cis elements that mediate transcription factor binding are abundant within genomes, but the rules governing occupancy of such motifs in chromatin are not understood. The transcription factor GATA-1 that regulates red blood cell development binds with high affinity to GATA motifs, and initial studies suggest that these motifs are often unavailable for occupancy in chromatin. Whereas GATA-2 regulates the differentiation of all blood cell lineages via GATA motif binding, the specificity of GATA-2 chromatin occupancy has not been studied. We found that conditionally active GATA-1 (ER-GATA-1) and GATA-2 occupy only a small subset of the conserved GATA motifs within the murine β-globin locus. Kinetic analyses in GATA-1-null cells indicated that ER-GATA-1 preferentially occupied GATA motifs at the locus control region (LCR), in which chromatin accessibility is largely GATA-1-independent. Subsequently, ER-GATA-1 increased promoter accessibility and occupied the β major promoter. ER-GATA-1 increased erythroid Krüppel-like factor and SWI/SNF chromatin remodeling complex occupancy at restricted LCR sites. These studies revealed three phases of β-globin locus activation: GATA-1-independent establishment of specific chromatin structure features, GATA-1-dependent LCR complex assembly, and GATA-1-dependent promoter complex assembly. The differential utilization of dispersed GATA motifs therefore establishes spatial/temporal regulation and underlies the multistep activation mechanism.

Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region


*Department of Pharmacology, University of Wisconsin Medical School, Madison, WI 53706; †Department of Cell Biology, University of Massachusetts Medical School, Worcester, MA 01655; and §Department of Molecular Cell and Developmental Biology, Mount Sinai School of Medicine, New York, NY 10029

Edited by Mark T. Groudine, Fred Hutchinson Cancer Research Center, Seattle, WA, and approved September 12, 2005 (received for review July 22, 2005)

Cis elements that mediate transcription factor binding are abundant within genomes, but the rules governing occupancy of such motifs in chromatin are not understood. The transcription factor GATA-1 that regulates red blood cell development binds with high affinity to GATA motifs, and initial studies suggest that these motifs are often unavailable for occupancy in chromatin. Whereas GATA-2 regulates the differentiation of all blood cell lineages via GATA motif binding, the specificity of GATA-2 chromatin occupancy has not been studied. We found that conditionally active GATA-1 (ER-GATA-1) and GATA-2 occupy only a small subset of the conserved GATA motifs within the murine β-globin locus. Kinetic analyses in GATA-1-null cells indicated that ER-GATA-1 preferentially occupied GATA motifs at the locus control region (LCR), in which chromatin accessibility is largely GATA-1-independent. Subsequently, ER-GATA-1 increased promoter accessibility and occupied the β major promoter. ER-GATA-1 increased erythroid Krüppel-like factor and SWI/SNF chromatin remodeling complex occupancy at restricted LCR sites. These studies revealed three phases of β-globin locus activation: GATA-1-independent establishment of specific chromatin structure features, GATA-1-dependent LCR complex assembly, and GATA-1-dependent promoter complex assembly. The differential utilization of dispersed GATA motifs therefore establishes spatial/temporal regulation and underlies the multistep activation mechanism.

by contrast to zinc fingers that solely mediate DNA binding, the GATA-1 zinc fingers are multifunctional. The GATA-1 C-terminal finger binds GATA motifs (A/T)GATA(A/G) [9, 10], and the N-terminal finger binds the nine zinc finger-containing protein Friend of GATA-1 (FOG-1) [11, 12]. The N finger also stabilizes DNA binding at certain GATA motifs [13–15]. GATA-1-mediated activation and repression can require FOG-1, although GATA-1 target genes can also be FOG-1 independent [12]. FOG-1 increases GATA-1 chromatin occupancy [16, 17] and is required for GATA switches in which GATA-1 displaces chromatin-bound GATA-2 [16]. Despite the multiple zinc fingers, intrinsic FOG-1 DNA binding activity has not been demonstrated. GATA-1 also binds the histone acetyltransferases CREB binding protein/p300 [18], which interact with a plethora of activators [19]. Elucidating mechanisms underlying GATA-1-mediated transcriptional regulation and identifying GATA-1 target genes should provide key insights into how GATA-1 regulates cell differentiation and proliferation.

