Paradoxical Role of DNA Methylation in Activation of FoxA2 Gene Expression during Endoderm Development*

Keren Bahar Halpern, Tal Vana, and Michael D. Walker

From the Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel

Background: The transcription factor FoxA2 is a key mediator of endoderm development and pancreas gene expression. The transcription factor FoxA2 is a master regulator of endoderm development and pancreatic beta cell gene expression. To elucidate the mechanisms underlying the activation of the FoxA2 gene during differentiation, we have compared the epigenetic status of undifferentiated human embryonic stem cells (hESCs), hESC-derived early endoderm stage cells (CXCR4+ cells), and pancreatic islet cells. Unexpectedly, a CpG island in the promoter region of the FoxA2 gene displayed paradoxically high levels of DNA methylation in expressing tissues (CXCR4+ cells) and pancreatic islet cells. Upon differentiation of stem cells, pluripotency is gradually lost as cells become lineage-specified and ultimately acquire cell-specific functions. The choice between pluripotency or differentiation depends on a specific set of genes, primarily transcription regulators. The classic “pluripotency factors” Oct4, Nanog, and SOX2 act to maintain the pluripotent state and to block differentiation. On the other hand, “master regulators” of differentiation act in the opposite fashion to promote differentiation.

Results: DNA methylation is paradoxically associated with active expression of the FoxA2 gene.

Conclusion: Developmental control genes such as FoxA2 may be regulated by a novel genetic mechanism.

Significance: Clarification of this mechanism may provide a better understanding of how developmental pathways are activated.

Development of multicellular organisms involves progression from pluripotency through multipotency, leading ultimately to terminal differentiation. This process is mediated by transcription factors and epigenetic modifications that orchestrate the accompanying changes in gene expression; in turn this creates inheritable cellular memories characteristic of both specific lineages and cell types (1, 2). Human embryonic stem cells (hESCs) are a valuable tool for studying differentiation and development (3), because they possess the potential to differentiate into cells of all three major germ layers, thus permitting access to cell populations corresponding to early stages of development. This provides an opportunity to define the mechanisms controlling establishment of the primary germ cell layers. This is valuable for expanding our understanding of basic developmental biology, as well as for establishing a practical source of desired cell populations for cell replacement therapy.

Upon differentiation of stem cells, pluripotency is gradually lost as cells become lineage-specified and ultimately acquire cell-specific functions. The choice between pluripotency or differentiation depends on a specific set of genes, primarily transcription regulators. The classic “pluripotency factors” Oct4, Nanog, and SOX2 act to maintain the pluripotent state and to block differentiation. On the other hand, “master regulators” of differentiation act in the opposite fashion to promote differentiation.

Genes of the FoxA family are expressed at an early stage of mouse development during the formation of definitive endoderm (5–8). The first to be activated is FoxA2, initially in the anterior primitive streak and the node at embryonic day 6.5 in the mouse (6, 9). FoxA2 (together with family member FoxA1) is required for proper development of endoderm-derived organs such as liver, pancreas, lung, and prostate and is thus considered a master regulator of early endoderm formation and endoderm lineage establishment (10). FoxA2 has been described as a “pioneer factor” that binds to the chromatin of a progenitor cell prior to target gene activation along with factors that help to modify the chromatin upon gene activation during cellular differentiation (10–12). FoxA2 is also expressed later in development, in endoderm-derived tissues, including liver, lung, stomach, pancreas, small intestine, and colon (13). In pancreatic islets, FoxA2 is a direct activator of Pdx1 (14), which places it near the top of the pancreatic transcription factor hierarchy. Currently, little information is available regarding the precise molecular mechanisms that control the activation of this gene during different stages of development (15).

In this study, we report that a CpG island in the promoter region of the FoxA2 gene paradoxically shows high levels of DNA methylation in endoderm lineage FoxA2-expressing tissues and low levels in nonexpressing tissues. Using reporter gene assay, this CpG island region was found to inhibit reporter
gene expression, an effect that was abolished by in vitro methylation of the inhibitory fragment. Furthermore, using a ChIP assay, we observed that this region is bound by the Polycomb group protein SUZ12 and the DNA methyltransferase DNMT3b in hESCs, where this region is unmethylated; this binding is accompanied by high levels of H3K27me3. Furthermore, inhibition of methylation by 5-aza-2′-deoxycytidine (5-aza-dC) or directed DNMT3b knockdown led to decreased activation of FoxA2 during stem cell differentiation toward endoderm progenitors. These results indicate that methylation of the CpG island plays a key role in FoxA2 gene activation in both early and late stages of FoxA2 expression.

**EXPERIMENTAL PROCEDURES**

**Cell Culture**

The following established cell lines were used in this study: 293T (human embryonic kidney cells) and MIN6 (mouse β cells) (16). 293T cells were grown in DMEM supplemented with 10% FCS and penicillin (200 IU/ml) with streptomycin (100 μg/ml). MIN6 cells were grown on Falcon tissue culture plates in DMEM supplemented with 15% FCS, penicillin/streptomycin, 2 mM L-glutamine, 5.6 mM glucose, and 0.5% β-mercaptoethanol. All experiments using MIN6 cells were performed with passages not higher than 33.

Pancreatic islets were obtained from the Geneva University Hospital, Cell Isolation and Transplantation Center through Juvenile Diabetes Research Foundation Award 31-2008-416 (European Consortium of Islet Transplantation, Islet for Basic Research program). Experiments were performed under approval of the Weizmann Institute Bioethics and the Embryonic and Stem Cell Oversight committee.

**hESC Culture and Differentiation**—HUES2 cells were obtained from Dr. N. Benvenisty and Dr. D. Melton. H9 cells were obtained from WiCell Research Institute (Madison, WI). Experiments with these cells were approved by the Embryonic and Stem Cell Oversight Committee of the Weizmann Institute of Science. HUES2 and H9 cell lines were cultured essentially as described (17).

