vitro, this site and a neighboring E-box DNA element bind GATA factors and
Megakaryocyte-specific GATA1 Expression

a pentameric hemopoietic transcription factor complex containing, at a minimum, GATA1-SCL/TAL-1-E2A-LMO2-LDB1 in erythroid and megakaryocytic cells (11). In vivo GATA1-SCL/TAL-1-E2A-LMO2-LDB1 binding is detected at mHS-3.5 in erythroid cells (3). Second, whereas the whole 312 bp of mHS-3.5 is required for megakaryocyte reporter gene expression, only the 5′ 133 bp is necessary for red cell reporter gene expression (11). These data suggest that the 3′ 179 bp of mHS-3.5 binds trans-acting factors that cooperate with proteins that bind the GATA site, to direct megakaryocyte-specific enhancer activity. Therefore, as a prelude to identifying trans-acting factors required for megakaryocyte-specific GATA1 expression, and megakaryocyte-specific gene activation in general, we set out to pinpoint the cis-acting sequences within mHS-3.5 mediating megakaryocyte-specific enhancer activity.

MATERIALS AND METHODS

Primary Cells and Cell Lines—All cell culture reagents were purchased from Invitrogen (Paisley, UK), unless indicated otherwise. ΔneoHS mice have been previously described (5). Briefly, a genomic region between 12,000 and 2600 bp upstream of the hemopoietic Gata1 transcription start site, and encompassing the mHS-3.5 element (Fig. 1B), was deleted by homologous recombination in ΔneoHS mice. To isolate mouse primary wild-type (WT)5 C57BL/6 or ΔneoHS megakaryocytes, mice were administered intraperitoneal 5-fluorouracil (5-FU, 0.15 g/kg, Sigma). After 8 days, bone marrow was flushed from the femurs and tibias and a LD depletion column (both from Miltenyi Biotech). More than 90% of the unlabeled cells were Ter119+ve red cells as assessed by flow cytometry and May-Grunevald Giemsa staining and expression of CD61 (BD Pharmingen), which was assessed by flow cytometry. Primary red cells were purified from the spleen of phylyhazidine-treated mice. Briefly, spleen cells labeled with unconjugated anti-B220 (cat. 553084), anti-Gr1 (cat. 553123), anti-Mac1 (cat. 553124) and anti-CD2 (cat. 553109) antibodies (all from BD Pharmingen, Oxford, UK) were removed using streptavidin-coated magnetic beads and a large cell depletion column (both from Miltenyi Biotech, Bergisch Galdbach, Germany). The remaining cell population was more than 95% megakaryocytes as assessed by May-Grunevald Giemsa staining and expression of CD61 (BD Pharmingen), which was assessed by flow cytometry. Primary red cells were purified from the spleen of phylyhazidine-treated mice. Briefly, spleen cells labeled with unconjugated anti-B220 (cat. 553084), anti-Gr1 (cat. 553123), anti-Mac1 (cat. 553308), and anti-CD2 (cat. 553109) antibodies (all from BD Pharmingen) were removed using goat anti-rat IgG coated magnetic beads and a LD depletion column (both from Miltenyi Biotech). More than 90% of the unlabeled cells were Ter119+ve red cells as assessed by flow cytometry and May-Grunevald Giemsa staining. Fetile livers were obtained from E13.5/14.5 C57BL/6 embryos. Culture and derivation of the murine megakaryoblastic leukemia cell line L8057 (14), C212 myeloblast cell line (15), and 3T3-L1 pre-adipocyte cell line (16) have been previously described.

Constructs—For constructs WT, GK3, GK8, and GS12 shown in Fig. 5, a PCR fragment extending from A to I (see Fig. 2B) was obtained using a common upstream primer 5′-TTGTTGACTGCAGATTTTGGATTCTGAGGGCTGCAAATGGGCTCCC-3′ and a specific downstream primer, which was either WT sequence or sequences corresponding to mutants GK3, GK8 or GS12 (see Fig. 3B). PCR products were then cloned into the Asp718 site of the 5′-3′-LacZ vector (10) using Asp718 sites that were introduced into the PCR primers. All constructs were sequenced prior to use.

Chromatin Immunoprecipitation (ChIP)—ChIP experiments on mouse primary megakaryocytes and erythroid cells were performed as previously described (3, 17), except that 5 × 10⁷ megakaryocytes were used for each immunoprecipitation. For the ZBP-89 ChIP, a polyclonal rabbit antiserum from Rockland (ref. 100–401-685, Gilbertsville, PA) was used.

Transgenics Procedures—Standard techniques were used to isolate transgene sequences for DNA purification and for pronuclear injection of CD-1 (Charles River Laboratories, MA and MRC Harwell, UK) and B6CBAF1/J (Jackson Laboratories, Bar Harbor, ME) fertilized eggs. Chi-meric fetuses (thereafter called transient transgenics) were sacrificed at E13.5–14.5, genotyped by PCR using LacZ and RapSyn primers as previously described (10) and analyzed for β-galactosidase expression as detailed below.