Major progress has been made in defining the GATA-1-instigated genetic network. Studies of protein–DNA interactions at the β-globin locus led to the identification of the first GATA-1 target genes, the β-globin genes [20, 21]. Chromatin immunoprecipitation (ChIP) analyses showed that GATA-1 occupies β-globin [22] and α-globin loci [23] regulatory regions. Gene profiling studies in GATA-1-null cells (G1E), with or without a conditionally active estrogen receptor ligand binding domain fusion to GATA-1 (ER-GATA-1), revealed many positively and negatively regulated genes [24]. Translating microarray data into direct targets can be complex, because ER-GATA-1 activation rapidly blocks proliferation [25], which may indirectly affect transcription. This problem is further complicated by the abundance of GATA motifs in the genome.

Genomic DNA regions occupied by factors in cells can be identified by ChIP coupled with microarray (26). However, given the novelty of this technology, little is known about false-positive and false-negative rates. Bioinformatics analysis of cis-element conservation may predict functional cis elements in certain contexts (27). ChIP, without microarray, analysis requires knowledge of a prospective DNA binding region (28, 29). Because bioinformatics predictions yield prospective binding regions, a coupled ChIP-bioinformatics approach allows one to definitively assess factor occupancy. Herein, we used this approach to analyze ER-GATA-1 and GATA-2 chromatin occupancy at the murine β-globin locus.

The murine β-globin locus contains Eγ and βH1 genes, active during embryogenesis, and the adult βmajor and βminor genes [30]. Upstream of Eγ resides the LCR (31, 32), consisting of four erythroid-specific DNaseI hypersensitive sites (HSs) (HS1–4) [33, 34]. GATA-1 occupies HS1–4 and the βmajor promoter in adult erythroid cells [22], but occupancy has not been tested at many conserved GATA motifs of the locus. βmajor transcriptional activation is associated with higher-order chromatin rearrangements within the LCR [35] and looping to bring the LCR in proximity of βmajor [36–38]. Looping requires GATA-1 [38] and the erythroid Krüppel-like factor (EKL) [37] that binds certain CACCCT motifs [39–41]. Although EKL occupancy in cells has not been described, studies with an EKL-null cell line [42, 43], altered specificity EKL mutants [40], and targeted deletions [44–46] indicate that EKL regulates β-globin transcription. EKL binds CREB binding protein/p300 [47] and the
Brahma-related gene 1 (BRG1) component of the SWI/SNF chromatin remodeling complex (43, 48, 49) and appears to function via establishing active chromatin. Targeted deletion of EKLF abrogates DNaseI hypersensitivity at HS3 and the \( \beta \)-globin promoter (40), and EKLF mediates chromatin remodeling in vitro (48).

The studies described herein investigate the specificities of ER-GATA-1 and GATA-2 interactions with GATA motifs, the relationship between ER-GATA-1 concentration/activity and chromatin occupancy, and how ER-GATA-1 influences specific steps in \( \beta \)-globin locus activation. ER-GATA-1 and GATA-2 occupied only a small subset of conserved GATA motifs, and EKLF occupancy was unpredictable from the distribution of conserved CACCC motifs. We established the spatial/temporal regulation of ER-GATA-1-instigated molecular events, which revealed ER-GATA-1-independent and -dependent phases of \( \beta \)-globin locus activation.

Methods

Cell Culture. G1E cells stably expressing ER-GATA-1 (50), mouse erythroleukemia, and FOG-1-null cells were maintained as described in Supporting Methods, which is published as supporting information on the PNAS web site.

Quantitative ChIP Assay. Quantitative ChIP analysis was performed as described in ref. 29 and in Supporting Methods.

Quantitative RT-PCR Assay. cDNA was prepared from 1 \( \mu \)g of purified total RNA. RT-PCRs (20 \( \mu \)l) contained 2 \( \mu \)l of cDNA solution with the appropriate primers. Product was measured by SYBR green fluorescence. Relative expression levels were determined from a standard curve of serial dilutions of cDNA samples.

Restriction Endonuclease Accessibility Assay. Assays were performed as described in ref. 52.

Protein Analysis. Protein analysis was conducted as described in Supporting Methods.

Primers and Antibodies. Primers and antibodies are described in Supporting Methods (see also Tables 1 and 2, which are published as supporting information on the PNAS web site).

