# Allosteric Activator Domain of Maintenance Human DNA (Cytosine-5) Methyltransferase and Its Role in Methylation Spreading

Sriharsa Pradhan* and Pierre-Olivier Estève

*New England Biolabs, 32 Tozer Road, Beverly, Massachusetts 01915*

*Received January 29, 2003; Revised Manuscript Received March 18, 2003*

ABSTRACT: The human maintenance DNA (cytosine-5) methyltransferase (hDNMT1) consists of a large N-terminal regulatory domain fused to a catalytic C-terminal domain by randomly repeated Gly-Lys dipeptides. Several N-terminal deletion mutants of hDNMT1 were made, purified, and tested for substrate specificity. Deletion mutants lacking 121, 501, 540, or 580 amino acids from the N-terminus still functioned as DNA methyltransferases, methylated CG sequences, and preferred hemimethylated to unmethylated DNA, as did the full-length hDNMT1. Methylated DNA stimulated methylation spreading on unmethylated CpG sequences for the full-length and the 121 amino acid deletion hDNMT1 equally well but not for the mutants lacking 501, 540, or 580 amino acids, indicating the presence of an allosteric activation determinant between amino acids 121 and 501. Peptides from the N- and C-termini bound methylated DNA independently. Point mutation analysis within the allosteric region revealed that amino acids 284−287 (KKHR) were involved in methylated DNA-mediated allosteric activation. Allosteric activation was reduced in the double point mutant enzymes D25 (K284A and K285A) and D12 (H286A and R287A). Retinoblastoma gene product (Rb), a negative regulator of DNA methylation, bound to the allosteric site of hDNMT1 and inhibited methylation, suggesting Rb may regulate methylation spreading.

DNA methylation is an epigenetic process in the vertebrate genome. The epigenetic mechanism depends on the presence of 5-methylcytosine in certain CG sites and DNA (cytosine-5) methyltransferases. The methylation reaction involves the enzymatic modification of cytosine to 5-methylcytosine by the transfer of a methyl group from *S*-adenosyl-L-methionine (AdoMet)[1] to carbon 5 (*1*, *2*). During this process, the cytosine is flipped 180° out of the DNA backbone into an active site pocket of the enzyme (*3*). After completion of the methyl transfer reaction the products, methylated DNA and *S*-adenosyl-L-homocysteine (AdoHcy), are released. There are two classes of DNA (cytosine-5) methyltransferases in mammals, de novo and maintenance methyltransferases. The major maintenance methyltransferase is DNMT1 (*4*) and has several isoforms, including an oocyte-specific form that lacks the first 118 amino acids (*5*) and a splice variant known as DNMT1b (*6*). Maintenance methylation ensures the propagation of tissue-specific methylation patterns established during mammalian development. DNMT1 has a 7−21-fold preference for hemimethylated DNA (*7*). The de novo methyltransferase in mammals has two isoforms, DNMT3a and DNMT3b (*8*). DNMT3a and DNMT3b have an equal preference for hemimethylated or unmethylated DNA (*8*). These enzymes are required for the wave of de novo methylation following embryo implantation and de novo methylation of newly integrated retroviral sequences in mouse ES cells lacking DNMT1 (*9*). Thus the pattern of mammalian methylation appears to be established and maintained by a set of at least three different DNA methyltransferases. Methylation of DNA in eukaryotes is also implicated in various other biological and developmental processes. The bulk of the methylation takes place during DNA replication in the S-phase of the cell cycle (*10*).

Eukaryotic maintenance DNA (cytosine-5) methyltransferase, *DNMT1*, has been cloned from various organisms such as mouse (*4*), human (*11*), *Arabidopsis* (*12*), and pea (*13*). The full-length mammalian DNMT1 is about 1620 amino acids in length with a molecular mass of 183.5 kDa (*14*). The enzyme consists of a large N-terminal regulatory region and a smaller C-terminal catalytic region linked by a run of Gly-Lys dipeptide repeats (*4*, *11*−*13*). The large N-terminal domain is unique to the eukaryotic DNA (cytosine-5) methyltransferases DNMT1, DNMT3a, and DNMT3b. The amino terminus of DNMT1 can be subdivided into two domains. One domain contains the nuclear localization signal (NLS) and binding sites for DNA methyltransferase associated protein (DNMP1; *15*) and proliferating cell nuclear antigen (PCNA; *16*). The PCNA binding region on hDNMT1 is between amino acids 162 and 174. The second region contains binding sites for replication foci (RF; *10*), Zn, and the retinoblastoma protein (Rb; *17*). Another retinoblastoma protein binding region has been identified recently at the amino-terminal region (*18*). De novo methyltransferases DNMT3a and DNMT3b also interact with DNMT1 at its amino-terminal region (*19*). The smaller C-terminal domain shows strong similarity to prokaryotic DNA (cytosine-5) methyltransferases and contains the ele-

* Corresponding author. Tel: (978) 927-5054 ext 227. Fax: (978) 921-1350. E-mail: pradhan@neb.com.