β-Galactosidase Assays—β-galactosidase expressing fetal liver cells (E13.5–14.5) were analyzed either visually by X-gal staining (Sigma) (10), or by flow-cytometry using fluorescein di-(β-d-galactopyranoside) (FDG, Sigma) and lineage-specific antibody staining. For FDG staining, fetal liver cells were loaded with FDG as previously described (18) and stained with biotin- or PE-conjugated antibodies directed against Ter119, Mac1, or CD61 surface markers or their respective isotype controls (all from BD Pharmingen). APC-conjugated streptavidin was used as secondary antibody for biotin-conjugated primary antibodies. FACS analysis was performed using a CyAn machine and Summit software (Dako Cytomation, Cambridge, UK). Hoechst 33258 Molecular Probes, Eugene, OR) was used to exclude dead cells.

Micrograph images were taken with a BX60 microscope and a 40/0.75 objective (Olympus, London, UK), using a Qicam camera (Qimaging, Burnaby, Canada) and Openlab software (Improvision, Coventry, UK).

Statistical Analysis of Reporter Gene Expression—Analysis to determine the statistical significance of differences in the frequencies of LacZ-expressing transgenic embryos when different mGata1-LacZ transgenes were tested was performed with a binomial test (Graphpad).

Electromobility Shift Assay (EMSA)—Nuclear extract preparations and DNA binding assays were performed as previously reported (11). Antibodies used were: anti-Sp1 (sc-59), anti-Sp3 (sc-644) (Santa Cruz Biotechnology, Calne, UK) and the anti-ZBP-89 antisera used in for ChIP. Sequences of oligonucleotides used in EMSA are shown in Fig. 3B.

Bioinformatics—Multispecies mHS-3.5 alignments were performed with ClustalW (MacVector Accearys, UK). The GenBank accession numbers for mouse, human, and dog Gata1 genes are AF 136574, AF 136573, and NW139919, respectively. Sequences of Fil1 genes were obtained from Ensembl database. Accession numbers for the Fil1 genes are: ENSMUSG00000016087 (Mus musculus), ENSRNOG0000008904 (Rattus norvegicus), ENSG00000151702 (Homo sapiens), and ENSP-TRG0000004467 (Pan troglodytes). 1500-bp upstream of the transcription start site(s) were analyzed using MacVector to look for sequences similar to the E1 region. Dog (Canis familiaris) and opossum (Monodelphis domestica) Fil1 genes are not annotated in the current releases and were located using BLAST software with the mouse Fil1 cDNA sequence and a highly conserved regulatory sequence in the Fil1 gene promoter (19).

RESULTS

mHS-3.5 is Required for Hyperacetylation of Histone H3 within the mGata1 Locus in Primary Megakaryocytes but Not Red Cells—To establish when during megakaryocytic differentiation mHS-3.5 is required for GATA1 expression, we isolated fetal primary common myeloid progenitors (CMP), megakaryocyte erythroid progenitors (MEP), megakaryocyte progenitors ( MkP), and primary megakaryo-

---

3 The abbreviations used are: WT, wild type; EMSA, electrophoretic mobility shift assay; X-gal, 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; CHIP, chromatin immunoprecipitation assay; FDG, fluorescein di-(β-D-galactopyranoside); FACS, fluorescent-activated cell sorting.
cytes from wild-type mice and mice with deletion of mHS-3.5 (ΔneoΔHS mice) and examined GATA1 mRNA levels by quantitative real-time Taqman PCR. GATA1 levels were similar in wild-type and ΔneoΔHS CMP and MEP but specifically decreased in ΔneoΔHS MkP and megakaryocytes to 5% of wild-type levels (data not shown and Ref. 20). This suggests that mHS-3.5 plays a critical non-redundant role at the level of a megakaryocyte progenitor and megakaryocyte precursors.

Next, we examined in vivo acetylation of histone H3 in WT and ΔneoΔHS primary megakaryocytes and red cells to determine how loss of mHS-3.5 affected chromatin structure in the two lineages (contrast Fig. 1, A with B). In both WT megakaryocytes and red cells there is a domain enriched for hyperacetylated histone H3 between the mIE promoter and mHS +3.5, the mGata1 intron cis element. In addition, a smaller peak of acetylation is seen at mHS-3.5. Lastly, a peak of enrichment of acetylated H3 is also detected at the ubiquitous DNase I hypersensitive sites mHS+20, that probably marks the promoter of the mHdac6 gene (3).

Importantly, only in megakaryocytes, but not red cells, is deletion of mHS-3.5 associated with a striking loss of the domain enriched in hyperacetylated histone H3 between mIE and mHS+3.5. This loss was specific to this region of the mGata1 locus as enrichment of hyperacetylation of H3 was maintained at mHS+20. These data highlight the critical, non-redundant role of mHS-3.5 in maintaining a domain of hyperacetylated chromatin in the mGata1 locus in and around the Gata1 gene in primary megakaryocytes but not red cells. In part this domain of acetylation may be maintained by binding of GATA1 and SCL (supplementary Fig. S1, A and B, respectively) to mHS-3.5 in primary megakaryocytes.

A 25-bp Element within mHS-3.5 Is Required for Megakaryocyte-specific Enhancer Activity—We then proceeded to more precisely map sequences within mHS-3.5 required for megakaryocyte enhancer activity. We previously showed that whereas a 317-bp region within mHS-3.5 (11) (Fig. 2A, construct A-D) was required for enhancer activity in megakaryocytes, the 5’ 133 bp (construct A-E) was active only in red cells and not megakaryocytes (Fig. 2A, construct A-E). In the region between E and D within mHS-3.5, there are a number of blocks of sequence conserved between human, dog, and mouse (Fig. 2B). Next, two additional 3’ mHS-3.5 deletional constructs were tested for enhancer activity in fetal liver cells from E13.5 transient transgenics embryos (Fig. 2A, constructs A-I and A-J). Both constructs were able to direct LacZ expression in fetal liver megakaryocytes, as well as red cells.

