Embryonic stem (ES)\textsuperscript{4} cells pose specific challenges to chromatin structure and regulation of gene expression. The propensity of self-renewal demands strict maintenance of heritable patterns of expression, whereas pluripotency implies plasticity and tolerance of regulatory decisions that dictate cell fate. The temporal balance between repression of genes, in stem cells, and their expression, during differentiation and tissue specificity, is struck in large part by epigenetic means. Globally, the chromatin of ES cells appears less condensed; repressive histone modifications, e.g. histone H3 methylated within its amino terminus at lysine 9 (H3K9me) and deacetylated histones H3 and H4, increase in parallel with differentiation (for examples, see Refs. 1–3). However, polycomb repressor complexes play an important role in stem cell maintenance by imposing repressive histone modifications such as methylated H3K27 (H3K27me) (4–6). Specific genes in ES cells lie silent, but poised for de novo expression during cell specification, within bivalent, chromatin domains of histone modifications, associated with both active chromatin and repressed chromatin, e.g. H3K4me and H3K27me, respectively (7).

Methylation of CpG sequences effectively silences gene expression in a heritable manner (8), and differentiation of ES cells is marked by regulated changes in DNA methylation (9, 10). Bisulfite sequencing of DNA methylation in ES cells revealed “windows” of hypomethylated CpG dinucleotides at regulatory regions, such as enhancers of genes expressed in differentiated cells (11, 12). This mechanism may offer access to DNA-binding proteins that are associated with primary induction of chromatin activation. In ES cells, Forkhead factor Foxd3, considered a self-renewal and pluripotency regulator, binds the albumin enhancer to initiate a progression of transcription factor binding and activation of Alb expression (11, 12). In committed endoderm, forkhead factor Foxa1 replaces Foxd3 at the albumin enhancer to initiate a progression of transcription factor binding and activation of Alb expression. Foxa1 is a “pioneer transcription factor” during embryonic development and hepatic specification, due to an ability of Foxa1 to engage condensed chromatin, bind nucleosome-assembled Foxa1 regulatory elements, and displace repressive linker histones (13–16).

Epigenetic control of genes that are silent in embryonic stem cells, but destined for expression during differentiation, includes distinctive hallmarks, such as simultaneous activating/repressing (bivalent) modifications of chromatin and DNA hypomethylation at enhancers of gene expression. Although α-fetoprotein (Afp) falls into this class of genes, as it is silent in pluripotent stem cells and activated during differentiation of endoderm, we find that Afp chromatin lacks bivalent histone modifications. However, critical regulatory sites for Afp activation, overlapping Foxa1/p53/Smad-binding elements, are located within a 300-bp region lacking DNA methylation, due to transposed elements underrepresented in CpG sequences: a short interspersed transposable element and a medium reiterated sequence element. Forkhead family member Foxa1 is activated by retinoic acid treatment of embryonic stem cells, binds its DNA consensus site within the short interspersed transposable/medium reiterated sequence elements, and displaces linker histone H1 from silent Afp chromatin. Small interfering RNA depletion of Foxa1 showed that Foxa1 is essential in providing chromatin access to transforming growth factor β-activated Smad2 and Smad4 and their subsequent DNA binding. Together these transcription factors establish highly acetylated chromatin and promote expression of Afp. Foxa1 acts as a pioneer transcription factor in de novo activation of Afp, by exploiting a lack of methylation at juxtaposed transposed elements, to bind and poise chromatin for intersection with transforming growth factor β signaling during differentiation of embryonic stem cells.
Molecular mechanisms of de novo gene activation, during differentiation of ES cells, are not well defined, and clearly may differ from tissue-specific gene regulation in somatic cells (11). Bivalency of chromatin and methylation-free windows across the genome do not account for control of all silent genes, which are expressed during differentiation of stem cells. Among these genes is α-fetoprotein (Afp), which is widely used as a marker of stem cell differentiation and endoderm lineage specification (17–19). During embryonic development, Afp is expressed concomitant with hepatic determination at the presumptive ventral foregut (20–22). Previously, we found that Foxa1 is an important, hepatic-expressed activator of Afp (23, 24). Robust expression of Afp shortly after birth occurs at highly acetylated chromatin, bound by Foxa1 at an intercalated Foxa1/p53/Smad binding element (SBE) within the distal promoter of Afp (centered at −850) (25). Expression of p53 in the liver increases after birth, and p53 displaces Foxa1 at the Foxa1/p53/SBE, due to its higher DNA binding affinity for the overlapping binding sites (23). Whether similar strategies are used to maintain repression or initiate activation of Afp in stem cells versus hepatic-specified cells, is unknown.

