CpG Sites Preferentially Methylated by Dnmt3a in Vivo

Received for publication, October 12, 2005, and in revised form, January 25, 2006 Published, JBC Papers in Press, January 26, 2006, DOI 10.1074/jbc.M511100200

Masahiro Oka, Nemanja Rodić, Jamie Graddy, Lung-Ji Chang, and Naohiro Terada

From the Department of Pathology, Department of Molecular Genetics and Microbiology, University of Florida College of Medicine, Gainesville, Florida 32610

Dnmt3a and Dnmt3b are two major de novo DNA methyltransferases essential for embryonic development in mammals. It has been shown that Dnmt3a and Dnmt3b have distinct substrate preferences for certain genomic loci, including major and minor satellite repeats. However, the exact target CpG sites where Dnmt3a and Dnmt3b catalyze DNA methylation remains largely unknown. To identify a CpG site that is specifically methylated by Dnmt3a or Dnmt3b, we screened methylated genomic loci by methylation sensitive restriction fingerprinting using genomic DNA from wild-type, Dnmt3a null, Dnmt3b null, and Dnmt3a-Dnmt3b double null ES cells. Interestingly, one of the CpG sites was preferentially methylated in wild-type and Dnmt3b null ES cells but not in Dnmt3a null or Dnmt3a-Dnmt3b double null ES cells, suggesting that the site-specific methylation was Dnmt3a-dependent. Sequencing results demonstrated that purified Dnmt3a protein has a DNA sequence preference regardless of its chromosomal location or structure. In fact, it has been shown that Dnmt3a null mice and Dnmt3b null mice in germ cell development when Dnmt3b was conditionally disrupted resulted in a partial loss of DNA methylation on endogenous C-type retroviral DNA in mouse embryonic fibroblasts (8). These results suggest that Dnmt3a and Dnmt3b have different target sites for DNA methylation depending on the cell types and the stage of development.

It is known that the minor satellite DNA regions of the genome, which consist of 50,000–100,000 of DNA repeats at the centromeric region, are specifically methylated by Dnmt3b but not by Dnmt3a (6, 9). In contrast, major satellite repeats, located in the pericentromeric region, are preferentially methylated by Dnmt3a (9). Furthermore, a CpG site in the Xist promoter region is revealed as a Dnmt3a-specific site of DNA methylation (9). In addition, Hsieh (10) found that the cloned ERNA region of the Epstein-Barr virus genome on the episome was a better substrate for Dnmt3a than the control luciferase coding region, suggesting that there is a substrate preference for the enzyme regardless of its chromosomal location or structure. In fact, it has been demonstrated that purified Dnmt3a protein has a DNA sequence preference when either a plasmid or an oligonucleotide is used as the methylation substrate in vitro (11, 12).

In the present study, we attempted to identify the genomic DNA loci that are preferentially methylated by Dnmt3a or Dnmt3b in vivo. This was initially accomplished using methylation-sensitive restriction fingerprinting (MSRF), which has been successfully utilized for the identification of differentially methylated genes in cancer cells (13) or during ES cell differentiation (14). Using the MSRF method, we identified a CpG site in the G isoform of the fibroblast growth factor 1 (Fgf-1) gene locus that is preferentially methylated by Dnmt3a. This paper will further discuss the prospect that the flanking DNA sequence surrounding the CpG sites may play a role in this selective methylation by Dnmt3a.

EXPERIMENTAL PROCEDURES

In Vitro ES Cell Differentiation—ES cells (J1, Dnmt3a null, Dnmt3b null, Dnmt3a-Dnmt3b null) were maintained on gelatin-coated dishes in ES cell maintenance medium containing 1000 units/ml recombinant mouse leukemia inhibitory factor (LIF) (ESGRO; Chemicon International) as described previously (15). ES cells were differentiated using a hanging drop method as described (15). After 48 h, embryoid bodies (EBs) were collected and attached to tissue culture plates. Medium was then replaced every other day.