**Flow Cytometry**

Cells were dissociated using enzyme-free Hanks’-based cell dissociation buffer (Invitrogen, 13150-016) for 15 min, followed by quenching with 10% FCS in PBS. Staining of cells was carried out in PBS containing 3% FCS using the following antibody (from BD Biosciences): phycoerythrin mouse anti-human CD184 (CXCR4, catalog number 555974). Propidium iodide (from BD Biosciences): phycoerythrin mouse anti-human CD184 (CXCR4, catalog number 555974). Propidium iodide (from BD Biosciences): phycoerythrin mouse anti-human CD184 (CXCR4, catalog number 555974).

**Mini Chromatin Immunoprecipitation (miniChIP)**

Preparation of Soluble Chromatin—Between 10⁵ and 10⁶ cells were resuspended in 1 ml of serum-containing medium and fixed by the addition of formaldehyde (1% final concentration) for 10 min at room temperature with gentle mixing. Fixation was stopped by adding freshly prepared glycine (0.125 M) and incubating for 5 min at room temperature. Next, the cells were spun down (8000 rpm 4 °C) and washed once (4000 rpm 4 °C) in cold PBS. Cells were either fast frozen on dry ice and stored at −70 °C or lysed in 1 ml of lysis buffer (25 mM Tris-HCl, pH 8.1, 150 mM NaCl, 1% Nonidet P-40, 0.5% SDS, 1% deoxycholate, 1% Triton X-100) for 10 min on ice, and the lysate was sonicated using a Bioruptor 300 (Diagenode) in an ice bath. To achieve an average fragmentation of 200–600 bp, 30–45 sonication cycles of 30 s, each with a 30-s interval between them, were required according to cell type (45 cycles for MIN6 cells and 30 cycles for hESC). Next, the lysed chromatin was centrifuged for 10 min at 13,000 rpm and 4 °C, and the supernatant was either stored at −70 °C or used directly for immunoprecipitation.

Immunoprecipitation Reaction—Sonicated chromatin from 10⁵ cells was precleared by adding 50 μg/ml salmon sperm DNA, 100 μg/ml yeast RNA, 1 μg/ml BSA (final concentrations), and 35 μl of 50% (w/v) protein A-Sepharose beads (in 150 mM TSE buffer: 2 mM EDTA, 1% Triton X-100, 20 μM Tris, pH 8.1, 0.1% SDS, and 150 mM NaCl) and incubating for 30 min at 4 °C on a rotating wheel. In addition, the desired amount of protein A-Sepharose beads was precleared using the same conditions. Following preclearance, the chromatin was centrifuged (13,000 rpm 15 min 4 °C), and the supernatant was transferred to a new tube. From 1.1 ml of precleared chromatin, 50 μl was set aside as input sample and stored at −20 °C, and 500 μl were taken for incubation with the desired antibodies or sera. The chromatin and antibodies were incubated overnight at 4 °C on a rotating wheel. On the next day, the extracts were incubated with 35 μl of 50% precleared protein A-Sepharose suspension for 2 h at 4 °C. Next, the beads were sequentially washed once with TSE-150, three times with TSE-500 (2 mM EDTA, 1% Triton X-100, 20 μM Tris, pH 8.1, 0.1% SDS, and 500 mM NaCl), and two times with TE (10 mM Tris-HCl, pH 8.1, 0.1 mM EDTA). The immune complexes were then eluted by incubating the beads with 100 μl of elution buffer (1% SDS, 0.1 mM NaHCO₃) for 30 min. The eluates and input samples were heated at 65 °C overnight to reverse the formaldehyde cross-links. The eluates were purified using MinElute columns (Qiagen) according to the manufacturer’s instructions and eluted in 35 μl of water. A 2-μl sample was used for amplification by quantitative real-time PCR as described (17).

Antibodies for ChIP Assay—The following antibodies were used in ChIP assays: α tri-methylated H3 ([Lys7] Millipore 07-449), α IgG (control antibody Abcam ab46540), αSUZ12 (Abcam ab12073), αDNMT3b (Abcam ab2851), and αDNMT3a (Abcam ab2850).

**Real Time Quantitative PCR**

Total RNA was extracted from cell cultures using the Trizol Reagent procedure (Molecular Research Center Inc.) according to the manufacturer’s instructions. RNA from sorted populations was isolated using RNeasy MinElute Cleanup kit (Qiagen 74204). DNA was eliminated using TURBO DNA-free kit (Ambion AM1907). DNase I-treated RNA was reverse transcribed into cDNAs using affinity Script reverse transcriptase (Agilent Technologies) with 10 μM random hexamer primers (Roche), as recommended by the manufacturer (Invitrogen). Transcript levels were measured using real time quantitative
DNA Methylation Activates FoxA2 Gene

PCR on a 7300 ABI real time PCR system using Power SYBR green PCR master mix (Applied Biosystems). The levels of each gene were normalized using β-actin as an endogenous control mRNA.

Bisulfite Sequencing Analysis

Bisulfite conversion of genomic DNA was done using the Epitect bisulfite kit (Qiagen 59104) according to the manufacturer’s instructions. PCR primers were designed using Methyl Primer Express® software v1.0. Specific fragments were amplified by PCR (Roche: FastStart Taq 250U (5 units/2.5 l) 12-032-953-001) using 20 ng of DNA following bisulfite treatment and 2.5 l of mixed primers (0.5 l/l) under the following conditions: 95 °C for 5 min; 5 cycles of 95 °C for 1 min, 53 °C for 3 min, and 72 °C for 3 min; 40 cycles of 95 °C for 30 s, 55 °C for 45 s, and 72 °C for 45 s; followed by an additional incubation at 72 °C for 7 min. Following PCR amplification, the DNA was extracted from the gel using the RBC gel extraction kit, cloned into the Promega pGEM-T vector system, transfected into competent bacteria that were grown on LBAgar plates containing ampicil-n, 0.2 mM X-gal, and 0.1 mM isopropyl b-thiogalactopyranoside. The DNA from positive clones was isolated using a mini plasmid isolation kit and sequenced using T7 or Sp6 prim-ers. Obtained sequences were aligned with unconverted genomic DNA sequence using the biq-analyzer tool, and DNA methylation levels were calculated. DNA samples were purchased from Biochain Inc.