Addition of retinoic acid (RA), which binds retinoic acid receptor (RAR), activates transcription of a number of genes and initiates in vitro differentiation of ES cells (27). We used this model of stem cell differentiation to determine the mechanisms of de novo activation of Afp. We found that, prior to RA-mediated activation, the distal promoter of silent Afp is nucleosome occupied but lacks the characteristics of bivalent chromatin, as well as repressive methylation of H3K9. Addition of RA mediates rapid and robust expression of Foxa1 (28), which is required for activation of Afp during stem cell differentiation. Foxa1 binds within a DNA regulatory region of Afp lacking DNA methylation; however, the methylation-free window of activator sites within Afp is not maintained by stem cell factor Foxd3. Rather, we find that the methylation-free window of Afp spans juxtaposed, CpG-underrepresented DNA transposable elements (TE): a short interspersed transposable element (SINE) and a medium reiterated sequence 1 (MER1) element (29). RA-induced Foxa1 exploits the lack of methylation at SINE/MER1 transposable elements and functions as a pioneer transcription factor: displacing linker histone H1 and altering repressed chromatin. By initiating chromatin accessibility, Foxa1 intersects with TGF-β-activated Smad proteins to activate silent chromatin and express Afp during differentiation of stem cells.

**EXPERIMENTAL PROCEDURES**

**ES Cell Culture Conditions**—Mouse wild type AB-1, a gift of R. Johnson (30), and Foxa1 null (31) embryonic stem cells were grown on plastic dishes (TRP) coated with 0.1% gelatin (Sigma). The culture medium was Dulbecco’s modified essential medium (HyClone, Logan, UT) supplemented with 20% fetal bovine serum (Gemini Bioproducts, West Sacramento, CA), 1% antibiotics (Mediatech Inc., Manassas, VA), a final concentration of 6 mM glutamine (Mediatech Inc.), 0.1 mM β-mercaptoethanol (Sigma), and 16.7 ng/ml of recombinant LIF (32). Cells were differentiated by culturing with the above medium without β-mercaptoethanol or LIF and with the addition of 1 μM retinoic acid (Sigma).
Foxa1 and Smad Proteins Activate Afp in ES Cells

Results

Retinoic Acid Induces Expression of Multiple Definitive Endoderm-enriched Genes in Mouse ES Cells—To investigate the regulation of Afp activation in ES cells, we used LIF withdrawal and addition of RA to differentiate mouse embryonic stem cells and induce definitive endoderm-directed gene expression (37, 38). Addition of RA to mouse ES cells, combined with withdrawal of LIF, caused progressive activation of the definitive endoderm-enriched genes Foxa1, Foxa2, HNF6, and Afp (Fig. 1A) (39, 40), decreased staining of the pluripotency marker alkaline phosphatase, and decreased expression of the pluripotency markers Foxd3, Ronin, Sox2, and Nanog (supplemental Fig. S1). Activated expression of the definitive endoderm-enriched gene Foxa1 begins within 24 h after RA treatment and peaks by 2 days, whereas other endoderm-enriched genes are expressed only after 2 and 3 days of treatment. This temporal pattern of RA-induced expression is consistent with a potential role for Foxa1 as an initiator of active chromatin structure and expression of endoderm-enriched genes.

Chromatin in the Afp Distal Promoter Is Not Bivalent or Methylated at H3K9—Afp is silent during mouse blastula formation and is later expressed in the gut endoderm and yolk sac beginning at day 7.5 of embryonic development (E7.5) (41). This developmental expression pattern is similar to genes that have bivalent chromatin modifications in pluripotent stem cells: H3K4 methylation, associated with active expression, and H3K27 methylation, associated with repression of expression (7). To determine whether Afp exhibits this combination of histone modifications at the distal promoter, we performed ChIP for H3K27me2 (Fig. 1B) and H3K4me3 (Fig. 1C). Consistent with genome-wide sequencing of ChIP-enriched DNA (ChIP-Seq) from mouse ES cells (42), we did not detect significant amounts of either H3K4me3 or H3K27me2 at the Afp distal promoter ($p > 0.3$) prior to RA treatment (Fig. 1, B and C).