These data suggest that a 25-bp fragment within mHS-3.5 (hereafter called E-I), present in construct A-I but not A-E, is required to direct LacZ expression in fetal liver megakaryocytes, as well as red cells. This data is consistent with the region being conserved through evolution (Fig. 2). The E-I Region Binds a Sequence-specific DNA Binding Activity Present in Mouse Primary Megakaryocytes but Not Primary Red Cells—We then examined in vitro DNA binding by nuclear proteins using an EMSA with a probe encompassing sequences E-I (Fig. 3A), to identify potential DNA binding transcriptional regulators responsible for megakaryocyte-specific enhancer activity. Five retarded bands were observed using nuclear extracts from the megakaryoblast cell line, L8057 (Fig. 3A, lane 1). Given the high GC content of the E-I region we hypothesized that some of the retarded bands may reflect binding by the Sp family of transcription factors. This was confirmed by supershift assays, and two DNA-protein complexes contained Sp1 and Sp3 (Fig. 3A, lanes 2 and 3). However, as we could not detect in vivo binding, by chromatin immunoprecipitation, of either Sp1 or Sp3 to mHS-3.5 in either primary red cells or primary megakaryocytes, (though binding of both factors was detected at mHS+20) (data not shown), we did not further investigate the role of Sp1 and Sp3.

The three remaining DNA-protein complexes were called A, B, and C. We then investigated the tissue-specific expression of these
complexes by EMSA using nuclear extracts from E13.5 fetal liver (which is mainly composed of erythroid cells), primary fetal liver derived megakaryocytes, C2C12 myoblasts and 3T3-L1 adipocytes (Fig. 3A, lanes 4–7). A number of other cell types were also studied (data not shown). As expected we detected binding of transcription factors Sp1 and Sp3 in all nuclear extracts, given the ubiquitous expression of Sp1 and Sp3 (21). The other three DNA binding activities (A, B, and C) were expressed in both hemopoietic and non-hemopoietic cells with cell-type specific differences in the abundance of the binding activity. Note-worthy for this study was that complex B was barely detected in fetal liver (mainly erythroid cells) but was easily detectable in primary fetal megakaryocytes.

To map where Sp1, Sp3, A, B, and C were binding in vitro, within the DNA fragment E-I, a series of 2-bp scanning mutations were made in wild-type sequence to generate a series of mutant oligonucleotide Fig. (3B) that were used as competitor in EMSA assays with wild-type sequence as probe. By recording which probes failed to abrogate binding of the 5 activities (Sp1, Sp3, A, B, and C), we could delineate the in vitro DNA binding sites of Sp1/3, A, B, and C in E-I (Fig. 3C; the data are summarized in Fig. 3D). For example, mutant oligonucleotide GK3 failed to abrogate binding of A alone (Fig. 3C, lanes 8 and 9), suggestive that the nucleotides which had been mutated were important for binding of complex A. Similarly, oligonucleotide GK8 fails to compete binding of C specifically (Fig. 3C, lanes 19 and 20), suggesting that the mutated nucleotides in oligonucleotide GK8 were important for binding of complex C. Because of the overlapping nature of the binding sites for A, B, and C (Fig. 3, C and D), none of the mutations tested failed to specifically disrupt binding factor B. Note that GS12 fails to disrupt binding of B and C (Fig. 3C, lanes 24 and 25). Finally, we used the mutant oligonucleotides GK3, GK8, and GS12 as probes in EMSA assays (Fig. 3E). This confirms that in vitro GK3 binds complex A poorly (Fig. 3E, lanes 4–6), GK8 fails to bind complex C (Fig. 3E, lanes 10–12) and finally that no binding of complexes B and C is seen with GS12 (Fig. 3E, lanes 13–15).

ZBP-89 Zinc Finger Transcription Factor Interacts with mHS-3.5 E-I Region—As the identities of A, B, and C were uncharacterized, we asked if GATA1 protein partners could be binding to E-I, based on the hypothesis that the E-I region may cooperate with the critical GATA binding site at the 5′-end of mHS-3.5, (see introduction), and proteins at these sites may physically interact. Recently, the zinc finger transcription factor ZBP-89 (22) was reported in a complex with GATA1 and to function in hemopoiesis (23). In vitro, ZBP-89 antisera, but not preimmune serum, specifically disrupts binding of A to mHS-3.5 E-I region (Fig. 4A). In addition, ChIP experiments confirm in vivo ZBP-89 binding at mHS-3.5 in L8057 megakaryoblasts (Fig. 4B). Interestingly, ZBP-89 binding was also detected at the IE promoter and the mHS+3.5 cis element. This suggests that activity A is, or at least contains, ZBP89 and that ZBP-89 binds to mHS-3.5 E-I region in vivo.