MSRF—The MSRF method was performed as described previously (14). Briefly, ES cells and differentiated EBs were harvested, and

In mammals, global DNA methylation is catalyzed mainly by three DNA methyltransferases, Dnmt1, Dnmt3a, and Dnmt3b. Dnmt1 has a high preference for hemimethylated DNA and is essential for maintaining methylation patterns during DNA replication (1, 2). On the other hand, Dnmt3a and Dnmt3b have a preference for both unmethylated and hemimethylated DNA (3, 4) and are responsible for de novo methylation during early development. Although Dnmt3a and Dnmt3b contain highly conserved amino acid sequences in their PWWP domains, cysteine-rich domains, and carboxyl-terminal catalytic domains, the overall homology of these proteins is relatively low (3, 5). Gene disruption studies revealed that Dnmt3a null mice and Dnmt3b null mice showed different developmental defects (6). Conditional knock-outs of Dnmt3a or Dnmt3b in germ cells also exhibited distinct phenotypes (7).

Dnmt3a disruption in male germ cells showed impaired spermatogenesis, while offspring from the Dnmt3a conditional mutant females died in utero. In both cases, Dnmt3a null cells displayed a lack of methylation at imprinting loci (7). In contrast, no apparent phenotype was observed in germ cell development when Dnmt3b was conditionally disrupted (7). In in vitro cell culture, a disruption of Dnmt3b, but not Dnmt3a, resulted in a partial loss of DNA methylation on endogenous C-type retroviral DNA in mouse embryonic fibroblasts (8). These results suggest that Dnmt3a and Dnmt3b have different target sites for DNA methylation depending on the cell types and the stage of development.

It is known that the minor satellite DNA regions of the genome, which consist of 50,000–100,000 of DNA repeats at the centromeric region, are specifically methylated by Dnmt3b but not by Dnmt3a (6, 9). However, the exact target CpG sites where Dnmt3a and Dnmt3b have a preference for both unmethylated and hemimethylated DNA and is essential for maintaining methylation patterns during DNA replication (1, 2). On the other hand, Dnmt3a and Dnmt3b have a preference for both unmethylated and hemimethylated DNA (3, 4) and are responsible for de novo methylation during early development. Although Dnmt3a and Dnmt3b contain highly conserved amino acid sequences in their PWWP domains, cysteine-rich domains, and carboxyl-terminal catalytic domains, the overall homology of these proteins is relatively low (3, 5). Gene disruption studies revealed that Dnmt3a null mice and Dnmt3b null mice showed different developmental defects (6). Conditional knock-outs of Dnmt3a or Dnmt3b in germ cells also exhibited distinct phenotypes (7).
genomic DNA was extracted using the Wizard genomic DNA purification kit (Promega). Extracted DNA was digested with MseI alone (New England Biolabs), or in combination with BstUI (New England Biolabs), at 10 units per 1 μg of DNA for each enzyme. The PCR reaction was performed in a 20-μl reaction mixture containing 100 ng of digested genomic DNA, 0.4 μM primers, 1.25 units of HotMasterTaq DNA polymerase (Eppendorf), 200 μM dNTP, 1 μCi/μl [α-32P]dCTP (3000 Ci/mmol), and 1X reaction buffer. The primers used in this study were as follows: Bs-1, 5’-AGCGGCGCGG-3’; Bs-2, 5’-GCCGGGCGCGA-3’; Bs-3, 5’-CGGGGGCGCGA-3’; Bs-4, 5’-ACCCACCCCGC-3’. The PCR reaction consisted of an initial denaturing step for 5 min at 94 °C and 30 cycles of the following temperature and times: 94 °C for 2 min, 40 °C for 1 min, and 72 °C for 2 min. As a final step, the reaction was incubated at 72 °C for 8 min. Each sample was then separated by electrophoresis on a 5% polyacrylamide gel. Wet gels were laid on Whatman 3MM filter paper, and exposed to x-ray film (Eastman Kodak Co.) at −80 °C. After developing the film, the DNA bands of interest were excised from the polyacrylamide gel and eluted by incubation of the gel fragments in 50 μl of sterile deionized water for 10 min at 100 °C. Eluted DNA was re-amplified by the identical primers and PCR conditions. Re-amplified DNA fragments were excised from the gel, ligated into a TA-cloning vector using pCRII-Topo cloning kit (Invitrogen), and sequenced. Using Bs-1 and Bs-4 primers, we identified the band that is reduced in Dnmt3a null ES cells compared with wild-type and Dnmt3b null ES cells. Sequencing results showed that the fragment contains Fgf-1.G 1st exon sequence.