[1] Abbreviations: hDNMT1, human DNA (cytosine-5) methyltransferase; Rb, retinoblastoma; FMR-1, fragile X mental retardation syndrome; SNRPN exon 1, small nuclear riboprotein-associated peptide N; AdoMet, *S*-adenosyl-L-methionine; AdoHcy, *S*-adenosyl-L-homocysteine.

10.1021/bi034160+ CCC: $25.00 
Published on Web 04/15/2003

ments necessary for catalysis including a target recognition domain (TRD) (*20*). Thus, the *DNMT1* gene appears to have evolved from the fusion of at least three genes (*21*). There is significant sequence similarity at the amino acid level between vertebrate DNMT1 enzymes, suggesting similar biological functions. However, details of how the N-terminus functions are poorly understood.

Despite having all of the elements necessary for catalysis, the C-terminal region of DNMT1 alone is not biochemically active, perhaps due to improper protein folding (*22*), suggesting that the N-terminus and the C-terminus are both essential for activity. Previously, we have shown that the fusion of an intact N-terminal region of the mouse DNMT1 to any one of three prokaryotic DNA methyltransferases resulted in an enzyme having the substrate specificity of the parental prokaryotic protein. Steady-state kinetic analysis of one of the enzymes, DNMT1-*Hha*I, revealed a catalytic efficiency for hemimethylated DNA about 2.5-fold greater than unmethylated DNA, although M. *Hha*I alone does not exhibit such a preference, suggesting that the amino-terminal region of DNMT1 has a crucial role in the hemimethylation reaction (*23*). Because the C-terminus alone is catalytically incompetent, the question remains as to what elements are critical for the methyl transfer reaction in mammalian DNMT1. In this work, deletion derivatives of hDNMT1 were used to examine the substrate specificity and essential amino acid sequences responsible for the enzyme's substrate preference for hemimethylated DNA. GST-hDNMT1 fusion peptides were used to identify various DNA binding sites of the enzyme. An allosteric binding region of hDNMT1 was identified using catalytically competent deletion derivative and point mutant enzymes. Methylated DNA binds to the allosteric activation site, which is situated ~1000 amino acids away from the catalytic center. Methylated DNA binding to the allosteric site induces a switch from a slow to a fast reaction velocity. The role of activated hDNMT1 in de novo methylation spreading and its regulation by trans-acting protein factors, such as retinoblastoma gene product binding at the allosteric site, is discussed.