Transient Transfections

Transfections of 293T or MIN6 cells were carried out using the cationic polymer transfection reagent jetPEI (Polyplus Transfection) according to the manufacturer’s instructions. Transfections of hESCs were carried out using the lipid-polymer mixture TransIT-2020 (Mirus) according to the manufacturer’s instructions. Briefly, hESCs were grown for 24 h on growth factor-reduced Matrigel. On the day of transfection, 200,000 cells/well in a 24-well plate coated with Matrigel were transfected with a mixture of 0.5 gDNA (consisting of 100 ng of reporter construct, 50 ng of internal control, and 350 ng of pUC18) and 2 l of TransIT-2020 reagent in optimem medium (according to the manufacturer’s protocol; Mirus). Cells were harvested 48 h after transfection, and cell extracts were subjected to protein concentration determination and luciferase reporter enzyme assays as described (19).

Plasmid Constructions

Plasmids for promoter activity measurements were constructed using pGL3-basic vector (Promega). The region upstream from FoxA2 (FoxA2 gene promoter region) was generated by PCR using primers 5’-CCGTTGGAGTGAATGGTGCCTCC-3’ (top) and 5’-GCCGGCTCCGCTCTCCG-3’ (bottom) with template of human genomic DNA. The PCR fragment was subcloned into CloneJET™ PCR cloning kit (Fermentas) and ligated to pGL3-basic at the BglII site (creating construct + Cpg2). Fragments from this vector were generated by PCR using the following primers: forward −CpG2 5’TCA-CAGGGCTAACCCAGAACAG-3’ (top) and the same bottom primer, and forward −E 5’-CATCATTGGATTCTTGATGATT CTTC-3’ (top) and the same bottom primer. The PCR fragments were subcloned into the pGL3 vector using BgIII and XhoI sites.

Plasmids for the in vitro methylation assay were constructed using pGL2-control vector (E1611; Promega, Madison, WI). Fragments from the FoxA2 gene promoter region in CloneJET™ vector were generated by PCR using the following primers: for 675 5’-CGGTGGAGTGAATGGTGCCTCC-3’ (top) and 5’-TCTGTTCGCTTCGCTCGCC-3’ (bottom), for 2a 5’-GGACACTTCATAAGAGGGCAT-3’ (bottom) and the same top primers, and for 2b 5’-ATGCCCTCTCTAT GAAGTGGTTCTC-3’ (top) and 5’-TCTGTTCGCTTCGCTCGCC-3’ (bottom). The PCR fragment was subcloned into pGL3-basic using BgIII and XhoI sites.

In Vitro Methylation—Plasmids were in vitro methylated using HhaI or HpaII methyltransferases (New England Biolabs) as follows: 5–10 g of plasmid were incubated for 1.5 h at 37 °C with the specified enzyme (1 unit/1 g of DNA) in the presence of 1× buffer and 0.32 mM S-adenosyl methionine. After the reaction, the DNA was phenol-chloroform-extracted, EtOH-p precipitated, and resuspended in water. Methylated DNA was digested with methylation-sensitive restriction enzymes (HhaI or HpaII, respectively; New England Biolabs) and resolved on agarose gel to estimate the levels of methylation. Methylated DNA was used in transient transfection assays.

siRNA Transfection—hESCs were grown for 2 days on growth factor-reduced Matrigel. Cells were transfected with DharmaFECT2 (Dharmacon) transfection reagent mixture, according to the manufacturer’s instructions. Briefly, 95 l of 1× siRNA buffer (Dharmacon B-002000-UB-100) were mixed with ON-TARGET PLUS pool siRNA (Dharmacon) (final concentration, 50 nM) and 100 l of DEME F12 (mix 1). Each siRNA pool contained four different siRNAs targeting DNMT3b. A nontarget siRNA pool was used as control. 4 l of DharmaFECT2 transfection reagent mixture were mixed with 196 l of DEME F12 (mix 2). Both mixtures were incubated separately for 5 min at room temperature and then combined and incubated for an additional 20 min at room temperature (mix 3). Mix 3 was combined with 2 × 105 hESCs in a 6-cm plate coated with Matrigel in the presence of 10 l ROCK inhibitor (Sigma Y-27632). After 48 h, cells were either harvested for RNA and protein purification or permitted to differentiate (20) for 3 days and then harvested for RNA.

RESULTS

DNA Methylation at the FoxA2 Locus—The FoxA2 locus contains three CpG islands: CpG1 and CpG2a are upstream of the minimal promoter, whereas CpG3 encompasses the transcription start site (TSS) and the majority of the transcribed region (Fig. 1A). To determine the methylation state of the FoxA2 gene during development, we performed bisulfite DNA sequencing. Genomic DNA of different cell types was treated with sodium bisulfite, and PCRs were performed to permit analysis of regions within and outside of the three CpG islands. Bisulfite sequencing of CpG 3 (10% of total length −321 bp of 3368 bp), which covers the FoxA2 TSS, was performed in four cell types: hESCs and human skeletal muscle, in which the FoxA2 gene is inactive; and hESC-derived definitive endoderm
(CXCR4+) and human pancreatic islets, in which the FoxA2 gene is active. No significant methylation was observed in any of these cell types (<2%; data not shown), suggesting that this CpG island does not have a regulatory role in FoxA2 gene expression. By contrast, analysis of CpG2a revealed a paradoxical pattern of methylation: low levels of methylation were observed in tissues that do not express FoxA2 (hESC and skeletal muscle), and high levels were observed in FoxA2-expressing tissues (pancreatic islets, liver, and colon) (Fig. 1B). Furthermore, CXCR4+ cells (early definitive endoderm cells in which FoxA2 expression is commencing) also exhibited high levels of methylation (Fig. 1B).