Western Blotting—Primary antibodies for the following proteins were incubated overnight at 4 °C and used at the specified dilutions: H3K9ac (Upstate 06-942) 1:1000, H3K4me2 (Upstate 07-030) 1:2000, H3K4me3 (Novus NB 500-173) 1:2000, H3K9me2 (Abcam ab7312, Upstate 07-212) 1:1000, H3 (Abcam ab32600, Upstate 07-030) 1:2000, H3K4me3 (Novus NB 500-173) 1:2000, H3K4me2 (Upstate DPC4) 1:500, β-actin (Genetex GTX30632, San Antonio, TX) 1:5000, Smad4 (Upstate DPC4) 1:500, Smad2 (Zymed Laboratories Inc. 51-1300) 1:2000, and phosphorylated Smad2 (P-Smad2) (Chemicon AB3849) 1:1000.

Co-immunoprecipitation—ES cells were lysed in NTEP buffer (150 mM NaCl, 25 mM Tris–HCl, pH 7.5, 5 mM EDTA, 0.5% Nonidet P-40, plus freshly added 1X protease inhibitor mixture Set I (Calbiochem) with 1 mM phenylmethylsulfonyl fluoride and phosphatase inhibitors, 100 mM NaF (Sigma) and 50 mM sodium vanadate (Sigma), were sonicated for a few seconds. The cell lysate was precleared with incubation with 20 μl of protein-A beads (50% slurry, Sigma) for 1 h. The precleared lysate was incubated overnight with antibodies: anti-Smad4 (DPC4, Upstate) and anti-rabbit IgG (Upstate). Next, 25 μl of protein A beads (50% slurry) were added and incubated for 2 h at 4 °C. Protein-bound beads were recovered by centrifugation and washed three times with NTEP plus 500 mM NaCl, then plus 0.5% SDS, then plus 500 mM NaCl, followed by NTEP buffer. Input lysate, equivalent to one-tenth of the immunoprecipitation lysate, was analyzed, alongside bead-bound proteins, by SDS-polyacrylamide gel electrophoresis and immunoblot analyses, as previously described (35). The primary antibodies for immunoblotting were as follows: mouse anti-Smad4 (Santa Cruz), mouse anti-Foxa1 (7 Hills Bioreagents, Cincinnati, OH), and rabbit anti-phospho-Smad2 (Cell Signaling, 3101).

Statistical Analyses—Graphing and calculations of means, standard deviations, and p value determinations were conducted using GraphPad Prism 5 software.

FIGURE 1. ES cells express Afp when treated with RA and chromatin at the distal promoter is not methylated at H3K4, Lys-9, or Lys-27. A, RT-PCR was performed on cells treated with RA for 0, 1, 2, 3, or 4 days. $C_t$ (threshold cycle) values for each of the indicated genes were normalized to 18S. * indicates significant difference from 0 days of RA treatment. B–D, chromatin immunoprecipitation was performed on ES cells maintained in LIF or treated with RA for 4 days. DNA from immunoprecipitations for H3K27me2 (B), H3K4me3 (C), and H3K9me2 (D) was analyzed by real time PCR for amplification of the Afp distal promoter with the Foxa3 and Trappc9 promoters as positive controls. Levels of each modification are expressed as a ratio to levels of histone H3, determined by a separate immunoprecipitation. The y axis in D is split to better indicate the low values in Afp samples. Error bars represent S.D. of at least three repetitions.
As a positive control, these modifications were detected at the Hoxa3 gene, which is known to have bivalent chromatin in ES cells and increased H3K4me3 and decreased H3K27me2 in differentiated cells (7). Although H3K4me3 was not detected at the distal promoter of Afp in undifferentiated ES cells, levels of this modification increased with RA treatment, consistent with Afp activation (Fig. 1C). We found no significant levels of H3K9me2 at Afp, before or after RA induction of differentiation; this histone modification is associated with repressed chromatin and gene silencing (43, 44), as at Trappc9 (38) (Fig. 1D).