FIGURE 2. A conserved 25-bp sequence is required for the megakaryocyte-specific enhancer activity of mHS-3.5 in transgenic mice. A, mouse Gata1 locus with the hemopoietic-specific IE promoter and five coding exons (depicted as block boxes) is shown on top. An arrow shows the location of mHS-3.5. Below, a map of the vector 5′-3′ LacZ is shown with the LacZ coding sequence fused in-frame with the initiator ATG of the Gata1 gene. Below left, four genomic DNA fragments (A-D to A-J) and their location in mHS-3.5 are shown: closed circle, GATA site, open circle, E-box, black box, conserved GC-rich sequence, gray boxes, mouse/human conserved sequences. These fragments were attached to the 5′-end of 5′-3′ LacZ to produce constructs A-D to A-J. To the right is shown the size of each attached fragment and a summary of B, the number of β-galactosidase expressing embryos/totai number transgenic embryos. An embryo was defined as expressing β-galactosidase if >0.5% of cells stained with X-gal substrate. Column 2 shows the range of β-galactosidase-expressing cells (%), 200 cells were counted. Results for A-D and A-E constructs are taken from Ref. 11 and is shown for clarity. Below, photographs of representative X-gal staining of fetal liver megakaryocytes (red arrowheads) and definitive erythroblasts (black arrows) from E14.5 embryos transgenic for A-I or A-J constructs. Original magnification is ×40. The black bar represents 10 μm. B, sequence alignment of mouse, human, and dog HS-3.5 elements. Positions of GATA1 and SCL/TAL-1 binding sites and borders of constructs A-D, A-E, A-I, and A-J are shown.
FIGURE 3. Sequence-specific DNA activities that bind the 25-bp region (E-I) in megakaryoblast nuclear extracts. A, protein-DNA complexes detected by EMSA bind the E-I region of mHS-3.5. In the first panel (lanes 1–3), nuclear extracts from L8057 cell were incubated with the WT probe (for the sequence see B), either alone (lanes 1) or with antibodies directed against Sp1 (lane 2) or Sp3 (lane 3). In the second panel, nuclear extracts of E13.5/14.5 mouse fetal liver (FL) (lane 4), mouse primary megakaryocytes (1° Meg.) (lane 5), C2C12 mouse myoblasts (lane 6), or 3T3-L1 mouse adipocytes (lane 7) were used. DNA-protein complexes A, B, and C are indicated on the side of the autoradiograph along with Sp1 (*) and Sp3 (**) containing complexes. B, nucleotide sequence of DNA fragments used to map the binding sites for complexes binding to E-I region. The wild-type sequence is shown on top and mutant sequences below. Right, summary of the EMSA analysis shown in C indicates the ability of each mutated sequence to bind complexes A, B, and C. DNA-protein complexes binding to the E-I region detected by EMSA. Nuclear extracts from L8057 cells were incubated with WT probe shown in B, either alone (lanes 1, 2, and 3) or with increasing amount of unlabeled competitors (open triangles, 20- or 50-fold molar excess) (lanes 2–11, 13–22, and 23–35). DNA binding activities are indicated as in A. D, summary of DNA binding sequences required for binding of Sp1, Sp3, and complexes A, B, and C to the E-I region. Black lines show bases required for DNA binding and dashed lines bases that affect but are not required for binding. E, WT or indicated mutated labeled probes were incubated with L8057 nuclear extracts, either alone (lanes 1, 4, 7, 10, and 13) or with antibodies directed against Sp1 (lanes 2, 5, 8, 11, and 14) or Sp3 (lanes 3, 6, 9, 12, and 15). DNA binding activities are indicated as in A.
DNA Binding Factor B Is Required for mHS-3.5 Megakaryocyte Specific Activity—To test whether the newly identified in vitro DNA binding activities were important for mHS-3.5 megakaryocyte- but not red cell-specific enhancer activity, further reporter constructs were tested that contained mutations corresponding to GK3, GK8, and GS12 made within the context of fragment A-I. These constructs were tested for enhancer activity in fetal liver cells from E13.5 transient transgenic mice embryos. On this occasion we tested for LacZ expression using FDG as substrate and flow cytometric detection of β-galactosidase activity. This quantitative analysis allows for detection of subtle changes in the number of LacZ-expressing cells and the expression level/cell in larger cell numbers. β-Galactosidase activity was monitored in Ter119 erythroid cells, CD61Mac1 megakaryocytes and, as control, CD61Mac1 neutrophils/macrophages. As shown in Fig. 5, we confirmed that the WT E-I region-containing construct (construct A-I in Fig. 2) is able to specifically direct LacZ expression to erythroid cells and megakaryocytes but not neutrophils/macrophages. Similarly, LacZ expression was detected from constructs containing mutations corresponding to GK3 and GK8 in fetal liver CD61Mac1 megakaryocytes. However, two points are worth noting. First, 3/11 transgenic embryos containing the GK3 mutation also exhibited LacZ expression in neutrophils/macrophages. This suggests that either transgenes bearing this mutation may be more susceptible to position effect or that the DNA binding activity A may repress transgene expression in lineages that do not normally express GATA1. Second, though transgenic embryos containing constructs with the GK8 mutation clearly demonstrated megakaryocyte LacZ expression, there was a suggestion that the percentage of LacZ-expressing megakaryocytes was lower, though this was not statistically significant.