Results and Discussion

ER-GATA-1, GATA-2, and EKLF Occupy Only a Small Subset of Conserved DNA Motifs in Chromatin. The canonical GATA motif WGATAR, and WGATA and GATAR motifs that bind GATA factors with high-affinity in vitro (9, 10), are abundant in chromatin. As the rules of GATA factor chromatin occupancy are not understood (8), we used highly specific anti-GATA-1 and -GATA-2 antibodies (Fig. 8, which is published as supporting information on the PNAS web site) to analyze ER-GATA-1 and GATA-2 occupancy at 72 conserved (mouse vs. human) WGATAR, WGATA, and GATAR motifs.
motifs (within 64 amplicons) of the β-globin locus in tamoxifen-treated G1E-ER-GATA-1 cells (Fig. 1A), ER-GATA-1 and GATA-1 occupy chromatin similarly (22). The conservation of murine β-globin locus sequence, relative to that of human and dog, and GATA motifs are shown in Fig. 1A. ER-GATA-1 occupancy was detected at nine GATA motif-containing amplicons (Fig. 1B). Analysis of the Eγ, βH1, βmaj, and βmin promoters lacking conserved GATA motifs revealed occupancy only at βmaj, which contains one nonconserved WGATAR motif. ER-GATA-1 occupancy was not detected at HS5, which contains two conserved WGATAR sequences. Similar results were obtained by using different G1E-ER-GATA-1 cell clonal lines expressing ER-GATA-1 at a level equivalent to or lower than mouse erythroleukemia cell GATA-1 (data not shown). Thus, ER-GATA-1 occupies a small subset of the conserved GATA motifs, and motif conservation and hypersensitivity are insufficient to predict occupancy.

ER-GATA-1 expression in G1E cells represses GATA-2 transcription (50, 53), GATA-1 and GATA-2 bind similar GATA motifs in vitro, although GATA-2 binds AGATCTTA with higher affinity than GATA-1 (10). To compare how ER-GATA-1 and GATA-2 discriminate among GATA motifs, GATA-2 occupancy was measured in untreated G1E-ER-GATA-1 cells (Fig. 1C). No qualitative differences between GATA-2 and ER-GATA-1 occupancy were detected, but signals at HS2, relative to HS1, HS3, and HS4, were ∼2-fold lower for GATA-2 vs. ER-GATA-1.

Activation of the β-globin genes requires multiple factors including GATA-1, p54 subunit of nuclear factor erythroid-2 (p54/NFE2), and EKLF. Because EKLF binds certain CACCC motifs in vitro (41), EKLF occupancy was measured at all 17 β-globin locus regions containing conserved CACCC motifs (Fig. 1A) and five functionally important regions lacking conserved CACCC motifs. Occupancy was detected at HS2, HS3, and the βmaj promoter and weakly at HS1 (Fig. 1D). No occupancy was detected at 18 regions, indicating that EKLF occupies a small subset of the motifs. Because GATA-1 and EKLF interact (54, 55), these factors might cooccupy or bind sites independently. The ER-GATA-1 and EKLF occupancy patterns differ in that little or no EKLF occupancy occurred at HS4 (Fig. 1D); ER-GATA-1 occupancy at HS4 was high. EKLF occupancy at HS1 was less than HS2 and HS3, whereas ER-GATA-1 occupancy at HS1 was equivalent to or higher than at other HSs. These results indicate that ER-GATA-1 and EKLF cooccupy HS2, HS3, and the promoter, but not all sites.

Although ER-GATA-1 Occupies the LCR and the βmaj Promoter in Definitive Erythroid Cells, ER-GATA-1 Preferentially Occupies the LCR When ER-GATA-1 Is Limiting. All WGATAR motifs and many derivatives thereof bind GATA factors with high-affinity in vitro (9, 10). Thus, ER-GATA-1 and GATA-2 occupancy of a small subset of conserved GATA motifs is unrelated to naked DNA binding affinities. To determine whether sites occupied by GATA factors in cells are differentially accessible to ER-GATA-1, we asked whether ER-GATA-1 occupies different sites with identical or distinct kinetics. Tamoxifen treatment of G1E-ER-GATA-1 cells for up to 20 h increased ER-GATA-1 levels/activity (Fig. 2 A and B). Low-level βmaj mRNA and primary transcripts were detected by 8 h after tamoxifen treatment. The transcripts were half-maximal by 14 h and further increased by 20 h (Fig. 2B). ER-GATA-1 occupancy at HS2, HS3, and HS4 was rapid, whereas promoter occupancy required 14 h (Fig. 2C–E) by using either 0.4% or 1% formaldehyde (Fig. 2E). G1E-ER-GATA-1 cells were also treated with increasing concentrations of tamoxifen for 12 h. Although ER-GATA-1 occupied HS4 and HS3, almost no promoter occupancy was detected (Fig. 9, which is published as supporting information on the PNAS web site). Analysis of GATA-1 binding to oligonucleotides containing GATA motifs from HS2 or the βmaj promoter by EMSA, by using a range of protein concentrations, revealed indistinguishable high-affinity binding in both cases, confirming that the βmaj promoter GATA motif mediates high-affinity binding (data not shown). The finding that ER-GATA-1 occupancy of the LCR vs. the promoter can be segregated
provides a strategy to elucidate spatial/temporal relationships among reactions that activate the β-globin locus.