Repetitive, Transposable Elements Establish a Region Lacking DNA Methylation—Afp chromatin is silent in pluripotent ES cells, and exhibited neither repressive H3K27me2 nor H3K9me2 at the distal promoter. We next asked if DNA methylation is present across upstream regulatory regions of the gene, potentially preventing critical activators of transcription from binding to DNA. Bisulfite sequencing analysis revealed that moderate to high levels of DNA methylation of CpG dinucleotides are present 230 bp upstream of the Foxa1-binding site (Fig. 2A). Downstream of the intercalated Foxa1/p53/SBE, more proximal to the start site of transcription, lies extensive methylation of DNA. In both cases, DNA methylation did not change significantly with RA-mediated differentiation. Strikingly, CpG sequences are absent from a 300-bp region of DNA spanning the intercalated Foxa1/p53/SBE (Fig. 2, A and B). Two repetitive TE are juxtaposed within this region: a B2-type SINE element upstream of a MER1 element, which covers the Foxa1/p53/SBE sites (Fig. 2B).

Work by Smale and colleagues (11) shows that the forkhead box protein and stem cell factor, Foxd3, binds within a window of hypomethylation that maps to a region encompassing a forkhead-binding site within the distal enhancer of the Alb gene. Foxd3 binding, prior to Alb expression, blocks DNA methylation and poises chromatin for Foxa1 interactions and activation of Alb, during differentiation. The Alb gene is highly similar to Afp, as they arose during evolution from a single gene, although their patterns of tissue-specific regulation diverged (45, 46). To
Transient transfection of siRNA oligos into mouse ES cells reduced Foxa1 and Smad4 mRNA and protein (Fig. 3A, S2), measured 4 days after beginning RA treatment. A 2-fold level of siRNA-mediated reduction of either Foxa1 or Smad4 mRNAs caused an ~2-fold decrease in transcription of Afp, compared with control scrambled siRNA. Measurements of RNA and protein levels in siRNA-treated and control cells indicate the specificity of siRNA depletion (Fig. 3A and supplemental Fig. S2). A critical role for Foxa1 in activating stem cell expression of Afp is illustrated by a comparison of Afp activation in RA-treated Foxa1-/- ES cells (31) compared with wild type cells. Foxa1-/- ES cells do not express Foxa1 and showed severely reduced RA-activated expression of Afp at 4 days of differentiation (Fig. 3B), when Afp is readily detectable in wild type ES cells (Figs. 1A and 3B).

To further assess the role of TGF-β signaling in activation of Afp, we used a TGF-β receptor inhibitor, SB-431542, which inhibits activin type 1 receptor ALK4, TGF-β type 1 receptor ALK5, and nodal type 1 receptor ALK7 proteins (55), all of which signal through Smad2 phosphorylated at Ser-465 and Ser-467 (P-Smad2) (56). Addition of SB-431542 to ES cells, treated with RA, decreases P-Smad2 to undetectable levels (Fig. 3C) and reduces expression of Afp, while not affecting expression of Foxa1, Smad2, or Smad4 (Fig. 3D). Thus, Foxa1, P-Smad2, and Smad4 are essential activators of expression of Afp during stem cell differentiation.

Foa1 and Smad Proteins Bind Chromatin during Activation of Afp—To determine whether Foxa1 and Smad2/4 directly activate Afp in differentiating ES cells, we performed ChIP and assessed binding of these proteins at Foxa1/p53/SBE in mouse ES cells, similar to previous studies in mouse liver (25). Here, we show that Foxa1 binds to the overlapping Foxa1/p53/SBE in the distal promoter of the Afp gene after 4 days of RA treatment but not in LIF-maintained ES cells (Fig. 4A). As a negative control, the p53 response element of the Nanog gene (57), which lacks an overlapping Foxa1 site, did not bind Foxa1.

Smad transcription factors act downstream of TGF-β signaling by binding DNA, as a complex between Smad4 and a phosphorylated R-Smad, e.g. P-Smad2 (56). ChIP experiments using antibodies against P-Smad2, as well as Smad4, indicated that a significant increase in Smad binding occurs when ES cells are treated with RA (Fig. 4, B and C), concomitant with Foxa1 binding to chromatin. Treatment of ES cells with RA could potentially increase TGF-β activity, as observed in human keratinocytes and leukemia cells (58–60), and in turn lead to elevated

determine whether Foxd3 binds the forkhead-binding site within the region flanked by CpG methylation at the Afp distal promoter, we performed ChIP for Foxd3 in LIF and RA-treated ES cells. Foxd3 binding was not detected at Afp, whereas it was detected at Alb (Fig. 2C). Lack of DNA methylation at the Afp regulatory sites, where Foxa1 binds in differentiating liver (23, 24), likely requires no prior “placeholder” binding of stem cell factor Foxd3, due to the presence of CpG-underrepresented, repetitive elements.