Most importantly, transgenic embryos with the GS12 mutation did not exhibit LacZ expression in CD61Mac1 fetal liver megakaryocytes, even though LacZ expression in erythroid cells was maintained. A statistical analysis using a binomial test showed that the frequencies of LacZ expression in transgenics embryos (Fig. 5, column 1) for the four constructs were not statistically different in erythroid cells. In contrast, the frequency of embryos transgenic for GS12 expressing LacZ in megakaryocytes was significantly lower.

As the GS12 mutation does not bind activities B and C (in contrast to the GK8 oligonucleotide that fails to bind only DNA-binding activity C in vitro), and as mice with the GK8 mutation are able to direct megakaryocyte LacZ expression, this suggests that, at a minimum, DNA binding activity B or a combination of activities B and C, are important for megakaryocyte-specific enhancer activity of E-I within mHS-3.5. A critically important caveat is that though mutations in oligonucleotides GK3, GK8, and GS12 disrupt binding of activities A, B, and C in vitro, this may not reflect events in vivo.

An E-I-like Region Binding Activity B Is Evolutionary Conserved in the Fli1 Gene Promoter—We then asked if sequences within the E-I region and, in particular, nucleotides important for activity B binding, were present in regulatory regions of other megakaryocyte genes. We found an evolutionary conserved E-I-like region in the promoters of the mouse, rat, human, chimpanzee, dog, and opossum Fli1 genes (Fig. 6A). The Fli1 gene is expressed in megakaryocytes and is required for their development (24). Interestingly, the most important nucleotides for activity B binding in the mHS-3.5 E-I region (see Fig. 3D) are especially well conserved in the Fli1 promoter. We then tested if the E-I like region in the Fli1 promoter was able to bind activity B by EMSA using the WT mHS-3.5 E-I region as probe and the Fli1 E-I like regions of several species as
cold competitors. As shown in Fig. 6B, all the Fli1 E-I like regions tested were able to compete with the mHS-3.5 E-I region for activity B binding. In addition, all Fli1 E-I like regions were also able to compete for the binding of activity A but not of activity C. Unlike all other species, the E-I like region of the opossum Fli1 promoter was unable to compete for the Sp1/Sp3 binding. To confirm the binding of activity B, we used the various Fli1 E-I like regions as probes in EMSA experiments (Fig. 6C).

**DISCUSSION**

To begin to address the question of what combinations of transcriptional regulators are required for the distinct erythroid and megakaryocytic gene expression programs, we have focused on the transcriptional regulation of the key erythroid-megakaryocytic transcription factor GATA1. GATA1 is absolutely required for terminal maturation of both lineages (see Introduction). One model for how GATA1 expression could be regulated in these two lineages is that common mechanisms ensure expression in both lineages. However, this is unlikely as the level of GATA1 mRNA is 5-fold higher in primary wild-type red cells compared with megakaryocytes,4 and when red cell GATA1 expression is reduced to 20% of wild-type levels, there is a block in red cell maturation (25). Therefore, in common with other transcription factors (where heterozygous mutant mice have a mutant phenotype from reduced gene expression), the level of GATA1 expression is also likely to be carefully controlled to ensure proper lineage maturation. These postulated differences in regulation of megakaryocyte and red cell GATA1 expression are strengthened by our previous observations that there are different sequence requirements for red cell and megakaryocyte GATA1 expres-
A. E-I  H/M/D  TGGGGAGGGAAATGGG

M. musculus  GAGGGGTGTTGGGAGG
R. norvegicus  AGGGGGTGGGAGG
H. sapiens  GAGGGGTGGGAGG
P. troglodytes  GAGGGGTGGGAGG
C. familiaris  GAGGGGTGGGAGG
M. domestica  GAGGGGAAGGGGAGG

B. FIGURE 6. An E-I-like sequence is conserved throughout evolution in the Fli1 promoter. A, on top, is shown the conserved E-I sequence present in human (H) mouse (M), and dog (D). Cross-species alignment of the E-I-like sequence found in the Fli1 promoter of the mouse (M. musculus), rat (R. norvegicus), human (H. sapiens), chimpanzee (P. troglodytes), dog (C. familiaris), and opossum (M. domestica) is shown below. Position of nucleotides required for factor B binding in mHS-3.5 are shaded and the position of each E-I-like sequence respective to the transcription start site of the Fli1 genes is shown to the right of each sequence. B, mHS-3.5 E-I region was used as probe in EMSA using L8057 nuclear extracts. In the left panel, binding reaction was performed in the absence (−) or presence either of antibodies against Sp1 or Sp3 or of a 50-fold molar excess of the unlabeled WT, GK3, GK8, or GS12 competitors, as in Fig. 3C. In the right panel, 50- or 150-fold molar excess of the unlabeled Fli1 E-I-like region of the indicated species was used as competitor. DNA binding activities are indicated as in Fig. 3C. C, mHS-3.5 E-I region or the Fli1 E-I-like region from the indicated species were used as probes in EMSA using L8057 nuclear extracts. The binding reaction was performed in the absence (−) or presence of a 50-fold molar excess of the WT or GS12 competitors as in Fig. 3C. The binding activity competed by WT, but not GS12, is shown by a black circle.