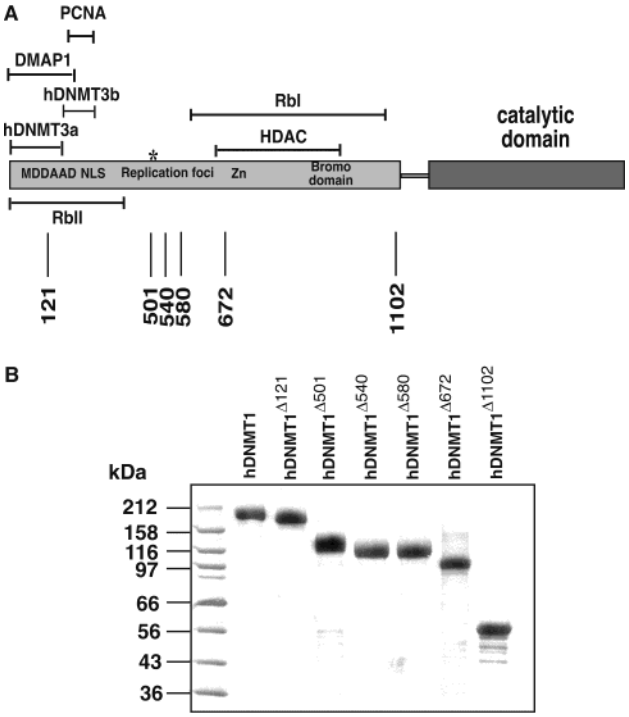


FIGURE 1: Functional organization of hDNMT1. (A) Functionally mapped regions of the full-length hDNMT1 are illustrated along with the carboxy-terminal catalytic domain. DNMT1 associated protein (DMAP1) binding (1−126), proliferative cell nuclear antigen (PCNA) binding (162−174), nuclear localization signal (NLS) (194−213), replication fork (RF) targeting peptide (320−567), and Zn binding region are represented. The repressor domain shows homology to trithorax-related proteins HRX and ALL1. Histone deacetylase (HDAC) binding (653−812) region lies in the repressor domain. Retinoblastoma gene product (Rb) binding (RbI, 416−913; RbII, 1−336) is shown. De novo methyltransferases DNMT3a and DNMT3b bind to the amino terminus of hDNMT1 (1−148 and 149−217). Phosphorylated S509 is indicated by an asterisk. Starting amino acids in the deletion mutants are indicated at the bottom. Methylated DNA-dependent allosteric activation domain (MDDAAD) is identified at the amino-terminal region of the enzyme (121−500). Amino acid numbers representing different binding domains are in parentheses. (B) Integrity and purity of the deletion hDNMTs. Mutants are indicated on the top of the gel. Protein markers in kDa are shown on the left.

## EXPERIMENTAL PROCEDURES

*Deletion hDNMT1 Mutant Transfer Vectors.* Human DNMT1 expression constructs were derived from pBK-SHMT5.0 (b) (gift of Prof. S. Baylin, Johns Hopkins University). This plasmid had the full-length hDNMT1 cDNA based on the previously published sequence (*11*). The cDNA was cloned into pVIC1 to give the transfer vector pVICHMT (*7*). Upon cotransfection pVICHMT expressed hDNMT1. All of the deletion mutants (Figure 1) were constructed using a PCR approach and cloned into pVIC1. The double point mutants D25 (H284A and R285A) and D12 (H286A and R287A) were made using mutant oligonucleotides as PCR primers. The genes were cloned into pVIC1 for protein expression (*7*).

*Insect Cell Culture, Viral Transfection, Recombinant Protein Expression, and Purification*. A pupal ovarian cell line (Sf9) from the worm *Spodoptera frugiperda* was used for cotransfection and expression of the hDNMT1 as described previously (*7*). For protein purification, infected cells [$(1-5) \times 10^8$] were resuspended in 15−30 mL of buffer M supplemented with protease inhibitor cocktail [50 mM Tris-HCl, pH 7.4, 1 mM $Na_2EDTA$, protease inhibitor cocktail containing 4-(2-aminoethyl)benzenesulfonyl fluoride, pepstatin A, E64, bestatin, leupeptin, and aprotinin (Sigma), 0.2% (v/v) per mL of cell extract, 7 $\mu$g/mL PMSF, and 500 mM NaCl]. Processing of the cells and purification of the enzymes were done as described previously (*7*).

*Expression and Purification of the hDNMT1 and Its Deletion Mutants*. Cotransfection of the hDNMT1 deletion mutant transfer vectors pVICDNMT1$^{\Delta121}$, pVICDNMT1$^{\Delta501}$, pVICDNMT1$^{\Delta540}$, pVICDNMT1$^{\Delta580}$, pVICDNMT1$^{\Delta672}$, pVLhisDNMT1$^{\Delta1102}$, with linear AcNPV DNA (BaculoGold DNA; Pharmingen) resulted in homologous recombination and integration of the polyhedrin promoter and the cDNAs carried by the construct. The optimal expression of the enzymes was 48 h postinfection as was that of the full-length DNMT1. All mutant enzymes were purified on chitin beads, were ~95% pure, and were of the expected size on a SDS−PAGE (4−12% Tris−glycine gradient) gel (*7*).

*Substrate DNA and Oligonucleotides.* All oligonucleotides were synthesized at New England Biolabs Inc. Methylated,

unmethylated, hemimethylated FMR-1, and SNRPN exon 1 locus sequences along with 5-fluoro-2′-deoxycytidine-containing DNA were described previously (*7*).