By inspecting the sequence downstream to CpG2a, we observed that the CG-rich region extends ~250 bp beyond the core CpG island. We designated this CpG-rich region as CpG2b and determined methylation levels. Again, we observed relatively low levels of DNA methylation in nonexpressing tissues (hESC and skeletal muscle) and higher levels in expressing tissues (CXCR4+ cells, islets, liver, and colon) (Fig. 1C). These results indicate that CpG2b (although it is not identified by the computer algorithm used) is differentially methylated by a similar mechanism as CpG2a: this unusual methylation pattern may play a role in activating FoxA2 gene expression.

Consistent with these results, in a genome-wide nonquantitative study (20), CpG2a showed an expression-dependent pattern of differential methylation: FoxA2-expressing tissues (pancreas, liver, and several types of cancer cell) showed high levels of methylation, whereas nonexpressing tissues (including brain, hESC, skeletal muscle, and sperm) showed low levels of methylation. Interestingly, CpG1 does not display the paradoxical methylation pattern seen with CpG2 (20) and hence is probably less important for regulation of the FoxA2 gene.

**CpG2 as a Negative Regulatory Element**—An effective approach to assess the possible role of CpG2 in regulating FoxA2 gene transcription is a reporter gene assay, because plasmids propagated in *Escherichia coli* have unmethylated CpGs, which can be efficiently methylated *in vitro* using the enzymes HhaI and HpaII (21). Upon transfection to mammalian cells, CpG sequences do not become methylated (21). Several DNA fragments derived from the FoxA2 promoter were ligated upstream to the luciferase reporter gene (Fig. 2A). Construct +CpG2 contains the entire fragment spanning CpG2 to the TSS. From +CpG2, a fragment of 675 bp that contains the entire CpG2 (2a+b) was deleted, leaving the core promoter and enhancer (22) of the FoxA2 gene (construct −CpG2). Construct −E contains only the core promoter without the enhancer region. The three constructs were transfected to undifferentiated hESCs, the mouse pancreatic beta cell line MIN6, and 293T cells, and luciferase levels were measured. The presence of the promoter fragment that contains CpG2 (construct +CpG2) did not produce a significant increase in activity in any cell type tested, as compared with the pGL3-basic vector (Fig. 2B). On the other hand, deletion of the 675 bp containing the CpG2 (construct −CpG2) led to a significant increase in promoter activity in all cell types, but most strikingly in MIN6 cells (Fig. 2B). This indicates 1) that the 675-bp fragment contains a negative regulatory element (note that the CpG island is unmethylated in these constructs, because plasmids are prepared from *E. coli* and 2)
that the −CpG2 fragment contains a positive regulatory element that is preferentially active in beta cells. Consistent with this, deletion of the enhancer region (construct −E) led to a decrease in promoter activity only in MIN6 cells (Fig. 2B).

**DNA Methylation Abolishes the Negative Effect of CpG2**—To extend this analysis, we tested whether the 675-bp fragment displays a negative effect when removed from the context of the intact promoter and whether this effect is influenced by DNA methylation. For this, we inserted the following three fragments to pGL2-control vector, a luciferase plasmid that contains the SV40 promoter and enhancer: the 675-bp fragment (CpG2a), a fragment containing only CpG2a, and a fragment containing only CpG2b (designated as a+b, a, and b, respectively in Fig. 2, A and C). These fragments were inserted between the SV40 enhancer and promoter, upstream of the luciferase gene. All plasmids were tested in unmethylated form and following in vitro methylation by HhaI or HpaII methylase.

In their unmethylated state, all three fragments produced a 30–50% decrease in promoter activity relative to control plasmid (Fig. 2D, black bars), indicating that all three fragments contain functional negative regulatory elements. Methylation with HpaII eliminated the negative activity only in the a+b fragment (Fig. 2D, white bars), perhaps because of the absence of HhaI sites in the CpG2b region (Fig. 2C). These results are consistent with the idea that methylation of CpG2 plays a positive role in regulating FoxA2 gene promoter activity.

**Binding Analysis of Repressive Proteins to the CpG Island**—DNA methylation typically correlates with gene silencing and can repress gene expression by directly blocking access of transcriptional regulatory factors to target DNA. In the FoxA2 gene, there is apparently a paradoxical opposite effect of DNA methylation on gene expression. DNA methylation at FoxA2 CpG2 occurs selectively in expressing tissues; methylation at the CpG island appears to activate gene expression, whereas in nonexpressing tissues, a lack of methylation inhibits gene expression. To address the mechanisms involved, we used ChIP assay to examine the binding of regulatory proteins and the histone methylation status of the CpG island region. As control, we examined the binding to the coding sequence (CDs) of the FoxA2 gene, which is not expected to be recognized selectively by transcription factors.

SUZ12 is a member of the Polycomb group protein family and is a component of the complex PRC2 (Polycomb repressive complex) that acts to repress transcription (23). To determine
whether SUZ12 binds to CpG2 in vivo, we performed ChIP analysis using hESCs, CXCR4+/H11001, or human islets cells (50,000 cells) with anti-SUZ12 antibody. SUZ12 occupancy on the CpG island region was compared with its binding to Pax1 CDs, which is known to be bound selectively by SUZ12 in ESCs (24) and hence serves as a positive control and also to FoxA2 CDs and Pax1 CpG island, that are not expected to bind the SUZ12 protein. In hESCs, SUZ12 binds FoxA2 CpG2 to a similar extent as to the Pax1 CDs (Fig. 3A). Furthermore SUZ12 occupancy in hESC was significantly higher than its occupancy in CXCR4+/H11001 and islet cells (Fig. 3A), whereas relatively low levels of binding were observed in the FoxA2 CDs and Pax1 CpG island in all of the cells tested.

DNA methylation is typically removed following zygote formation and re-established around the time of implantation; the closely related de novo DNA methyltransferases DNMT3a and DNMT3b are required for this process (25, 26). During embryonic development, DNMTs are recruited by PRC2 proteins such as EZH2 and SUZ12 and act to methylate the DNA (27). DNMT3b occupancy on CpG2 was compared with its binding to Pax1 CDs and to FoxA2 CDs and Pax1 CpG island, which are not expected to bind the DNMT3b protein. In hESCs, DNMT3b binds the FoxA2 CpG2 to a similar extent as the Pax1 CDs. DNMT3b occupancy in hESCs was significantly higher than its occupancy in CXCR4+ and islet cells (Fig. 3A), whereas relatively low levels of binding were observed in the FoxA2 CDs and Pax1 CpG island in all of the cells tested.