At least two portions of mHS-3.5 are required for megakaryocyte GATA1 expression. First, previous experiments showed that a GATA site (see Fig. 2B) within mHS-3.5 was required for enhancer function (11, 12). In vivo, in primary megakaryocytes, this site can bind GATA1 (supplementary Fig. S1A), consistent with models that GATA1 posi-
Megakaryocyte-specific GATA1 Expression

In conclusion, we report the identification and characterization of a 25-bp element within the critical mGATA1 enhancer mHS-3.5 region required to direct megakaryocyte reporter gene expression in transgenic mice. In vitro this complex element binds to at least one DNA binding activity present in primary megakaryocytes but not red cells and mutations that impair binding of this activity cripple megakaryocyte mHS-3.5 function. Identification of the protein(s) that constitute this activity, B, will shed new light on the differential regulation of megakaryocyte versus erythroid gene expression and ultimately how these two lineages are differentially specified from a common progenitor.

Acknowledgments—We thank Professor Bill Wood and Jackie Sloan-Stanley for the megakaryocyte purification protocol, and Anne Azizberger and Chris Kuhl for help with flow cytometric protocols. We thank the Vyas and Porcher laboratories for helpful discussions. We thank Drs. de Bruijn and Professor Wood for reading the manuscript.

REFERENCES
1. Akashi, K., Traver, D., Miyamoto, T., and Weissman, I. L. (2000) Nature 404, 193–197
2. Iwasaki, H., Mizuno, S., Wells, R. A., Cantor, A. B., Watanabe, S., and Akashi, K. (2003) Immunity 19, 451–462
3. Valverde-Garduno, V., Guyot, B., Anguita, E., Hamlett, I., Porcher, C., and Vyas, P.

VOLUME 281 • NUMBER 19
MAY 12, 2006

Part of GATA1 function may be to establish a domain of chromatin enriched for acetylated histone H3 in and around the Gata1 gene, as is the case for the β-globin gene (27, 28). In addition, we also detected in vivo binding by the bHLH transcription factor SCL/TAL-1 (supplementary Fig. S1B) that could potentially bind an E-box adjacent to the GATA site consistent with the notion of SCL/TAL-1 regulating GATA1 expression. Though mutation of the E-box in mHS-3.5 does not abrogate reporter gene expression (11), SCL/TAL-1 can be recruited to cis elements in the absence of direct DNA binding (3, 29).

However, it is likely that GATA and SCL/TAL-1 binding is required for both red cell and megakaryocyte GATA1 expression as both regulators bind mHS-3.5 in both lineages and mutation of the GATA site renders mHS-3.5 inactive in transgenic mouse reporter gene assays in red cells and megakaryocytes.

In contrast, the second region in mHS-3.5 required for megakaryocyte function, but not red cell, mHS-3.5 enhancer function in transgenic mice is a 25 bp fragment, E-I. Within E-I the 5’ 17 bp are fully conserved through evolution from human to mouse.

In vitro mouse E-I binds Sp1 and Sp3 and three DNA binding activities A, B, and C. Activity A is likely to contain or may correspond entirely to the Kruppel-type zinc finger protein ZBP-89 (also known as zfp 148/BERF-1/BFCOL-1/mtb) that has been shown to function both as a transcriptional activator and repressor (22, 30–32). Though we have not directly tested the functional role of Sp1 and Sp3 in mediating mHS-3.5 enhancer activity, we failed to detect in vivo Sp1 or Sp3 binding to mHS-3.5 (despite clear binding of these factors at 23 kilobases away at the putative promoter of the gene neighboring mGata1, Hdac6).

In contrast, in vivo ZBP-89 binds not only mHS-3.5 but also mIE and mHS+3.5 in the L8057 megakaryoblastic cell line. However, results from mice transgenic for GK3 (that abolishes in vitro ZBP-89 binding) suggests that ZBP-89 is not a critical determinant of mHS-3.5 directed megakaryocyte expression though may potentially stop inappropriate GATA1 expression in non-GATA1 expressing cell types (see Fig. 5). The absence of a major role for ZBP-89 in directing megakaryocyte GATA1 expression is also indirectly supported by the low/absent ZBP-89 in vitro binding activity at the mouse E-I mHS-3.5 region in nuclear extracts from primary megakaryocytes (Fig. 2A). However, one caveat of this interpretation is that disruption of binding in vitro (by the GK3 mutation) may not disrupt in vivo binding. Moreover, as ZBP-89 binds to mIE and mHS+3.5 in vivo, in addition to mHS-3.5, one cannot exclude that ZBP-89 regulates GATA1 expression via any combination of in vivo binding sites. Finally, recent data suggest that in megakaryocytic cells ZBP-89 binds GATA1 (which autoregulates its own transcription, Ref. 26). Importantly, knock-down of ZBP-89 expression in zebrafish embryos ablates thrombocyte development and disruption of the ZBP-89 locus in mouse ES cells impairs megakaryocyte differentiation (23). It will be important to document megakaryocyte and red cell GATA1 expression in these mutant fish and mice and also ascertain if GATA1 is a direct transcriptional target of ZBP-89.

In contrast to A, the nature of activities B and C remain unknown. The discrepancy between in vitro and in vivo Sp1 and Sp3 binding data raise the specter that binding of activities B and C in the E-I region could also be in vitro artifacts. However, combined results from the combination of mutants GS12 (where in vitro binding of B and C is disrupted) and GK8 (in vitro binding of C alone is disrupted) suggest that either activity B alone, or activities B and C together, are important in mediating mHS-3.5 megakaryocyte-specific (but not red cell-specific) enhancer function in transgenic mice. This coupled with specific in vitro binding of complex B in primary megakaryocytes, but not primary red cells suggests activity B may be critical for megakaryocyte-specific mHS-3.5 enhancer function.