*DNA (Cytosine-5) Methyltransferase Assay and Data Analysis.* A typical reaction contained *S*-adenosyl-L-[*methyl*-$^3$H]methionine (AdoMet) (specific activity 15 Ci/mmol; Amersham), substrate DNA, and enzyme in assay buffer [50 mM Tris-HCl, pH 7.8, 1 mM $Na_2EDTA$, pH 8.0, 1 mM DTT, 7 $\mu$g/mL PMSF, 5% glycerol (v/v) and 100 $\mu$g/mL BSA]. For reaction rate analysis, 5 $\mu$M AdoMet and 1−5 nM enzyme were incubated with a fixed amount of DNA (5 $\mu$M CG). The substrates were FMR-1 unmethylated, methylated upper strand, and both strand methylated DNA. At fixed time intervals (0, 5, 10, 15, 20, 25, and 30 min), 25 $\mu$L of the reaction was transferred to an Eppendorf tube, and the reactions were stopped by immediately transferring the tubes to an ethanol−dry ice bath or spotted directly onto a DE81 circle (Millipore). The samples were processed as described previously (*7*, *24*).

For allosteric activation of the enzyme by methylated DNA, 10−30 nM enzyme, 5 $\mu$M CG (SNRPN exon 1, unmethylated), 5 $\mu$M AdoMet, and various concentrations of methylated CG (SNRPN exon 1, methylated) were used. The assay was started by the addition of cold DNA, enzyme, and tritiated AdoMet first with an increasing concentration of methylated DNA. For double reciprocal plots poly(dI-dC)·poly(dI-dC) and FMR-1 lower strand methylated DNA substrates were used.

To study the effect of Rb on allosteric activation, an increasing concentration (0, 80, 160, 240 nM) of Rb2 (MBP-Rb fusion protein containing amino acids 701−872 of Rb) and a fixed concentration of wild-type hDNMT1 (40 nM) were incubated at room temperature. Methylated DNA (5 $\mu$M), unmethylated DNA (5 $\mu$M), and AdoMet (7 $\mu$M) were added to the above mixture, and the reaction was continued at 37 °C for 30 min. The reaction was stopped and processed as described above. Steady-state kinetic experiments were done in duplicate, and mean values were calculated. The data obtained were analyzed by regression analysis using the GraphPad PRISM program (GraphPad Software Inc.). Kinetic constants represent estimated values.

*5-Fluoro-2′-deoxycytidine Assay.* Duplex oligonucleotides containing 5-fluoro-2′-deoxycytidine (FdC) were $^{32}$P end-labeled using polynucleotide kinase and [$\gamma$-$^{32}$P]ATP. hDNMT1 or deletion mutant enzymes (10−50 nM) were used with 5 nM fluorocytosine-containing oligonucleotide duplex in the presence or absence of 100 $\mu$M cold AdoMet in 1× buffer M at 37 °C for 30 min. The reaction was stopped by addition of 3× SDS loading dye and boiling the mixture for 5 min at 95 °C. DNA−enzyme complexes were analyzed as described previously (*7*).

*Generating GST Fusion hDNMT1 Peptides and DNA Binding Assay.* Specific DNA fragments were cloned in-frame with GST in pGEX5.1 (Amersham Pharmacia Biotech) and expressed in *Escherichia coli* strain ER2502 (New England Biolabs). Recombinant protein expression was induced at 30 °C by the addition of IPTG to 0.2 mM. Cells were harvested, lysed in buffer (1× PBS with 1% Triton X-100), clarified, and captured on GST beads. Details on constructs are available on request.

A typical reaction for the DNMT1−DNA binding assay consisted of 100 nM fusion protein, 3.2 $\mu$M poly(dA-dT)·poly(dA-dT), and 5 nM radioactive CG and/or methyl CGs in binding buffer containing 10 mM HEPES, pH 7.5, 10% glycerol (v/v), 50 mM KCl, 0.1 mM EDTA, 1 mM DTT, 2.5 mM $MgCl_2$, and 0.2% Triton X-100 (v/v). The mixture was incubated at 37 °C for 30 min, loaded onto a spin column, and washed twice with 1× PBS containing 160 mM NaCl at room temperature. The bound DNA was measured for radioactivity. The cpm values measured for the control GST were deducted from all of the fusion protein measurements.