Foa1 and Smad4Activate Expression of Afp in Differentiating ES Cells—Development repression of Afp, in postnatal liver and in vitro, is imposed by p53 and TGF-β1-regulated transcription factors (23, 25, 35, 36, 47–49). P-Smad2 and Smad4 proteins, transcription effectors activated by TGF-β signaling, are bound to the Foxa1/p53/SBE (Fig. 2B) of postnatal repressed Afp, concomitant with an autoregulatory, co-repressor of TGF-β-mediated regulation, Snod (35). Whether TGF-β signaling and Smad proteins, in the absence of the co-repressor Snod, are involved in activation of Afp has never been determined. The juxtaposition of the SBE- and Foxa-binding sites at ~850 of Afp, as well as positive correlation between DNA binding of forkhead and Smad proteins in TGF-β-mediated activation of multiple genes (50–54), led us to examine the potential roles of Foxa1 and Smad proteins in de novo activation of Afp during stem cell differentiation.

For this purpose, we used siRNA oligos to deplete Foxa1 and Smad4 during RA-mediated differentiation of mouse ES cells.

**FIGURE 3.** Foxa1, Smad4, and TGF-β signaling are required for activation of Afp expression. A, RT-PCR was performed on RNA extracted from RA-treated ES cells transfected with Foxa1-targeted siRNA oligos, Smad4-targeted siRNA oligos, or scrambled siRNA oligos. C, values from amplification of the indicated genes were normalized to 18S and graphed relative to the non-target siRNA. Error bars represent standard deviation from at least three repetitions. B, RT-PCR was performed on RNA extracted from wild type and RA-treated ES cells treated with RA for 0 or 4 days. C, values from amplification of the indicated genes were normalized to 18S and graphed relative to the 0 day wild type sample. C, Western blots for P-Smad2 and β-actin were performed on protein lysates from ES cells treated with RA ± SB431542 for 4 days. D, RT-PCR was performed on RNA extracted from ES cells treated with RA ± SB431542 for 4 days. Expression of the indicated genes was normalized to 18S and graphed relative to the RA-treated sample.
phosphorylation of Smad2 and increased binding of P-Smad2/Smad4 to the Afp distal promoter. RA treatment transiently increases Smad2 RNA levels over 2 days of treatment but, by 4 days of RA, levels of Smad2 RNA and protein are equivalent to untreated ES cells (Fig. 4D and supplemental Figs. S3). Therefore, although P-Smad2 protein and Smad2 RNA are readily detectable in ES cells maintained in LIF, P-Smad2 is bound to chromatin at the distal promoter of Afp only after RA-induced Foxa1 expression and binding of Foxa1 to Afp chromatin. Levels of P-Smad2 actually decreased in RA-treated ES cells (Fig. 4D), but interaction with Afp chromatin occurs at this time due to Foxa1. Importantly, Foxa1 and Smad4 are bound to the Afp distal promoter simultaneously, as demonstrated through sequential ChIP assays of these two proteins in RA-treated ES cells (Fig. 4E).

As a pioneer transcription factor, Foxa1 increases accessibility of repressed chromatin compacted by the linker histone H1, as established and well characterized at the Alb enhancer during embryonic development (13). We determined that the DNA binding of Foxa1 reduces nucleosome occupancy and decreases levels of histone H1 at the Afp distal promoter, during RA-mediated differentiation of ES cells. We performed ChIP assays of histones H3 and H1 and saw a significant drop in nucleosomes occupying this region, as well as loss of linker histone H1 (Fig. 4F), concurrent with RA-induced Foxa1 binding (Fig. 4A). Thus, differentiating ES cells show increased binding of Foxa1 and Smads alongside lower nucleosome occupancy, as represented by H3, and decreased linker histone H1.

nucleosome presence is found when the TGF-β pathway was inhibited (Fig. 5B). Depletion of Foxa1, by siRNA, is clearly effective in reducing Foxa1 protein bound at the Afp regulatory region, and likewise Foxa1 binding is essential for access to chromatin and binding of P-Smad2 and Smad4 proteins (Fig. 5C).