**Bisulfite Sequencing and Combined Bisulfite Restriction Analysis (COBRA)**—DNA was extracted from ES cells, EBs, and adult tissues using the DNA Wizard genomic DNA purification kit (Promega). A bisulfite reaction was performed using the EZ DNA methylation kit (Zymo Research). Two μg of genomic DNA were used for conversion with bisulfite. Approximately 80 ng of bisulfite-converted DNA were used as a template for each PCR analysis. Primers used for COBRA analysis and bisulfite sequencing were as follows: for Fgf-1.G, 5’-TAGTGCTTTGTTAGAGGATAT-3’ and 5’-ACCACACACACCAATCTCAGATAT-3’; for Fgf-1.B, 5’-TATATTTTTTGAGTGTTAGTATGTG-3’ and 5’-AAAAATCTCTCTCTCATCTCCTT-3’; for Fgf-1.A, 5’-GGGTTGTTGGTAGTTGGAGGATAT-3’ and 5’-TATTGCTAAAAAACAAAAAAA-3’. For bisulfite sequencing, the PCR fragments were cloned into pCRII-TOPO cloning vector (Invitrogen), and individual clones were sequenced.

**Lentiviral Vector Construction and Infection**—Lentiviral vector, pTYF-EF-FLAG-Dnmt3a2, and FLAG-Dnmt3b1 that contains Dnmt3 under the control of the EF1-α promoter were described previously (15). To express the Dnmt3a-Dnmt3b chimeric protein, the plasmid was made as follows. Briefly, the amino terminus coding region of Dnmt3a (amino acid sequence 1–393) was amplified by PCR using forward and reverse primers containing BamHI and EcoRI sites, respectively. The carboxy terminus region of Dnmt3b (amino acid sequence 544–893) was PCR amplified using primers contain EcoRI and Spel, respectively. The amplified fragments were digested with the restriction enzyme and gel-purified, and both of the fragments were inserted into the BamHI-Spel site of pTYF-EF-FLAG vector (15, 16) to create an amino-terminally FLAG-tagged Dnmt3a-3b chimeric protein expression vector. A Dnmt3a-b3 chimeric protein expression vector was similarly made using the amino-terminal Dnmt3b fragment (amino acid sequence 1–543) and the carboxy-terminal fragment of Dnmt3a (amino acid sequence 394–689). The constructed plasmid was packaged into lentivirus by co-transfection with P(N) (a helper plasmid) and pHEF-VSVG (an envelope coding plasmid) (17). Transduction of the viral vectors was performed as described previously (17). Expression of the FLAG-tagged proteins was confirmed by both immunoblotting and immunostaining.

**Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)**—Total RNA was isolated from adult mouse tissues using TRizol reagent (Invitrogen). Two μg of RNA were used as a template for RT reactions using SuperScript synthesis for the first strand synthesis kit (Invitrogen). Sequences of forward and reverse primer pairs were as follows: Fgf-1.G (5’-CTAGGAAGTAAGGACGGATT-3’ and 3’-ACAGCTCCCGTC-TTCTTG-3’), Fgf-1.B (5’-GCTCTGTTGTAATAGGATG-3’ and 3’-ACAGCTCCGGTTCTCTTG-3’), Fgf-1.A (5’-GCGAAGAGGCCACCTGTAAA-3’ and 5’-ACAGCTCCGGTTCTCTTG-3’).

**Southern Blotting**—DNA from wild-type ES cells, Dnmt3a-Dnmt3b null ES cells, and lentivirus-infected Dnmt3a-Dnmt3b null ES cells were isolated using the DNA Wizard genomic DNA purification kit (Promega). Fifteen μg of DNA were digested with HpyCH4IV or HpaII (New England Biolabs), separated on 1% agarose gel, and analyzed by Southern hybridization. Probes were amplified by PCR using the following primers: for major satellite, 5’-CTGATAGGGTGAATGGATG-3’ and 5’-CCGTTAGGTATGTTGGAGGATAT-3’; for minor satellite, 5’-TCTGAGGTTAGGGTGAATGGATG-3’ and 5’-CCGTTAGGTATGTTGGAGGATAT-3’. Amplified DNA was cloned into pCRII-TOPO cloning vector (Invitrogen) and verified by sequencing.

**RESULTS**

**Identification of a CpG-rich DNA Sequence That Undergoes de Novo DNA Methylation by Dnmt3a during ES Cell Differentiation**—To isolate a DNA locus that is specifically methylated by Dnmt3a or Dnmt3b, we utilized MSRF using genomic DNA from wild-type, Dnmt3a null, Dnmt3b null, and Dnmt3a-Dnmt3b double null ES cells as the substrate. Isolated DNA was digested by a combination of methylation-sensitive (BstUI) and -insensitive (MseI) restriction enzymes and amplified by PCR. Using various combinations of four different 10-mer primers, we identified a band that is amplified less in Dnmt3a null ES cells and Dnmt3a-Dnmt3b double null ES cells (Fig. 1A). The bands were excised from the gels, and the sequences were determined. Nine out of 14 clones had the same sequence, which mapped to the first exon of the Fgf-1.G isoform on mouse chromosome 18 (Chr18:39090252–39090501) and contained the BstUI site (Fig. 1B).