We then asked if sequences within the E-I region were present in regulatory regions of other megakaryocyte genes. Binding of A and B was detected at an evolutionarily conserved E-I like element located close to a GATA site, near the transcriptional start site of another megakaryocyte transcription factor gene Fli1 (19). This leads to the hypothesis that activity B may be epistatically upstream of GATA1 and Fli1, in the cascade of gene activation orchestrating megakaryocyte-specific gene expression.

The binding site for activity B is a bipartite G/A-rich sequence. Though a core ETS binding site consensus (GGGAA) is present within the activity B binding site, competition and transgenics experiments suggest that the ETS site is not required for activity B binding since mutation GK8, which disrupts the core ETS site, is still able to bind activity B (Fig. 3B) and doesn’t significantly affect LacZ expression in megakaryocytes (Fig. 5). Nonetheless, given the importance of the ETS family of transcription factors in megakaryocyte gene expression (24, 33–37), we tested if ETS family members were present in activity B by using specific antibodies against ETS1, ETS2, FLI1, PEA3, ELF1, ELK1, ERG1, ETF1, or PU.1 (data not shown) in EMSA supershift assays, but observed no immunoreactivity against the retarded bands A, B, or C.

We also looked for binding sites for the transcription factors Meis1 and RUNX1 (TGCAGG and TGTGGT, respectively) in E-I. Both regulators are required for megakaryocyte but not red cell maturation (38, 39) and have been shown to regulate megakaryocyte-specific genes (40, 41). However, we could not detect binding sites for these factors in the E-I region by EMSA and supershift assays.

Computer-based search of consensus binding sites showed that factor B site shows similarities to that of MZF1 (myeloid zinc finger 1), a zinc-finger transcription factor which binds two G/A-rich DNA sequences (AGTGAGGGA and CGGGnGAGGGGGAA) via two separate zinc-finger clusters within the protein (42). The core GGGGA sequence, which is most important for DNA binding, is present in region E-I. However, the second part of the B site, where the mutation in GS12 mutant is located, does not match well with the MZF1 consensus. Moreover, MZF1 is mainly expressed in immature myeloid cells and MZF1 mutant mice and embryos do not have an overt megakaryocyte phenotype (43).

In conclusion, we report the identification and characterization of a 25-bp element within the critical mGATA1 enhancer mHS-3.5 required to direct megakaryocyte reporter gene expression in transgenic mice. In vitro this complex element binds to at least one DNA binding activity present in primary megakaryocytes but not red cells and mutations that impair binding of this activity cripple megakaryocyte mHS-3.5 function. Identification of the protein(s) that constitute this activity, B, will shed light on the differential regulation of megakaryocyte versus erythroid gene expression and ultimately how these two lineages are differentially specified from a common progenitor.
Megakaryocyte-specific GATA1 Expression