DNA methylation is typically removed following zygote formation and re-established around the time of implantation; the closely related de novo DNA methyltransferases DNMT3a and DNMT3b are required for this process (25, 26). During embryonic development, DNMTs are recruited by PRC2 proteins such as EZH2 and SUZ12 and act to methylate the DNA (27). DNMT3b occupancy on CpG2 was compared with its binding to Pax1 CDs and to FoxA2 CDs and Pax1 CpG island, which are not expected to bind the DNMT3b protein. In hESCs, DNMT3b binds the FoxA2 CpG2 to a similar extent as the Pax1 CDs. DNMT3b occupancy in hESCs was significantly higher than its occupancy in CXCR4+ and islet cells (Fig. 3A). A low extent of binding was observed for the FoxA2 CDs and Pax1 CpG island in all of the cells tested.

DNMT3a occupancy on the FoxA2 CpG2 region was also tested and compared with its binding to the CpG island in the DMRT2 gene, a target of DNMT3a in hESCs (28), and with FoxA2 CDs and Pax1 CpG island, which are not expected to bind the DNMT3a protein (Fig. 3C). In hESCs, DNMT3a occupancy on the DMRT2 CpG island was significantly higher than its occupancy in CXCR4+ and islet cells. On the other hand, a low level of binding was observed in FoxA2 CpG2, CDs, and Pax1 CpG island in all cells tested. We therefore conclude that in hESCs, DNMT3b, but not DNMT3a binds preferentially to FoxA2 CpG2.

Polycomb group proteins mediate di- and trimethylation of histone H3K27 (H3K27me2 and H3K27me3) (23), a stable epigenetic mark associated with gene silencing. The levels of H3K27me3 were examined by ChIP in hESCs, CXCR4+, and islet cells using a specific antibody (Fig. 3D). In accordance with gene expression patterns, high levels of K27 methylation were observed at FoxA2 CpG2 and the protein coding region (CDs) in hESCs as compared with CXCR4+ and islet cells. For the Pax1 gene (CpG island and CDs), we observed the expected high levels of K27 methylation in all cells tested (because this gene is silent in these cells).

Inhibition of DNMTs during Differentiation—To test the hypothesis that DNA methylation is involved in activation of FoxA2 gene transcription during differentiation, we treated cells with 5-aza-dC, a potent inhibitor of DNA methylation. We tested the effect of 5-aza-dC on expression of the pluripotency marker Oct4 and the definitive endoderm genes CXCR4, SOX17, and FoxA2. Oct4 expression was down-regulated by ~40% (29), whereas CXCR4 and SOX17 were up-regulated (5- and 4,000-fold, respectively) following exposure to differentiation conditions for 3 days (18), in the absence or in the presence of 5-aza-dC (Fig. 4A), demonstrating that under these condi-
DNA Methylation Activates FoxA2 Gene

5-aza-dC does not block the differentiation process. FoxA2 expression, on the other hand was up-regulated by 600-fold following differentiation, but this effect was significantly inhibited (3-fold) by treatment with 5-aza-dC (Fig. 4A). These results suggest that DNA methylation is necessary for efficient activation of FoxA2 gene expression in differentiating hESCs; this may involve methylation of CpG2 by DNMT3b.

To directly test the role of DNMT3b in activation of FoxA2 gene expression, we inhibited the expression of DNMT3b by siRNA knockdown. Following siRNA transfection, DNMT3b mRNA levels were measured using quantitative real time PCR. Transfection of undifferentiated hESCs with DNMT3b siRNA led to a 60% reduction in DNMT3b mRNA levels as compared with control siRNA (data not shown). hESCs were exposed to siRNA directed against DNMT3b, or control siRNA, for 2 days followed by 3 days of differentiation. Following this combined protocol of gene knockdown and differentiation, RNA was extracted, and gene expression levels were measured using real time PCR. Treatment of cells with either control siRNA or siRNA directed against DNMT3b did not affect global differentiation as seen by increased expression of the endoderm markers CXCR4 and SOX17 (Fig. 4B). As expected, following differ-

FIGURE 4. Effect of 5-aza-2′-deoxycytidine and DNMT knockdown on FoxA2 gene expression. A, hESCs (HUES2) were cultured in nondifferentiating and differentiating conditions in the absence or presence of 5 μM 5-aza-2′-deoxycytidine. Expression of Oct4 (panel i), CXCR4 (panel ii), Sox17 (panel iii), and FoxA2 (panel iv) was determined by real time quantitative PCR. β-Actin was used as endogenous control. Expression levels are indicated relative to undifferentiated control. Expression levels are indicated relative to undifferentiated (untreated) HUES2 cells. Each data point represents the mean ± S.E. (n = 3). *, p < 0.05. B, hESC cells (H9) were transfected with either DNMT3b siRNA or control siRNA. After 2 days, cells were exposed to differentiation conditions. Expression of CXCR4 (panel i), Sox17 (panel ii), DNMT3b (panel iii), and FoxA2 (panel iv) was measured by real time quantitative PCR in undifferentiated cells and differentiated cells treated with control siRNA (si CT) or siRNA directed against DNMT3b (si DNMT3b). The expression of β-actin was used as endogenous control. Expression levels are indicated relative to undifferentiated hESCs. The results shown represent the means ± S.E. (n = 3). *, p < 0.05. AZA, 5-aza-2′-deoxycytidine; Diff., differentiated; Undiff., undifferentiated.
entiation, levels of DNMT3b were reduced (to 17%, siCT); treatment with siDNMT3b led to a further reduction (to 7%) (Fig. 4B). Levels of FoxA2, on the other hand, were elevated as expected following differentiation but were significantly decreased (~2-fold) following treatment with the siRNA directed against DNMT3b (Fig. 4B).