We confirmed the results using COBRA (Fig. 2A). Bisulfite treatment converts unmethylated cytosines to thymidine residues. Therefore, both CpG sites within the BstUI recognition site (CGCG) must be methylated to be digested by BstUI in COBRA. As a template for the bisulfite-PCR, we used genomic DNA from undifferentiated ES cells or from differentiated EBs (for 2 or 4 days). Differentiation of ES cells was confirmed by monitoring the expression of marker genes, as described previously (Ref. 18 and data not shown). It has been reported that the global level of DNA methylation increased after ES cell differentiation (19), which supposedly reflects the wave of de novo methylation during early embryogenesis (20). Digestion of the bisulfite-treated PCR fragment with BstUI showed that Dnmt3a null and Dnmt3a-Dnmt3b double null ES cells had only slight DNA methylation in the Fgf-1.G-BstUI locus when they were either undifferentiated or differentiated. In contrast, wild-type and Dnmt3b null ES cells showed substantial DNA methylation when they were undifferentiated, and their methylation level was further increased after cell differentiation. We further examined more detailed DNA methylation patterns by bisulfite sequencing (Fig. 2B). As expected from the COBRA results, two CpG sequences...
within the BstUI site in the first exon of Fgf-1. G (±276 and ±278) were methylated in wild-type cells. Similarly, both CpG sites in the BstUI site were highly methylated in Dnmt3b null ES cells. In contrast, a significant decrease in methylation of CpG within the BstUI site was observed in Dnmt3a null ES cells. Interestingly, the second CpG site within the BstUI site (±278) was more dominantly hypomethylated (87.5% of clones) than the first CpG (±276, 50% of clones) in Dnmt3a null ES cells. In COBRA, BstUI could digest bisulfite-treated PCR products only
if both of the CpG sites are methylated. Therefore, only 6% of the fragments for Dnmt3a null ES cells are expected to be digested with BstUI, which appears consistent with the COBRA data above (Fig. 2A). In Dnmt3a-Dnmt3b double null ES cells, CpG sites were totally unmethylated. We then analyzed the CpG sites neighboring the BstUI site. As shown in Fig. 2B, two CpG sites upstream of the BstUI site (+232 and +247) were highly methylated in wild-type cells. These upstream CpG sites were relatively highly methylated in Dnmt3a null ES cells. These results suggest that some CpG sites within the Fgf-1 gene are preferentially methylated by Dnmt3a in vivo.

DNA Methylation in Other Isoforms of Fgf-1 Gene during ES Cell Differentiation—Since the murine Fgf-1 gene has multiple isoforms resulting from different transcription initiation sites (21), we analyzed CpG methylation within other Fgf-1 gene isoforms using COBRA (Fig. 1). The Fgf-1.A isoform is predominantly expressed in the heart, and the Fgf-1.B isoform is mainly expressed in the brain (21). A HpyCH4IV site within the 5′-upstream region of Fgf-1.A isoform (Fgf-1.A-HpyCH4IV: Chr18:39153367–39153370) and a HinfI site within the first exon of Fgf-1.B isoform (Fgf-1.B-HinfI: Chr18:39142391–39142395) were identified as sites suitable for COBRA. Interestingly, the Fgf-1.A-HpyCH4IV site was also hypomethylated in Dnmt3a null ES cells as well as the Fgf-1.G-BstUI site. In contrast, methylation of the Fgf-1.B-HinfI site was slightly higher in Dnmt3a null ES cells when compared with Dnmt3b null ES cells (Fig. 2A).