(2004) Blood 104, 3106–3116
4. Fujisawa, Y., Browne, C. P., Cannizz, K., Goff, S. C., and Orkin, S. H. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 12355–12358
5. Shvidrasani, R. A., Fujisawa, Y., McDevitt, M. A., and Orkin, S. H. (1997) EMBO J. 16, 3965–3973
6. Takahashi, S., Komoto, T., Suwabe, N., Yoh, K., Nakajima, O., Nishimura, S., Kuroha, T., and Yamamoto, M. (1998) Blood 92, 434–442
7. Vyas, P., Ault, K., Jackson, C. W., and Orkin, S. H. (1999) Blood 93, 2867–2875
8. Ronchi, A., Hauf, C., Kershaw, E., Giardino, R., Ricciardi-Castagnoli, P., Cross, M., Ghysdael, J., and Ottolenghi, S. (1997) Cell. Biol. 121, 135–145
9. Vyas, P., Ault, K., Jackson, C. W., Orkin, S. H., and Shivdasani, R. A. (1999) Proc. Natl. Acad. Sci. U. S. A. 94, 4487–4492
10. McDevitt, M. A., Fujiwara, Y., Shvidrasani, R. A., and Orkin, S. H. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 7976–7981
11. Vyas, P., McDevitt, M. A., Cantor, A. B., Katz, S. G., Fujiwara, Y., and Orkin, S. H. (1999) Development 126, 2799–2811
12. Nishimura, S., Takahashi, S., Kuroha, T., Suwabe, N., Nagasawa, T., Trainor, C., and Yamamoto, M. (2000) Mol. Cell. Biol. 20, 713–723
13. Villeval, J. L., Cohen-Solal, K., Tulliez, M., Giraudier, S., Guichard, J., Burstein, S. A., Vyas, P., Ault, K., Jackson, C. W., Orkin, S. H., and Shivdasani, R. A. (1999) Blood 93, 289–298
14. Ishida, Y., Levin, J., Baker, G., Stenberg, P. E., Yamada, Y., Sasaki, H., and Orkin, S. H. (1999) Exp. Hematol. 27, 136–143
15. Yaffe, D., and Saxel, O. (1977) Nature 269, 523–526
16. Ishida, Y., Levin, J., Baker, G., Stenberg, P. E., Yamada, Y., Sasaki, H., and Inoue, T. (1993) Exp. Hematol. 21, 289–295
17. Yaffe, D., and Sazel, O. (1977) Nature 269, 523–526
18. Green, H., and Meuth, M. (1974) Cell 1, 127–133
19. Guyot, B., Valverde-Garduno, V., Porcher, C., and Vyas, P. (2004) Blood 104, 89–91
20. Fiering, S. N., Roederer, M., Nolan, G. P., Mclenn, D. R., Parks, D. R., and Herzenberg, L. A. (1991) Cytometry 12, 291–301
21. Starck, J., Doubekovski, A., Sarrazin, S., Gonnet, C., Rao, G., Skoulitchi, A., Godet, J., Duxanet-Fourt, I., and Morle, F. (1999) Mol. Cell. Biol. 19, 121–135
22. Kuhl, C., Atzberger, A., Iborra, Niewsland, B., Porcher, C., and Vyas, P. (2005) Mol. Cell. Biol. 25, 8592–8606
23. Suske, G. (1999) Gene (Amst.) 234, 291–300
24. Merchant, J. L., Iyer, G. R., Taylor, B. R., Kitchen, J. R., Mortensen, E. R., Wang, Z., Flintoff, R. J., Michel, J. B., and Bassel-Duby, R. (1996) Mol. Cell. Biol. 16, 6644–6653
25. Waa, A. J., Morin, T. B., Choe, S.-K., Schindler, Y. L., Sullivan, M. R., Fujiwara, Y., Paw, B., and Cantor, A. B. (2005) Blood 106, 828
26. Hart, A., Melet, F., Grossfeld, P., Chien, K., Jones, C., Tunnacchile, A., Favier, R., and Bernstein, A. (2000) Immunity 13, 167–177
27. McDevitt, M. A., Shvidrasani, R. A., Fujiwara, Y., Yang, H., and Orkin, S. H. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 6781–6785
28. Kobayashi, M., Nishikawa, K., and Yamamoto, M. (2001) Development 128, 2341–2350
29. Letting, D. L., Rakowski, C., Weiss, M. J., and Blobel, G. A. (2003) Mol. Cell. Biol. 23, 1334–1340
30. Kiekbaer, C. M., Grass, J. A., Johnson, K. D., Boyer, M. E., and Bresnick, E. H. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 14309–14134
31. Anguita, E., Hughes, J., Heyworth, C., Blobel, G. A., Woods, W. G., and Higgs, D. R. (2004) EMBO J. 23, 2841–2852
32. Yamada, A., Takaki, S., Hayashi, F., Georgopoulos, K., Perlmuter, R. M., and Takanaka, K. (2001) J. Biol. Chem. 276, 18082–18089
33. Park, H., Shelley, C. S., and Arnaout, M. A. (2003) Blood 101, 894–902
34. Boopathi, E., Lenka, N., Prabu, S. K., Fang, J. K., Wilkinson, F., Atchison, M., Giallongo, A., and Avadhanali, N. G. (2004) J. Biol. Chem. 279, 35242–35254
35. Lemarchandel, V., Ghysdael, J., Mignotte, V., Rahuel, C., and Romeo, P. H. (1993) Mol. Cell. Biol. 13, 668–676
36. Deveaux, S., Filipe, A., Lemarchandel, V., Ghysdael, J., Romeo, P. H., and Mignotte, V. (1996) Blood 87, 4678–4685
37. Wang, X., Crispino, J. D., Letting, D. L., Nakazawa, M., Poncz, M., and Blobel, G. A. (2002) EMBO J. 21, 5225–5234
38. Eisbacher, M., Holmes, M. L., Newton, A., Hogg, P. J., Khachigian, L. M., Crossley, M., and Chong, B. H. (2003) Mol. Cell. Biol. 23, 3427–3441
39. Jackers, P., Szalai, G., Moussa, O., and Watson, D. K. (2004) J. Biol. Chem. 279, 52183–52190
40. Hira, T., Spence, S. E., Rachel, R. A., Fujita, M., Nakamura, T., Ward, J. M., Devor-Henneman, D. E., Saiki, Y., Kutsuna, H., Tessarollo, L., Jenkins, N. A., and Copeland, N. G. (2004) EMBO J. 23, 450–459
41. Ichikawa, M., Asai, T., Satoh, T., Seo, S., Yamazaki, I., Yamagata, T., Mitani, K., Chiba, S., Ogawa, S., Kurokawa, M., and Hirai, H. (2004) Nat Med 10, 299–304
42. Okada, Y., Naga, R., Satoh, T., Matsuura, E., Minami, T., Morita, I., and Dui, T. (2003) Blood 101, 4748–4756
43. Elagib, K. E., Rake, F. K., Mogass, M., Khetawat, R., Delehanty, L. L., and Goldfarb, A. N. (2003) Blood 101, 4333–4341
44. Morris, J. F., Hromas, R., and Rauscher, F. J. 3rd. (1994) Mol. Cell. Biol. 14, 1786–1795
45. Gaboli, M., Kost, P. A., Gurrerri, C., Cattorette, G., Ronchetti, S., Cordon-Cardo, C., Bruxmeyer, H. E., Hromas, R., and Pandolfo, P. F. (2001) Genes Dev. 15, 1625–1630