Foxa1 and Smad Proteins Activate Afp in ES Cells

**FIGURE 4. Activated Smad proteins bind chromatin at the Afp distal promoter only when Foxa1 is bound and silent chromatin is altered.** A–C, E, and F, chromatin immunoprecipitation was performed on ES cells maintained in LIF or treated with RA for 4 days. DNA from immunoprecipitations for Foxa1 (A), Smad4 (B), P-Smad2 (C), H3 and H1 (F) was analyzed by real time PCR for binding to the Afp distal promoter and to the Nanog p53 response element. Results were graphed as percent bound minus IgG relative to input (A–C) or as normalized to LIF (F). E, re-ChIP analysis reveals simultaneous association of Foxa1 and Smad4 at the Afp distal promoter in RA-treated ES cells, but not in LIF-treated ES cells. A 1° ChIP was done with Smad4 and reactions were pooled, released from protein A beads, and then reimmunoprecipitated for IgG or Foxa1. Error bars represent S.D. of at least three repetitions; ns, not significant. D, whole cell lysates from ES cells maintained with LIF or differentiated with RA were probed with antibodies to Smad4, Smad2, P-Smad2, and β-actin. Triangles indicate differential loading.
Foxa1 and Smad Proteins Activate Afp in ES Cells

Afp expression is controlled by signaling pathways during differentiation of ES cells. Foxa1 is a pioneer factor that plays a role in maintaining chromatin structure, which is important for gene activation. Smad proteins, specifically Smad2 and Smad4, are involved in TGF-β signaling and are critical for the activation of Afp.

**FIGURE 5.** Foxa1 acts upstream of Smads to mediate Smad binding, nucleosome occupancy reduction, and H3K9 acetylation. Chromatin immunoprecipitation was performed on ES cells treated with RA for 4 days and transfected with either Foxa1-targeted or scrambled siRNA oligos (A, C, and H) or treated with SB431542 or dimethyl sulfoxide (B, F, and I). Transfection of the siRNA oligos for Foxa1 or for a non-target control and treatment with SB431542 or dimethyl sulfoxide occurred coincident with addition of RA. DNA from immunoprecipitations was analyzed by real-time PCR for binding to the Afp distal promoter. Levels of histone H3, determined by a separate immunoprecipitation, are expressed as a ratio to levels of histone H3, determined by a separate immunoprecipitation. Error bars represent S.D. from at least three repetitions. ns, not significant.

Pioneer Function of Foxa1 in Embryonic Stem Cells—Studies of the Foxa family of proteins by Zaret and colleagues (13, 14, 62) led to the development of a pioneer transcription factor model of de novo transcription activation in somatic cells. In this model, a protein interacts favorably with its target sequence in highly condensed chromatin and alters the chromatin structure such that a local domain becomes accessible and remodeled independently of ATP requiring enzymatic activity (13, 14, 62). More recently, the pioneer factor mechanism was elaborated in embryonic stem cells to include DNA methylation and regulation of enhancers and critical activation sites (11). Bisulfite sequencing of genomic DNA from pluripotent, embryonic stem cells revealed that windows of unmethylated DNA occur at regulatory sites, with significant roles in gene expression during differentiation of stem cells. Foxd3, a self-renewal and pluripotency factor that belongs to the forkhead family, plays a major role in maintaining methylation-free regions by binding to forkhead regulatory elements. The hypo-methylated status of the Alb enhancer is maintained by Foxd3 and, with differentiation, Foxa1 displaces Foxd3 prior to activation of Alb (11). Loss of Foxd3, by Cre-mediated excision in ES cells (12), induces differentiation along multiple lineages, including trophoderm, endoderm, and mesoderm, despite continued expression of known pluripotency factors Oct4, Sox2, and Nanog. Thus, pioneer functions of Foxa1 must be modified to include displacement of a stem cell factor, previously bound to silent chromatin, during activation of Alb in stem cells (11).

Transposable Elements in Regulation of Gene Expression—Genomic insertion of TE can have a direct influence on nearby

Additional text discussing mechanisms of gene regulation and the role of Foxa1 and Smad proteins in Afp expression is included, focusing on the critical role of Foxa1 in altering chromatin and facilitating the binding of Smad2/4 proteins to their DNA-binding sites, recruit enzyme complexes that acetylate nucleosomes. Thus, Foxa1 and TGF-β effectors, P-Smad2 and Smad4, are key players in a functional pathway, leading to de novo activation of Afp during differentiation of ES cells: RA-mediated induction of Foxa1 expression, binding of Foxa1 within a nucleosome occupied but unmethylated SINE/MER1 regulatory unit, alteration of chromatin structure at this region, accessibility and binding of Smad2/4 proteins, followed by acetylation of chromatin and activation of gene expression.