Fgf-1 DNA Methylation Patterns in Adult Tissues—We also examined the gene expression and DNA methylation status of Fgf-1 isoforms in mouse adult tissues (Fig. 3, A and B), since DNA methylation is generally associated with gene inactivation. The RT-PCR results revealed that the Fgf-1.G gene was highly expressed in liver and moderately expressed in kidney, as reported previously (22).
analysis by COBRA showed that the Fgf-1.G-BstUI site was hypomethylated in liver, where Fgf-1.G gene expression was high. In kidney, the Fgf-1.G-BstUI site methylation was lower than in other tissues that do not express Fgf-1.G. Bisulfite sequencing further confirmed the COBRA data (Fig. 3C). Overall, there was a tendency for DNA methylation levels to be higher in tissues where the Fgf-1.G gene was not expressed. Similarly, the Fgf-1.A-HpyCH4IV site was hypomethylated in heart and kidney, where the Fgf-1.A gene was expressed. The correlation was less clear with Fgf-1.B. The Fgf-1.B-Hinfl locus was highly methylated in testis and only slightly methylated in other somatic organs.

Reconstitution of DNA Methylation by Ectopic Expression of Dnmt3a—To further demonstrate that Dnmt3a is responsible for site-specific DNA methylation within the Fgf-1 gene, we reconstituted the expression of Dnmt3a in Dnmt3a-Dnmt3b double null ES cells and looked for a recovery in methylation (Fig. 4). Overall, there was a tendency for DNA methylation levels to be higher in tissues where the Fgf-1.G gene was not expressed. Similarly, the Fgf-1.A-HpyCH4IV site was hypomethylated in heart and kidney, where the Fgf-1.A gene was expressed. The correlation was less clear with Fgf-1.B. The Fgf-1.B-Hinfl locus was highly methylated in testis and only slightly methylated in other somatic organs.

The Regulatory Domain of Dnmt3a Is Important for the Site-specific DNA Methylation—Dnmt3a and Dnmt3b each contain a highly conserved catalytic domain at their carboxyl termini. In contrast, the amino-terminal halves of the proteins are more divergent and show low homology. To examine which domain is more important to the substrate specificity, we made a chimeric construct of Dnmt3a and Dnmt3b, namely, Dnmt3a-3b, which contains the Dnmt3a regulatory domain with the Dnmt3b catalytic domain, and Dnmt3b-3a, which contains the Dnmt3b regulatory domain with the Dnmt3a catalytic domain (Fig. 5A). These chimeric genes were ectopically expressed in Dnmt3a-Dnmt3b double null ES cells using lentiviral vectors, and their DNA methylation activity was assayed. As shown in Fig. 5B, DNA methylation in major satellite DNA repeats was preferentially recovered by ectopic expression of Dnmt3a as reported previously (9). Dnmt3a-3b, but not Dnmt3b-3a, partially recovered DNA methylation in major satellite DNA. Therefore, the regulatory domain of Dnmt3a is more important than the catalytic domain to determine the substrate specificity of Dnmt3a to major satellite repeats. Using these constructs, we examined the DNA methylation of Fgf-1 sites (Fig. 5C). The Fgf-1.A-HpyCH4IV site was methylated by the transduction of Dnmt3a-3b in Dnmt3a-Dnmt3b double null ES cells regardless of differentiation status. In contrast, Dnmt3b-3a expression did not restore DNA methylation. These results suggest that the regulatory domain, rather than the catalytic domain, of Dnmt3a is important for the site-specific DNA methylation at the Fgf-1.A site. It should be noted that methylation of the Fgf-1.G-BstUI site was not observed by the expression of either one of the chimeric constructs. Fgf-1.B-Hinfl site was similarly methylated by Dnmt3a, Dnmt3b, and Dnmt3a-3b (data not shown).

DISCUSSION

Using MSRF, we have identified a BstUI recognition site within the first exon of the Fgf-1.G gene locus as a CpG site that is methylated preferentially by Dnmt3a. Similarly, we found that a HpyCH4IV site in
the 5'-flanking region of the Fgf-1.A gene was predominantly methylated by Dnmt3a. Furthermore, reconstitution of Dnmt3a in double null ES cells restored the methylation of these sites, whereas the expression of Dnmt3b did not. These results demonstrated that these CpG sites in the Fgf-1 gene are specific target sites for Dnmt3a.

Genome-wide methylation analysis by restriction landmark genomic scanning, using the NotI restriction enzyme, suggested that methylation patterns at CpG islands are quite similar between wild-type, Dnmt3a null, and Dnmt3b null ES cells (23). However, by MSRF screening, we identified the genomic locus that is specifically methylated by Dnmt3a.