**DISCUSSION**

We have identified a CpG island located in the promoter region of the FoxA2 gene that appears to directly regulate expression of the gene according to its methylation status. Although DNA methylation typically correlates with gene silencing, in this case, we observed the opposite pattern, namely high levels of DNA methylation in expressing tissues (CXCR4+, islets, liver, and colon) and low levels in nonexpressing tissues. Many tissue-specific genes and genes overexpressed in tumors show methylation of CpG islands downstream of the TSS (30), showing that methylation at these regions does not block transcription (31). However, most DNA methylation studies have focused on proximal promoter regions: consistently, single-gene studies have shown that promoter methylation correlates with reduced transcriptional efficiency.

Given the low extent of methylation of CpG islands in the human genome, we assume that at the onset of gastrulation, an as yet unknown mechanism distinguishes FoxA2 CpG2 for methylation in cells of the endoderm lineage, as opposed to the bulk of CpG islands in the genome. A key question arising from our study is what mechanisms lead to this selective methylation of CpG2. One possibility is binding of factors that direct methylation to this region. A plausible candidate for this is the Polycomb group complex (PRCs). Polycomb group proteins are considered as negative regulators of gene expression: they are recruited to the promoter region and act to repress transcription by recruitment of DNA methyltransferases that directly methylate the DNA and promote binding of chromatin remodeling proteins. Indeed, using ChIP assay, we observed that the CpG2 region is bound by the Polycomb group protein SUZ12 and the DNA methyltransferase DNMT3b preferentially in undifferentiated hESCs as compared with CXCR4+ or islets cells. PRC proteins may be recruited to the genomic loci by sequence-specific factors that bind directly to the DNA. Consistent with this, it has been shown that de novo methylation patterns can be directed by specific cis-acting methylation-determining regions that precisely recapitulate methylation patterns that arise during differentiation (32); this study also showed that methylation-determining regions can bind transcription factors and protect the region from undergoing methylation. This scenario may apply in hESCs or nonendodermal tissues, where the CpG is unmethylated, whereas in endodermal tissues, such factors may be absent or unable to bind, allowing the recruitment of PRCs and DNMTs that methylate the region.

Recently, a methodology was reported (33, 34) that combines bisulfite conversion with ChIP and deep sequencing, thereby allowing a direct genome-wide interrogation of two epigenetic marks on the same DNA molecule. It was shown that depending on CpG density, H3K27me3 and DNA methylation either co-occur (low CpG density) or are mutually exclusive (high CpG density). Because CpG2 has high CpG density, it may be expected to show mutually exclusive H3K27me3 modification and DNA methylation. Indeed CpG2 shows high levels of H3K27me3 when the CpG island is unmethylated (hESC) and lower levels of H3K27me3 when the CpG island is methylated (CXCR4+ and islets). The low levels of H3K27me3 in FoxA2-expressing tissues may help to explain how the FoxA2 gene is active, even though its promoter is methylated. Further analysis of the chromatin state around the CpG2 region is needed to understand how the DNA methylation does not repress FoxA2 gene expression.

Typically, DNA methylation of gene promoters is associated with transcriptional silencing (35). Only in rare cases does gene methylation correlate with expression. Generally such methylation is located at regions far upstream of the TSS, e.g. the IL-8 gene and the imprinted Igf2 gene (36). In a recent genome-wide methylation analysis (20), it was shown that tissue-specific methylation of CpG islands in intragenic regions is associated with gene activation. It was suggested that these sites may bind methylation-sensitive repressors that suppress distant promoters. Alternatively, these regions may contain promoters for antisense RNAs that are active when this region is unmethylated (37). This assumption was supported by the presence of active chromatin modifications (H3K4me3) that correlate with TSS sites (38).

Thus, direct evidence for transcriptionally active yet heavily methylated promoters is rare (39, 40). In the FoxA2 gene promoter, CpG2 was shown to be methylated preferentially in expressing tissues. In addition, using reporter gene assay, we showed that the CpG island in its unmethylated form has a negative effect on transcription: this effect is abolished when the CpG island is methylated, consistent with the notion that methylation may block binding of repressing proteins. The presence of methylated DNA within the promoter region of a gene in expressing tissues raises the question of how methylation contributes to the activation of transcription. Two plausible mechanisms observed under situations of imprinting (41) are alternate promoter usage and presence of an insulator. In the alternate promoter model, methylation blocks access of the transcription machinery to the CpG island allowing the use of the basic core promoter of the gene, whereas in tissues where the CpG island is unmethylated, transcription initiates from the CpG island in preference to the basic core promoter. In the case of FoxA2, bioinformatics analysis did not identify promoter elements in the CpG2 region (data not shown). In the second mechanism, the methylated region contains a binding site for an insulator protein such as CTCF that blocks the effect of an enhancer located upstream of this region on a downstream gene. In the case of FoxA2, this possibility is more likely because in genome-wide chromatin state analysis done in human pancreatic islets (42), there is some evidence for such an upstream enhancer according to the presence of H3K4me around the CpG1 region. In addition, in a genome wide analysis of CTCF binding done in human hESCs, a binding site for CTCF was identified around the CpG2 region (43).

Based on these results, we hypothesize that in hESCs, a lack of DNA methylation permits binding of a sequence-specific repressor protein that mediates the binding to this region of DNA Methylation Activates FoxA2 Gene
SUZ12, a component of PRC2 that includes also the protein EED and EZH2. This protein directly methylates H3K27, a modification that also contributes to the repression of the gene in hESC. SUZ12 in turn recruits DNMT3b. We propose that upon differentiation, DNMT activation leads to CpG island methylation, which in turn causes loss of repressor protein binding. This is consistent with results showing that following dense DNA hypermethylation of promoters, EZH2 is no longer