**DISCUSSION**

Pioneer Function of Foxa1 in Embryonic Stem Cells—Studies of the Foxa family of proteins by Zaret and colleagues (13, 14, 62) led to the development of a pioneer transcription factor model of de novo transcription activation in somatic cells. In this model, a protein interacts favorably with its target sequence in highly condensed chromatin and alters the chromatin structure such that a local domain becomes accessible and remodeled independently of ATP requiring enzymatic activity (13, 14, 62). More recently, the pioneer factor mechanism was elaborated in embryonic stem cells to include DNA methylation and regulation of enhancers and critical activation sites (11). Bisulfite sequencing of genomic DNA from pluripotent, embryonic stem cells revealed that windows of unmethylated DNA occur at regulatory sites, with significant roles in gene expression during differentiation of stem cells. Foxd3, a self-renewal and pluripotency factor that belongs to the forkhead family, plays a major role in maintaining methylation-free regions by binding to forkhead regulatory elements. The hypo-methylated status of the Alb enhancer is maintained by Foxd3 and, with differentiation, Foxa1 displaces Foxd3 prior to activation of Alb (11). Loss of Foxd3, by Cre-mediated excision in ES cells (12), induces differentiation along multiple lineages, including trophoderm, endoderm, and mesoderm, despite continued expression of known pluripotency factors Oct4, Sox2, and Nanog. Thus, pioneer functions of Foxa1 must be modified to include displacement of a stem cell factor, previously bound to silent chromatin, during activation of Alb in stem cells (11).

Transposable Elements in Regulation of Gene Expression—Genomic insertion of TE can have a direct influence on nearby
gene expression, generally due to introduction of insertional mutations, regulatory promoters, or termination sequences (63, 64). Repetitive elements may induce genomic instability, and are often sites of DNA methylation, as an evolutionary means of silencing their effects (65). We find that neighboring SINE B2-type and MER-1 elements, together ~300 nucleotides, play a significant role in expression of Afp in mouse ES cells. A Foxa1/p53/SBE response unit lies within this region, and is absent of CpG sequences and free of DNA methylation. Lack of substrate, rather than active prevention, is the source of methylation-free DNA, upstream and at the binding site of Foxa1. Binding of Foxd3, prior to stem cell differentiation and replacement by Foxa1, does not occur at the Afp distal promoter, indirectly supporting a role for Foxd3 in prevention of DNA methylation at Foxa1 DNA-binding sites (11). CpG dinucleotide sequences occur immediately downstream of the Foxa1/p53/SBE unit, and these are highly methylated in both pluripotent and differentiating stem cells. These findings support the idea that TE insertion at Afp facilitates activation of expression in differentiating stem cells.

Interestingly, recent analyses of repetitive DNA elements and their impact on gene expression suggest that endogenous retroviral elements are a frequent source of transcription factor-binding sites during the evolution of a given genome (63). In primates, functional tumor suppressor p53 regulatory elements are highly represented among these insertion elements, which may introduce multiple regulatory sites as master regulatory units (63). Inclusion of a binding site for a pioneer transcription factor, such as Foxa1, would contribute a powerful ally in evolution of a p53-regulatory network (66), as attributed to retroelement insertion (63).

Our previous studies showed that the composite Foxa1/p53/SBE plays a major role in regulation of AFP during hepatic development, as well as derepression/reactivation in hepatocellular carcinoma or regeneration of liver. In vitro, recombinant Foxa1 binds this site to relieve repression of Afp chromatin structure and expression (24). Over a developmental time course of Afp repression, p53 displaces Foxa1 from their shared regulatory site and anchors a number of co-repressor proteins, which lead to repression of chromatin structure by histone deacetylation, and methylation of H3K9 and HP1 interaction (25, 35, 48, 49). Reversion of Afp expression to an active state, induced by partial hepatectomy and regeneration of liver, is marked by loss of p53 binding and renewal of Foxa1 interaction with activated Afp chromatin (48). Whether a Foxa1-binding site, taken out of context from a region defined by CpG-free TE, would have similar regulatory responsiveness to multiple regulatory stimuli is unknown.