FIGURE 5. The regulatory domain of Dnmt3a is important for site-specific DNA methylation. A, schematic illustration of chimeric molecules of Dnmt3a and Dnmt3b. Dnmt3a-3b contains the regulatory domain of Dnmt3a and the catalytic domain of Dnmt3b, while Dnmt3b-3a makes use of the regulatory domain of Dnmt3b and the catalytic domain of Dnmt3a. B, DNA methylation analysis for major satellite DNA repeats. Dnmt3a-Dnmt3b double null ES cells were infected with Dnmt3a, Dnmt3b, Dnmt3a-3b, or Dnmt3b-3a lentivirus as indicated. After 4 days, genomic DNA was isolated, digested with HpyCH4IV, and hybridized to the DNA probe for major satellite repeats. Note that methylation of major satellite repeats were selectively restored by Dnmt3a or Dnmt3a-3b expression. C, DNA methylation of the Fgf-1.A site by chimeric Dnmt3. After 2 days of lentiviral infection in the double null ES cells, they were maintained undifferentiated (Undiff.) for additional 4 days or differentiated as embryoid bodies for 4 days (Diff.). Genomic DNA was isolated and used for COBRA for the Fgf-1.A methylation site. Note that methylation of the Fgf-1.A COBRA site was selectively restored by Dnmt3a or Dnmt3a-3b expression. U, unmethylated DNA derived fragment; M, methylated DNA derived fragment.
CpG Sites Methylated by Dnmt3a

Even though genome-wide screening has a clear merit, restriction landmark genomic scanning restricts its analysis for the CpG sites located within the recognition site of the restriction enzyme used. In contrast, MSRF screening could be useful for the analysis of CpG methylation, regardless of flanking sequences.

It remains undetermined how Dnmt3a catalyzes DNA methylation of the specific target sites in the genome. One hypothesis is that Dnmt3a binds to the sequence-specific DNA binding molecules, which brings the complex to the specific genomic locus. In fact, Dnmt3a has been shown to bind to a sequence-specific DNA-binding protein (24–26) and act as a co-repressor (25, 26), although it has not been demonstrated whether Dnmt3a did actually catalyze DNA methylation at this specific site.

The other hypothesis is that Dnmt3a itself possesses a substrate DNA sequence preference. It has been reported that the cloned EBNA region of the Epstein-Barr virus genome on the episome was a better substrate in vivo for Dnmt3a than a control luciferase coding region on the same episome (10). Handa et al. (11) reported that there is a preference for the flanking sequence surrounding methylated CpG sites for Dnmt3a when they analyzed DNA methylation in vitro using a purified Dnmt3a and oligonucleotide DNA substrates. According to the study, the RCGY consensus sequence was a more preferred substrate for methylation by Dnmt3a. In contrast, YCGR was not a preferred substrate for Dnmt3a. It should be noted that, however, the study has not demonstrated a differential sequence preference between Dnmt3a and Dnmt3b. In our study, all the sites preferentially methylated by Dnmt3a, namely, two CpG sites within Fgf-1-G-BstUI site and one at the Fgf-1-A-HpyCH4IV site, contain RCGY sequences. Furthermore, the Fgf-1-B-Hinfl site, which was not selectively methylated by Dnmt3a, contains the YCGR sequence. Thus, our results support the previous in vitro studies that Dnmt3a may possess a substrate preference to RCGY even in vivo. Of interest, the Fgf-1-B-Hinfl YCGR site was more highly methylated in Dnmt3a null ES cells than in Dnmt3b null ES cells (Fig. 2A). Furthermore, Dnmt3b reconstitution caused a slightly higher degree of DNA methylation at the Fgf-1-B-Hinfl YCGR site than Dnmt3a (Fig. 4A). Therefore, the YCGR site might be a preferred substrate for Dnmt3b in vivo. In fact, Dnmt3b is known to be highly expressed in adult testis (3, 5), where the methylation level of Fgf-1-B-Hinfl site was very high (Fig. 3B). The major satellite repeats selectively methylated by Dnmt3a (Fig. 5) include ACCT (HpyCH4IV or Mael recognition site), which consists of a RCGY sequence. The minor satellite repeats, which are known to be selectively methylated by Dnmt3b, on the other hand, contain a YCGR sequence. Taken together, these data suggest that the substrate preference of Dnmt3a or Dnmt3b may play an important role in determining which genomic loci are differentially methylated in vivo.