To demonstrate a direct interaction between the allosteric activation domain of hDNMT1 and Rb, glutathione−Sepharose bead bound GST fusion hDNMT1 containing amino acids 1−510, GST-DNMT7 (pGEXD7 clone) protein (30 nM), was incubated with Rb2 (120 nM) for 1 h. The unbound Rb2 was washed away. To determine if binding of methylated DNA to the allosteric site of hDNMT1 interferes with Rb interaction methylated SNRPN exon 1 or FMR-1, DNA was added to the GST-DNMT7 fusion first and then 120 nM Rb2 was added. The DNMT1−Rb complex was separated on a 10−20% Tris−glycine gel, blotted on a PVDF (Millipore) membrane, and probed with anti-MBP to detect MBP-Rb fusion (Rb2) in the complexes. DNA binding assays on the binary complex were performed by incubating the complex with 50 ng of radiolabeled DNA. The unbound DNAs were washed away, and bound radioactivity was measured.

## RESULTS

*Catalytically Active Deletion Mutant hDNMTs*. Full-length hDNMT1 as well as the deletion mutants hDNMT1$^{\Delta 121}$, hDNMT1$^{\Delta 501}$, hDNMT1$^{\Delta 540}$, and hDNMT1$^{\Delta 580}$ (Figure 1) was able to incorporate methyl groups into the poly(dI-dC)·poly(dI-dC) substrate, whereas hDNMT1$^{\Delta 672}$ and hDNMT1$^{\Delta 1102}$ were catalytically inactive. Four mutants, hDNMT1$^{\Delta 121}$, hDNMT1$^{\Delta 501}$, hDNMT1$^{\Delta 580}$, and hDNMT1$^{\Delta 672}$, along with the full-length hDNMT1 were chosen for evaluation of the target specificity. Hemimethylated oligonucleotide duplexes containing 5-fluorocytosine in place of target cytosine were used. The complementary strand has a methylated cytosine, ensuring that the FdC-containing strand is the only methyl acceptor during the assay. The reaction mechanism requires an initial covalent attachment of the cysteine residue in motif IV of the methyltransferase to the target cytosine through position 6 in the presence of the cofactor AdoMet (*2*). Following methyl group transfer to C5, abstraction of a proton from C5 allows the enzyme to be released by $\beta$-elimination. Since fluorine is a poor leaving group, the $\beta$-elimination is inhibited, and the enzyme remains attached to the DNA. The enzyme−DNA complexes are resistant to denaturing SDS gel electrophoresis and thus aid in detection via autoradiography if DNA is radioactive. Three different hemimethylated oligonucleotide sets containing 5-fluorocytosine (F) in either CG (FG), CCG (FCG), or CWG (FWG) sequences were tested for covalent complex formation either in the presence or in the absence of AdoMet. The recombinant full-length hDNMT1 and Sf9 cell extracts were used as positive and negative controls, respectively. The results of this assay are shown in Figure 2A,B. Full-length hDNMT1 as well as the deletion mutants (hDNMT1$^{\Delta 121}$, hDNMT1$^{\Delta 501}$, and hDNMT1$^{\Delta 580}$) formed covalent complexes with oligonucleotides containing FG in the