**FIGURE 5. Proposed mechanism of activation of FoxA2 gene expression during development.** A, summary of DNA methylation and protein occupancy states in different cell types. B, a schematic illustration of the FoxA2 activation model proposed in the text. CpG dinucleotides are presented as black and white lollypops (methylated and unmethylated respectively). Polycomb repressive complex 2 (PRC2), DNMT3b, and RNA polymerase II (Pol II) are indicated. Repressive (H3K27me3) or activating (H3K4me2, H3Ac) histone modifications are shown. Hexagons represent activating transcription factors.
required to maintain DNA methylation and therefore is released (44). A combination of activating transcription factors that begin to be expressed in CXCR4+ cells, including transcription factors that activate the FoxA2 enhancer (22), together with alterations in chromatin modifications such as H3K27 trimethylation (repressive) and H3 acetylation or H3K4 dimethylation (activating) may then lead to activation of FoxA2 gene expression (Fig. 5). It is possible that pioneer factors play a role in this process, perhaps by recognizing the locus while still in its closed chromatin conformation or by interacting with the methylated CpG2. We have performed a bioinformatic analysis to detect potential transcription factor binding sites in CpG2 using MatInspector software. Among the factors identified are NEUROD, PTF1, PAX4/6, C/EBP, HNF6, and HNF1, which are implicated in regulation of pancreas and liver genes (45).

An additional important question to be resolved is whether DNMT3b is recruited by the Polycomb proteins but is inactive in hESCs or is recruited later. According to ChIP assay, the DNMT3b is already present on the CpG2 region in hESCs and may thus contribute to the low levels of methylation (20%) observed in these cells. The signal that activates DNMT to methylate CpG2 is unknown. Because Polycomb complexes contain several enzymatic activities, e.g. histone methylation, ubiquitination, and sumoylation, a post-translational modification that directly or indirectly alters DNMT activity and/or recruitment is a plausible mechanism.

To our knowledge, this study is the first to identify methylation-dependent gene activation as a mechanism for activating developmental regulators. Methylation at the promoter region of the FoxA2 gene both in early stages of development (e.g. CXCR4+, definitive endoderm) and in several endoderm-derived tissues (liver, pancreas, and colon) set this case as a key developmental and lineage commitment mechanism in endoderm fate decision during development. More detailed molecular dissection of this process will permit identification of the participants and determine whether such mechanisms control activation of additional developmental regulators and cell lineages.

Acknowledgments—We thank Prof. Chaim Cedar, Prof. Yehudit Bergman, Prof. Rivka Dikstein, and Dr. Dror Sharivkin for valuable discussions; Prof. Nissim Benvenisty and Prof. Doug Melton for providing HUES2 cells; Prof. J. Miyazaki for MIN6 cells; Sara Weiss for expert technical assistance; and Dr. Elena Ainbinder (Weizmann Institute Stem Cell Core Unit) for important input.

REFERENCES

1. Chambers, I. (2004) The molecular basis of pluripotency in mouse embryonic stem cells. Cloning Stem Cells 6, 386–391
2. O’Shea, K. S. (2004) Self-renewal vs. differentiation of mouse embryonic stem cells. Biol. Reprod. 71, 1755–1765
3. Soria, B., Skoudy, A., and Martin, F. (2001) From stem cells to beta cells: new strategies in cell therapy of diabetes mellitus. Diabetologia 44, 407–415
4. Chen, L., and Daley, G. Q. (2008) Molecular basis of pluripotency. Hum. Mol. Genet. 17, R23–R27
5. Ang, S. L., Wierda, A., Wong, D., Stevens, K. A., Cascio, S., Rossant, J., and Zaret, K. S. (1993) The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. Development 119, 1301–1315
6. Monaghan, A. P., Kaestner, K. H., Grau, E., and Schütz, G. (1993) Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. Development 119, 567–578
7. Sasaki, H., and Hogan, B. L. (1993) Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. Development 118, 47–59
8. Schier, A. F. (2003) Nodal signaling in vertebrate development. Annu. Rev. Cell Dev. Biol. 19, 589–621
9. Kaestner, K. H., Hiemisch, H., Luckow, B., and Schütz, G. (1994) The HNF-3 gene family of transcription factors in mice: gene structure, cDNA sequence, and mRNA distribution. Genomics 20, 377–385
10. Li, Z., White, P., Tuteja, G., Rubins, S., Sackett, S., and Kaestner, K. H. (2009) Foxa1 and Foxa2 regulate bile duct development in mice. J. Clin. Invest. 119, 1537–1545
11. Xu, J., Pope, S. D., Izaikehi, A. R., Attema, J. L., Papathanasiou, P., Watts, I. A., Zaret, K. S., Weissman, I. L., and Smale, S. T. (2007) Pioneer factor interactions and unmethylated CpG dinucleotides mark silent tissue-specific enhancers in embryonic stem cells. Proc. Natl. Acad. Sci. U.S.A. 104, 12377–12382
12. Zaret, K. S., and Carroll, J. S. (2011) Pioneer transcription factors: establishing competence for gene expression. Genes Dev. 25, 2227–2241
13. Besnard, V., Wirt, S. E., Hull, W. M., and Whitsett, J. A. (2004) Immunohistochemical localization of Foxa1 and Foxa2 in mouse embryos and adult tissues. Gene Expr. Patterns 5, 193–208
14. Lee, C. S., Sund, N. J., VATMANIKI, M. Z., Matsuhashi, F. M., Stoffers, D. A., and Kaestner, K. H. (2002) Foxa2 controls Pdx1 gene expression in pancreatic beta-cells in vivo. Diabetes 51, 2546–2551
15. Nishizaki, Y., Shimazu, K., Kondoh, H., and Sasaki, H. (2001) Identification of essential sequence motifs in the node/notochord enhancer of Foxa2 (Hnf3beta) gene that are conserved across vertebrate species. Mech. Dev. 102, 57–66
16. Miyazaki, J., Araki, K., Yamato, E., Ikenaga, H., Asano, T., Shibasaki, Y., Oka, Y., and Yamamura, K. (1990) Establishment of a pancreatic beta cell line that retains glucose-inducible insulin secretion: special reference to expression of glucose transporter isoforms. Endocrinology 127, 126–132
17. Sharivkin, R., Walker, M. D., and Soen, Y. (2012) Proteomics-based dissection of human endoderm progenitors by differential cell capture on antibody array. Mol. Cell. Proteomics 11, 586–595
18. Kroon, E., Martinson, L. A., Kadoya, K., Bang, A. G., Kelly, O. G., Eliazer, S., Young, H., Richardson, M., Smart, N. G., Cunningham, J., Agulnick, A. D., D’Amour, K. A., Carpenter, M. K., and Baetge, E. E. (2008) Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. Nat. Biotechnol. 26, 443–452
19. Bahar Halpern, K., Veprik, A., Rubins, N., Naaman, O., and Walker, M. D. (2012) GPR41 gene expression is mediated by internal ribosome entry site (IRES)-dependent translation of bicistronic mRNA encoding GPR40 and GPR41 proteins. J. Biol. Chem. 287, 20154–20163
20. Straussman, R., Nejman, D., Roberts, D., Steinfeld, I., Blum, B., Benvenisty, N., Simon, I., Yakhini, Z., and Cedar, H. (2009) Developmental programming of CpG island methylation profiles in the human genome. Nat. Struct. Mol. Biol. 16, 564–571
21. Pollack, Y., Stein, R., Razin, A., and Cedar, H. (1980) Methyltransferase of foreign DNA sequences in eukaryotic cells. J. Biol. Chem. 255, 265–276
22. Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H., and Helin, K. (2006) Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev. 20, 1123–1136
23. Delcuve, G. P., Rastegar, M., and Dave, J. R. (2009) Epigenetic control.