Dnmt3a and Dnmt3b each consist of an NH2-terminal regulatory domain, which contains the PWPP domain, a cysteine-rich domain, and a COOH-terminal catalytic domain (3, 5). The catalytic domain of either Dnmt3a or Dnmt3b by itself has been shown to possess a similar substrate preference as a full-length Dnmt3a in vitro (11). Therefore, the catalytic domain may not be critical to determine the substrate specificity of Dnmt3a and Dnmt3b in vivo. Indeed, our experiments using chimeric Dnmt3 constructs between Dnmt3a and Dnmt3b revealed that the amino-terminal regulatory domain of Dnmt3a was more important for the site-specific methylation of major satellite repeats and the Fgf-1-A-HpyCH4IV locus (Fig. 5, B and C). The PWPP domain has been shown to have a high affinity for DNA (27–29) and is required for the methylation of the major satellite repeats (29). However, considering the non-sequence-specific binding property of the Dnmt3 PWPP domain (29), it may be difficult to assume that this domain is directly involved in enzyme targeting to specific DNA loci, such as the Fgf-1 gene.

FGF-1 is a member of the fibroblast growth factor family, which possesses broad mitogenic activities, and is involved in a variety of biological processes, including cell proliferation, migration, differentiation, and survival (30–35). The Fgf-1 gene contains multiple promoters associated with distinct first exons, which result in multiple transcript isoforms containing common sequences except for the first exons (36). These promoters are regulated in a tissue-dependent manner both in mouse and human (21). Since ectopic expression of Fgf-1 induces intimal hyperplasia in vivo (37), their strict gene regulation appears to be important in normal embryonic development and tissue homeostasis. In this paper, we also showed that the degree of DNA methylation inversely correlates with their expression pattern in adult mouse tissues. DNA methylation may be one of the mechanisms involved in the silencing of gene expression by preventing the access of transcription factors. Our data imply that a particular Dnmt3 member might play a role in the regulation of isoform determination of the Fgf-1 gene during embryonic development.

In conclusion, we have identified CpG sites preferentially methylated by Dnmt3a in vivo. Although the exact molecular mechanisms underlying this selective methylation remain undetermined, the regulatory domain of Dnmt3a likely plays a predominant role. Further, common flanking sequences (RCGY) around CpG may be involved in this selective methylation by Dnmt3a in vivo. Site-specific methylation by a specific member of Dnmt3 may play a role in the orchestration of gene expression during embryonic development.

Acknowledgments—We thank En Li for providing us with Dnmt null ES cells and Keith Robertson, Amy Meacham, Sarah Kehoe, Jeffrey Brower, and Catherine Ketcham for helpful discussions and critical reading of the manuscript.

REFERENCES

1. Leonardth, H., Page, A. W., Weier, H. U., and Bestor, T. H. (1992) Cell 71, 865–873
2. Pradhan, S., Bacolla, A., Wells, R. D., and Roberts, R. J. (1999) J. Biol. Chem. 274, 33002–33010
3. Okano, M., Xie, S., and Li, E. (1998) Nat. Genet. 19, 219–220
4. Aoki, A., Suetake, I., Miyagawa, J., Fujio, T., Chijiwai, T., Sasaki, H., and Tajima, S. (2001) Nucleic Acids Res. 29, 3506–3512
5. Xie, S., Wang, Z., Okano, M., Nogami, M., Li, Y., He, W. W., Okumura, K., and Li, E. (1999) Gene (Amst.) 236, 87–95
6. Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) Cell 99, 247–257
7. Kaneda, M., Okano, M., Hata, K., Sato, T., Tsujimoto, N., Li, E., and Sasaki, H. (2004) Nature 429, 900–903
8. Dodge, I. E., Okano, M., Dick, F., Tsujimoto, N., Chen, T., Wang, S., Ueda, Y., Dyson, N., and Li, E. (2005) J. Biol. Chem. 280, 17986–17991
9. Chen, T., Ueda, Y., Dodge, I. E., Wang, Z., and Li, E. (2003) Mol. Cell. Biol. 23, 5594–5605
10. Hsieh, C. I. (1999) Mol. Cell. Biol. 19, 8211–8218
11. Handa, Y., and Jeltsch, A. (2005) J. Mol. Biol. 348, 1103–1112
12. Lin, I. G., Han, L., Taghva, A., O’Brien, L. E., and Hsieh, C. I. (2002) Mol. Cell. Biol. 22, 704–723
13. Huang, T. H., Laux, D. E., Hamlins, B. C., Tran, P., Tran, H., and Lubahn, D. B. (1997) Cancer Res. 57, 1030–1034
14. Rodic, N., Oka, M., Hamazaki, T., Murawski, M. R., Jorgensen, M., Matouk, D. M., Resnick, J. L., Li, E., and Terada, N. (2005) Stem Cells 23, 1314–1323
15. Oka, M., Meacham, A. M., Hamazaki, T., Rodic, N., Chang, L. J., and Terada, N. (2005) Oncogene 24, 3091–3099
16. Zaisa, A. K., Son, S., and Chang, L. J. (2002) J. Virology 76, 7209–7219
17. Iwakuma, T., Cui, Y., and Chang, L. J. (1999) Virology 261, 120–132
18. Oka, M., Tagoku, K., Russell, T. L., Nakano, Y., Hamazaki, T., Meyer, E. M., Yokota, T., and Terada, N. (2002) Mol. Cell. Biol. 12, 1274–1281
19. Lee, J. H., Hart, S. R., and Skalnik, D. G. (2004) Genesis 38, 32–38
20. Li, E. (2002) Nat. Rev. Genet. 3, 662–673
21. Chiu, I. M., Touhalisky, K., and Baran, C. (2001) Prog. Nucleic Acids Res. Mol. Biol. 70, 155–174
22. Zhang, Y., Madiai, F., and Hackshaw, K. V. (2001) Biochim. Biophys. Acta 1521, 45–58
23. Ikeda, H., and Terada, N. (2002) Mol. Cell. Biol. 22, 5594–5605
24. Handa, Y., and Jeltsch, A. (2005) J. Mol. Biol. 348, 1103–1112
25. Lin, I. G., Han, L., Taghva, A., O’Brien, L. E., and Hsieh, C. I. (2002) Mol. Cell. Biol. 22, 704–723
CpG Sites Methylated by Dnmt3a