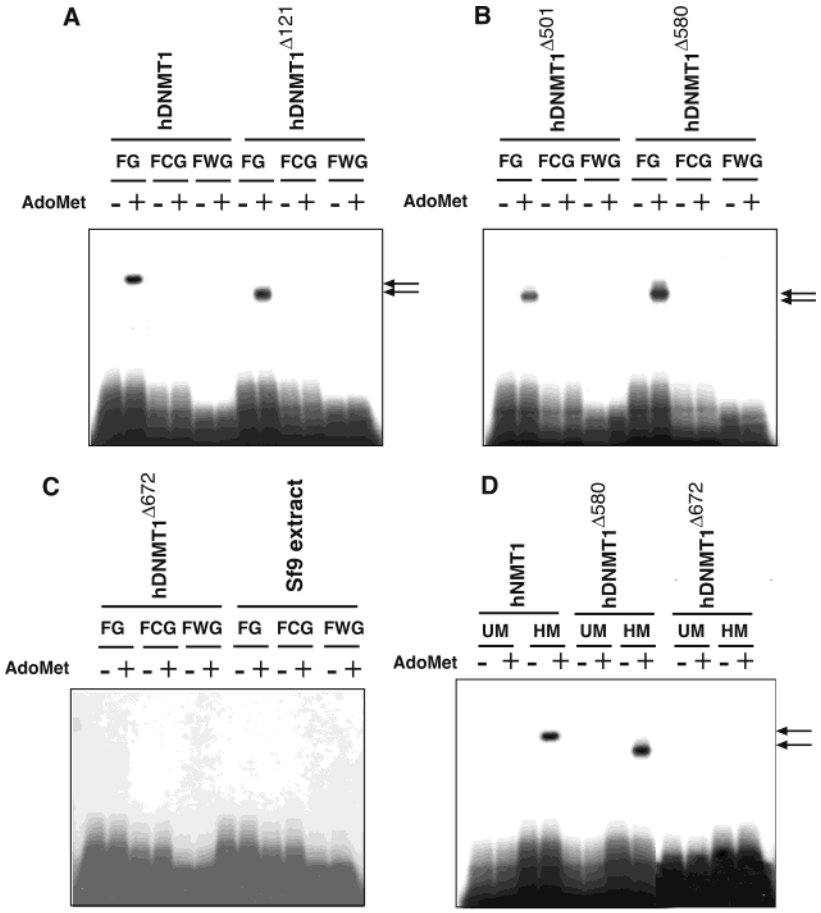


FIGURE 2: Catalytically active hDNMT1 and its amino-terminal deletion mutants. (A−C) Covalent cross-linking with hemimethylated substrates. Radioactive duplex oligonucleotides were covalently cross-linked with control hDNMT1 enzyme or the mutants as indicated above the gel. The presence or absence of AdoMet in the reaction is indicated with either a (+) or (−) sign above each lane. The target sites are indicated on the top as FG, FCG, or FWG. F represents 5-fluorocytosine and W is either A or T. DNA−protein cross-linked bands are indicated by arrows. Sf9 extracts indicate Sf9 cells expressing an empty pVIC1 backbone DNA. (D) Covalent cross-linking of unmethylated vs hemimethylated substrates. Radioactive duplex oligonucleotides were covalently cross-linked with control hDNMT1 enzyme or the mutant (hDNMT1$^{\Delta 580}$ and hDNMT1$^{\Delta 672}$) as indicated above the gel. The presence or absence of AdoMet in the reaction is indicated with either a (+) or (−) sign above each lane. UM or HM represents unmethylated or hemimethylated oligonucleotide substrate, respectively. All lanes contain equal amounts of radioactive oligonucleotide duplex. Unbound oligonucleotides are at the bottom of the gel.

presence of AdoMet. FCG and FWG sequences did not form strong complexes. Mutant DNMT1 lacking 672 amino acids also did not form covalent complexes in the presence of AdoMet with the hemimethylated FG oligonucleotide duplex (Figure 2C). This result confirms that hDNMT1 deletion mutants retain the CG sequence specificity comparable to the full-length hDNMT1.

Mutant hDNMT1$^{\Delta 580}$ was analyzed further for maintenance methylation using the FdC assay. Covalent complex formation between unmethylated and hemimethylated oligonucleotides containing 5-fluorocytosine (FG) was compared under similar experimental conditions. The results of this assay are shown in Figure 2D. Both full-length hDNMT1 and hDNMT1$^{\Delta 580}$ but not hDNMT1$^{\Delta 672}$ formed covalent complexes in the presence of AdoMet with the hemimethylated oligonucleotide substrate and were visible within 16 h of autoradiography. However, very weak signals were detected after 7 days of exposure (data not shown) for unmethylated oligonucleotide containing FdC. These results are consistent with previous kinetic data on a similar mutant, DNMT1$^{\Delta 501}$, which indicated that the dissociation constant for hemimethylated DNA was lower than unmethylated DNA (*31*).

*N-Terminal Deletion Recombinant hDNMT1 Prefers Maintenance Methylation*. Full-length hDNMT1 is known as the maintenance methyltransferase. It colocalizes with the replication foci during DNA replication along with several other protein factors (Figure 1). To understand the role of the amino terminus in enzymatic catalysis, the methylation rates of full-length and amino-terminal deletion mutant enzymes were compared using unmethylated, hemimethylated, and fully methylated FMR-1 repeat oligonucleotide duplexes as substrates under steady-state kinetic conditions (*7*). As expected, fully methylated oligonucleotide duplexes were poor substrates (data not shown). The reaction rates for hemimethylated and unmethylated oligonucleotides were measured, and the ratio of the two was calculated. As previously reported, the full-length hDNMT1 shows a 7-fold preference for hemimethylated DNA over unmethylated DNA (*7*). The preference for hemimethylated DNA by hDNMT1, hDNMT1$^{\Delta 121}$, hDNMT1$^{\Delta 501}$, and hDNMT1$^{\Delta 580}$ was 6.7-, 6.2-, 6.2-, and 9.2-fold, respectively (Table 1). This result further supports the conclusion that the N-terminal 580 amino acid residues of the enzyme have no influence on the maintenance methylation reaction in the present assay system.