GATA-1 Mediates Spatially Restricted EKLF and BRG1 Recruitment.
Because ER-GATA-1 elevates FOG-1 and EKLF mRNA (24), delayed ER-GATA-1 occupancy of the promoter might require the induction of these components. Cycloheximide cannot be used to test whether ER-GATA-1 occupancy at the promoter requires protein synthesis, because promoter occupancy requires ~14 h, cycloheximide elicited toxicity after several hours, and factors such as GATA-2 with short half-lives do not persist upon cycloheximide treatment (data not shown). RT-PCR analysis was conducted to determine whether FOG-1 and EKLF mRNA increases under conditions in which ER-GATA-1 differentially occupies the LCR and promoter. FOG-1 mRNA was constant (data not shown), but EKLF mRNA (Fig. 3A) and protein (Fig. 3B and C) increased by 8 h after tamoxifen treatment. ER-GATA-1 is not absolutely required for EKLF synthesis, because cells lacking ER-GATA-1 activity express EKLF (Fig. 3B and C).

Because ER-GATA-1 elevates EKLF before ER-GATA-1 occupancy at the β major promoter (Figs. 2E and 3B), and EKLF occupies the promoter upon ER-GATA-1 activation for 20 h (Fig.
EKLF might be required for ER-GATA-1 to occupy the promoter. Alternatively, EKLF might occupy the promoter after ER-GATA-1. EKLF resembled ER-GATA-1 in rapidly occupying HS2 and HS3 and subsequently occupying the promoter (Fig. 3D); no EKLF occupancy was detected at the active promoter (Fig. 11A) and EKLF occupancy levels correlated with the inactive necdin promoter. The ER-GATA-1 and EKLF occupancy at HS1 and HS2 were constant during the tamoxifen treatment. In untreated G1E-ER-GATA-1 cells, BRG1 occupancy was detected at HS4, HS3, HS2, βmajo1 promoter, and βmajo1 intron 2 (β12) (Fig. 4A). Thus, ER-GATA-1 activity and maximal EKLF levels are not required to recruit BRG1 to the locus. However, ER-GATA-1 activation for 20 h increased BRG1 occupancy at HS3 (Fig. 4A), which progressively increased upon ER-GATA-1 activation (Fig. 4B). Because such binding of BRG1 to the βmajo1 promoter slightly increased at 20 h (Fig. 4B).

The hematopoietic factor NF-E2 also occupies the LCR and, to a lesser extent, the βmajo1 promoter (51, 59–61). p45/NF-E2 occupies HS2 in an ER-GATA-1-independent manner, whereas occupancy at HS1 and HS3 is ER-GATA-1-dependent (22). We asked whether ER-GATA-1-dependent p45/NF-E2 occupancy occurs concomitant with ER-GATA-1 occupancy at the LCR, at the promoter, or during the intervening time. p45/NF-E2 levels and occupancy at HS2 were constant during the tamoxifen treatment (Fig. 11A and B, which is published as supporting information on the PNAS web site). Occupancy increased more rapidly at HS3 vs. the promoter (Fig. 11B). Because ER-GATA-1 enhances Pol II recruitment to the LCR and promoter (22, 62) and p45/NF-E2 enhances Pol II recruitment to the promoter (51), we tested whether these spatially distinct events have shared or distinct kinetics. EKLF, p45/NF-E2, and Pol II occupied the LCR before the promoter, similar to ER-GATA-1 (Fig. 9C).