23. Hattori, N., Abe, T., Suzuki, M., Matsuyama, T., Yoshida, S., Li, E., and Shiota, K. (2004) Genome Res. 14, 1733–1740
24. Di Croce, L., Raker, V. A., Corsaro, M., Fazi, F., Fanelli, M., Faretta, M., Fuks, F., Lo Coco, F., Kouzarides, T., Nervi, C., Minucci, S., and Pelicci, P. G. (2002) Science 295, 1079–1082
25. Fuks, F., Burgers, W. A., Godin, N., Kasai, M., and Kouzarides, T. (2001) EMBO J. 20, 2536–2544
26. Brenner, C., Deplus, R., Didelot, C., Loriot, A., Vire, E., De Smet, C., Gutierrez, A., Danovi, D., Bernard, D., Boon, T., Pelicci, P. G., Amati, B., Kouzarides, T., de Launoit, Y., Di Croce, L., and Fuks, F. (2005) EMBO J. 24, 336–346
27. Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002) Nat. Struct. Biol. 9, 217–224
28. Ge, Y. Z., Pu, M. T., Gowher, H., Wu, H. P., Ding, J. P., Jeltsch, A., and Xu, G. L. (2004) J. Biol. Chem. 279, 25447–25454
29. Chen, T., Tsujimoto, N., and Li, E. (2004) Mol. Cell. Biol. 24, 9048–9058
30. Rydel, R. E., and Greene, L. A. (1987) J. Neurosci. 7, 3639–3653
31. Davis, J. B., and Stroobant, P. (1990) J. Cell Biol. 110, 1353–1360
32. Unsicker, K., Reichert-Preibsch, H., Schmidt, R., Pettrmann, B., Labourdette, G., and Sensenbrenner, M. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 5459–5463
33. Lipton, S. A., Wagner, J. A., Madison, R. D., and D’Amore, P. A. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2388–2392
34. Fishel, R. S., Theurani, V., Eisenberg, S. J., Shai, S. Y., Corson, M. A., Nabel, E. G., Bernstein, K. E., and Berk, B. C. (1995) J. Clin. Invest. 95, 377–387
35. Renaud, F., Oliver, L., Desset, S., Tassin, J., Romquin, N., Courtois, Y., and Laurent, M. (1994) J. Cell. Physiol. 158, 435–443
36. Myers, R. L., Payson, R. A., Chotani, M. A., Deaven, L. L., and Chiu, I. M. (1993) Oncogene 8, 341–349
37. Nabel, E. G., Yang, Z. Y., Plautz, G., Forough, R., Zhan, X., Haudenschild, C. C., Maciag, T., and Nabel, G. J. (1993) Nature 362, 844–846