Focka1 Is a Gatekeeper of Chromatin Regulation—TGF-β1, Activin, and Nodal ligands in the TGF-β family, signal through activated Smad2/3, which are phosphorylated in undifferentiated ES cells and exhibit nuclear/cytoplasmic shuttling in response to TGF-β (67, 68). Smad proteins are weak DNA-binding proteins, often relying on additional proteins, e.g. forkhead family members, for stable interaction with DNA (52). Cross-talk between forkhead factors and TGF-β signaling at this level has more than one mechanism. For example, Forkhead factor Foxh1 (Fast-2) interacts directly with Smad2 through its Smad-interaction motif and forms a TGF-β-inducible, Foxh1-Smad2-Smad4 complex that binds DNA (50, 52). Additionally, Foxo3a directly interacts with P-Smad3 to activate expression of the cell cycle regulator p21Cip1 in epithelial cells (51), which is blocked by the forkhead box protein Foxg1 (51). However, Foxa1 does not contain a detectable Smad-interaction motif (70) domain of Smad protein interaction, and our data do not support a model of Foxa1/Smad protein-protein interaction in activation of Afp. We find that Smad2 phosphorylation is greater in undifferentiated mouse ES cells than in RA-differentiated cells. Despite the availability of P-Smad2 and Smad4, there is no measurable binding to their regulatory sites until Foxa1 is expressed and binds to chromatin.

The role for Foxa1 in creating access to chromatin for transcription activator binding was defined by seminal studies of Alb activation during hepatic development (15). Our current study of stem cell differentiation supports the pioneer model of transcription activation for Foxa1, established in vitro and in somatic cells, where histone H1 is displaced and nucleosome occupancy or density is decreased (13, 14). H3K4me2 is important for Foxa1 binding to estrogen-response elements in MCF7 breast cancer cells, and H3K4me2 is implicated in cell-type-specific recruitment of Foxa1 to chromatin (71). In pluripotent ES cells, we find no H3K4me2 at the Foxa1-binding site in Afp chromatin. However, with RA treatment and differentiation, chromatin modification by H3K4me2 occurs concomitantly with Foxa1 binding (supplemental Fig. S3). Although the chromatin encompassing the distal promoter of Afp lacks detectable activating or repressive modifications of histones, and the Foxa1/p53/SBE response element resides within a methylation-free TE, Foxa1 is essential for simultaneous binding with Smad transcription factors to the Afp distal promoter. Cross-talk between TGF-β signaling and Foxa1, in activation of Afp, occurs at the level of increased activation of Afp chromatin by H3K9 acetylation, once Smad proteins are bound to chromatin.

TGF-β signaling is implicated in both maintenance of stem cells and their differentiation (69, 72). The time course of intersections between forkhead factors and these pathways may in part direct signaling effectors to specific target genes and contribute to the pleiotropic effects of TGF-β ligands in stem cells. Taken together with the ability of specific forkhead factors, e.g. Foxa1 (16) and Fox1 (26), to interact with chromatin as stable epigenetic regulators through mitosis, additive or cooperative interactions between forkhead proteins, and other signaling pathways may have long term consequences during differentiation of stem cells.

Acknowledgments—We are grateful to S. Stratton, W.-W. Tsai, and other members of the Dent laboratory, as well as R. Behringer, M. Galko, S. Dent, and members of the Dent laboratory for advice, reagents, and discussions.

REFERENCES
1. Keohane, A. M., O’Neill, L. P., Belyaev, N. D., Lavender, J. S., and Turner, B. M. (1996) Dev. Biol. 180, 618–630
Foxa1 and Smad Proteins Activate Afp in ES Cells

(2008) Proc. Natl. Acad. Sci. U.S.A. 105, 944–949
67. James, D., Levine, A. J., Besser, D., and Hemmati-Brivanlou, A. (2005) Development 132, 1273–1282
68. Varelas, X., Sakuma, R., Samavarchi-Tehrani, P., Peerani, R., Rao, B. M., Dembowy, J., Yaffe, M. B., Zandstra, P. W., and Wrana, J. L. (2008) Nat. Cell Biol. 10, 837–848
69. Watabe, T., and Miyazono, K. (2009) Cell Res. 19, 103–115
70. Randall, R. A., Howell, M., Page, C. S., Daly, A., Bates, P. A., and Hill, C. S. (2004) Mol. Cell. Biol. 24, 1106–1121
71. Lupien, M., Eeckhoute, J., Meyer, C. A., Wang, Q., Zhang, Y., Li, W., Carroll, J. S., Liu, X. S., and Brown, M. (2008) Cell 132, 958–970
72. Pucéat, M. (2007) Cardiovasc. Res. 74, 256–261