DNA Methylation Activates FoxA2 Gene

AUGUST 22, 2014•VOLUME 289•NUMBER 34

JOURNAL OF BIOLOGICAL CHEMISTRY 23891
DNA Methylation Activates FoxA2 Gene

J. Cell. Physiol. 219, 243–250

26. Miranda, T. B., and Jones, P. A. (2007) DNA methylation: the nuts and bolts of repression. J. Cell. Physiol. 213, 384–390

27. Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J. M., Bollen, M., Esteller, M., Di Croce, L., de Launoit, Y., and Fuks, F. (2006) The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439, 871–874

28. Wu, H., Coskun, V., Tao, J., Xie, W., Ge, W., Yoshikawa, K., Li, E., Zhang, Y., and Sun, Y. E. (2010) Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. Science 329, 444–448

29. D’Amour, K. A., Agulnick, A. D., Eliazer, S., Kelly, O. G., Kroon, E., and Baetge, E. E. (2005) Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat. Biotechnol. 23, 1534–1541

30. Larsen, F., Gundersen, G., Lopez, R., and Prydz, H. (1992) CpG islands as gene markers in the human genome. Genomics 13, 1095–1107

31. Jones, P. A., and Laird, P. W. (1999) Cancer epigenetics comes of age. Nat. Genet. 21, 163–167

32. Lienert, F., Wirbelauer, C., Som, I., Dean, A., Mohn, F., and Schübeler, D. (2011) Identification of genetic elements that autonomously determine DNA methylation states. Nat. Genet. 43, 1091–1097

33. Brinkman, A. B., Gu, H., Bartels, S. J., Zhang, Y., Matarese, F., Simmer, F., Marks, H., Bock, C., Gnerke, A., Meissner, A., and Stunnenberg, H. G. (2012) Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res. 22, 1128–1138

34. Statham, A. L., Robinson, M. D., Song, J. Z., Coolen, M. W., Stirzaker, C., and Clark, S. J. (2012) Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. Genome Res. 22, 1120–1127

35. Jaenisch, R., and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet. 33, 245–254

36. Eden, S., Constancia, M., Hashimshony, T., Dean, W., Goldstein, B., Johnson, A. C., Keshet, I., Reik, W., and Cedar, H. (2001) An upstream repressor element plays a role in Igf2 imprinting. EMBO J. 20, 3518–3525

37. He, Y., Vogelstein, B., Velculescu, V. E., Papadopoulos, N., and Kinzler, K. W. (2008) The antisense transcriptomes of human cells. Science 322, 1855–1857

38. Zhao, X. D., Han, X., Chew, J. L., Liu, J., Chiu, K. P., Choo, A., Orlov, Y. L., Sung, W. K., Shahab, A., Kuznetsov, V. A., Bourque, G., Oh, S., Ruan, Y., Ng, H. H., and Wei, C. L. (2007) Whole-genome mapping of histone H3 lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell 1, 286–298

39. Hantusch, B., Kalt, R., Krieger, S., Puri, C., and Kerjaschki, D. (2007) Sp1/Sp3 and DNA-methylation contribute to basal transcriptional activation of human podoplanin in MG63 versus Saos-2 osteoblastic cells. BMC Mol. Biol. 8, 20

40. Niesen, M. I., Osborne, A. R., Yang, H., Rastogi, S., Chellappan, S., Cheng, J. Q., Boss, J. M., and Blanck, G. (2005) Activation of a methylated promoter mediated by a sequence-specific DNA-binding protein, RFX. J. Biol. Chem. 280, 38914–38922

41. Constância, M., Pickard, B., Kelsey, G., and Reik, W. (1998) Imprinting mechanisms. Genome Res. 8, 881–900

42. Bhandare, R., Schug, J., Le Lay, J., Fox, A., Smirnova, O., Liu, C., Naji, A., and Kaestner, K. H. (2010) Genome-wide analysis of histone modifications in human pancreatic islets. Genome Res. 20, 428–433

43. Kunarso, G., Chia, N. Y., Jeyakani, J., Hwang, C., Lu, X., Chan, Y. S., Ng, H. H., and Bourque, G. (2010) Transposable elements have rewired the core regulatory network of human embryonic stem cells. Nat. Genet. 42, 631–634

44. Varghese, S., Rabkin, S. D., Nielsen, G. P., MacGarvey, U., Liu, R., and Martuza, R. L. (2007) Systemic therapy of spontaneous prostate cancer in transgenic mice with oncolytic herpes simplex viruses. Cancer Res. 67, 9371–9379

45. Zaret, K. S., and Grompe, M. (2008) Generation and regeneration of cells of the liver and pancreas. Science 322, 1490–1494