**GATA-1-Dependent and -Independent Components of β-Globin Locus Chromatin Architecture.** ER-GATA-1 increases histone acetylation at the LCR and the βmajo1 promoter (63, 64), but β-globin locus chromatin structure in cells, with or without GATA-1, has not been studied. Differences in GATA motif accessibility in chromatin might underlie the early and late phases of LCR and promoter occupancy, respectively. In this regard, GATA motifs are required for human LCR fragments to form HSs in transgenic mice (65). We used a restriction endonuclease accessibility assay (52, 66) to determine whether LCR and βmajo1 promoter accessibility is ER-GATA-1 dependent. Accessibility was measured in nuclei from G1E-ER-GATA-1 cells treated with tamoxifen for increasing times (Fig. 5A). MboI and HaeIII sites at HS4 and HS2, respectively, were cleaved on a high percentage of templates equivalently in nuclei from control and tamoxifen-treated cells (Fig. 5 C and D). By contrast, HaeIII cleavage at HS3 and the βmajo1 promoter was increased by ER-GATA-1 (Fig. 5 B and C). These results suggest that a promoter chromatin impediment underlies the delayed ER-GATA-1, EKLF, p45/NF-E2, and Pol II occupancy kinetics at the promoter, and this impediment must be overcome to assemble the promoter complex. However, ER-GATA-1 rapidly occupies HS3, despite the fact that HS3 accessibility is enhanced by ER-GATA-1. The HS3 scenario might be analogous to GATA-4 occupancy of a reconstituted, condensed chromatin template, which induced chromatin unfolding (67).

Based on the role of histone acetylation in increasing chromatin accessibility (68), insufficient acetylation might cause the factor access blockade, or other epigenetic marks mediating active chromatin structure in cells, with or without GATA-1, has not been studied. Differences in GATA motif accessibility in chromatin might underlie the early and late phases of LCR and promoter occupancy, respectively. In this regard, GATA motifs are required for human LCR fragments to form HSs in transgenic mice (65). We used a restriction endonuclease accessibility assay (52, 66) to determine whether LCR and βmajo1 promoter accessibility is ER-GATA-1 dependent. Accessibility was measured in nuclei from G1E-ER-GATA-1 cells treated with tamoxifen for increasing times (Fig. 5A). MboI and HaeIII sites at HS4 and HS2, respectively, were cleaved on a high percentage of templates equivalently in nuclei from control and tamoxifen-treated cells (Fig. 5 C and D). By contrast, HaeIII cleavage at HS3 and the βmajo1 promoter was increased by ER-GATA-1 (Fig. 5 B and C). These results suggest that a promoter chromatin impediment underlies the delayed ER-GATA-1, EKLF, p45/NF-E2, and Pol II occupancy kinetics at the promoter, and this impediment must be overcome to assemble the promoter complex. However, ER-GATA-1 rapidly occupies HS3, despite the fact that HS3 accessibility is enhanced by ER-GATA-1. The HS3 scenario might be analogous to GATA-4 occupancy of a reconstituted, condensed chromatin template, which induced chromatin unfolding (67).
lysine 79 (H3-meK79), which has been implicated in counteracting GATA-1 activity, ER-GATA-1-independent molecular events consequent to factor occupancy at the promoter therefore correlates with increased H3-meK79 at the promoter and the ORF. The impedi-

ment to factor occupancy at the promoter therefore correlates with reduced HaeIII accessibility, acH3, and H3-meK79.

Multistep Mechanism of β-globin Locus Transcriptional Activation.

The preferential occupancy at the LCR and the differential kinetics of LCR vs. promoter occupancy have important implications for β-globin locus activation. It is attractive to propose that a GATA-1 level sufficient to occupy the LCR is achieved during hematopoiesis, before occupancy of both the LCR and the promoter. Hyper-sensitivity and elevated histone acetylation at HS2 also precede the third phase of activation. Because targeted deletion of the LCR does not abrogate p45/NF-E2 (60) and GATA-1 (38) occupancy at the promoter, phase 2 is not required for phase 3. This multistep mechanism exemplifies the complexities of how cell type-specific activators instigate diverse reactions with stringent spatial/temporal control, thereby activating a chromatin domain. Further dissection of the intrachromosomal reactions and the interphase relationships, in the context of the three-dimensional nuclear milieu, is expected will yield a comprehensive understanding of how cellular factors function combinatorially to regulate hemoglobin synthesis.

We thank Soumen Pal for critical comments. We acknowledge support from National Institutes of Health (NIH) Grants DK50107 (to E.H.B.), DK55700 (to E.H.B.), and DK46665 (to J.B.H.). H.I. is a Predoctoral Fellow of the American Heart Association. J.M. acknowledges prior support from NIH Grant R01GM6488 and current support from NIH Grant T32HL07999.