Table 1: Rate of Methylation for Double-Stranded Unmethylated and Hemimethylated Oligonucleotides of Full-Length and Amino-Terminal Deletion Mutant hDNMT1[a]

| DNMTs | rate of methylation (cpm) | | methylation preference |
|---|---|---|---|
| | $(MGG/GCC)_{12}$ | $(CGG/GCC)_{12}$ | HM/UM |
| hDNMT1 | 60 | 9 | 6.7 |
| $hDNMT1^{\Delta 121}$ | 62 | 10 | 6.2 |
| $hDNMT1^{\Delta 501}$ | 50 | 8 | 6.2 |
| $hDNMT1^{\Delta 581}$ | 55 | 6 | 9.2 |

[a] M = 5 methylcytosine, HM = hemimethylated substrate [$(MGG/GCC)_{12}$], and UM = unmethylated substrate [$(CGG/GCC)_{12}$].

*The N-Terminus of hDNMT1 Has a Methylated DNA-Dependent Allosteric Activation Domain*. Previous studies showed stimulation of DNA methylation by fully methylated random sequence DNA (*24*), suggesting the presence of an allosteric activation domain in hDNMT1. The rate of methylation was compared between full-length hDNMT1 and its deletion mutants containing the conventional target recognition domain in the presence of unmethylated SNRPN exon 1 substrate and an increasing amount of methylated SNRPN exon 1 as the allosteric activator. Addition of a mixture of AdoMet and unmethylated DNA allows the cofactor AdoMet to initiate the methylation reaction (*2*). An increase in the concentration of methylated DNA increased the methyl group incorporation up to ~2.5-fold by both hDNMT1 and $hDNMT1^{\Delta 121}$ (Figure 3A). This stimulation was significantly reduced with mutant enzymes $hDNMT1^{\Delta 501}$, $hDNMT1^{\Delta 540}$, and $hDNMT1^{\Delta 580}$. The methylation rate curve as a function of methylated DNA is hyperbolic for both hDNMT1 and $hDNMT1^{\Delta 121}$ (Figure 3A). Substituting methylated DNA with unmethylated SNRPN exon 1 DNA or poly(dA-dT)·poly(dA-dT) failed to yield a hyperbolic curve, showing that methylated DNA is the allosteric activator. The observation of a hyperbolic activation curve for $hDNMT1^{\Delta 121}$ but not with $hDNMT1^{\Delta 501}$ and $hDNMT1^{\Delta 580}$ suggests that the first 121−501 amino acids are essential for allosteric activation.

*The N- and C-Termini of hDNMT1 Bind DNA Independently*. To identify the methylated DNA binding regions of hDNMT1, several GST fusion proteins containing different segments of hDNMT1, which together covered its entire length, were used. A GST fusion containing the entire amino terminus of hDNMT1 was insoluble in *E. coli* and was not used. The fusion proteins were captured on GST beads, and the recovered protein was over 90% pure. The beads were incubated with $^{32}$P-labeled fully methylated FMR-1/SNRPN exon 1 DNA either in the presence or in the absence of poly(dA-dT)·poly(dA-dT), a nonspecific DNA. Radioactivity captured on the beads was measured. The DNA binding activity of GST fusion proteins is summarized in Table 2. The GST fusion proteins containing amino acids 1−261 and 365−510 were not able to bind methylated DNA, whereas fusion proteins comprising amino acids 1−322, 1−336, and 1−510 did. Thus amino acids 261−365 were identified as a DNA binding region. Moreover, fusion proteins containing amino acids 432−836 were also able to bind DNA. These results identify a second DNA binding site between amino acids 510 and 789. This region includes the Zn binding motif of hDNMT1. The catalytic domain of hDNMT1 was identified as the third DNA binding region since fusion peptides encoding either amino acids 1102−1616 or amino acids 1374−1616 of hDNMT1 were able to bind DNA. Similar methylated DNA binding activities were observed for prokaryotic DNA (cytosine-5) methyltransferase, M. *Msp*I (*25*). Thus methylated DNA binding properties of the fusions

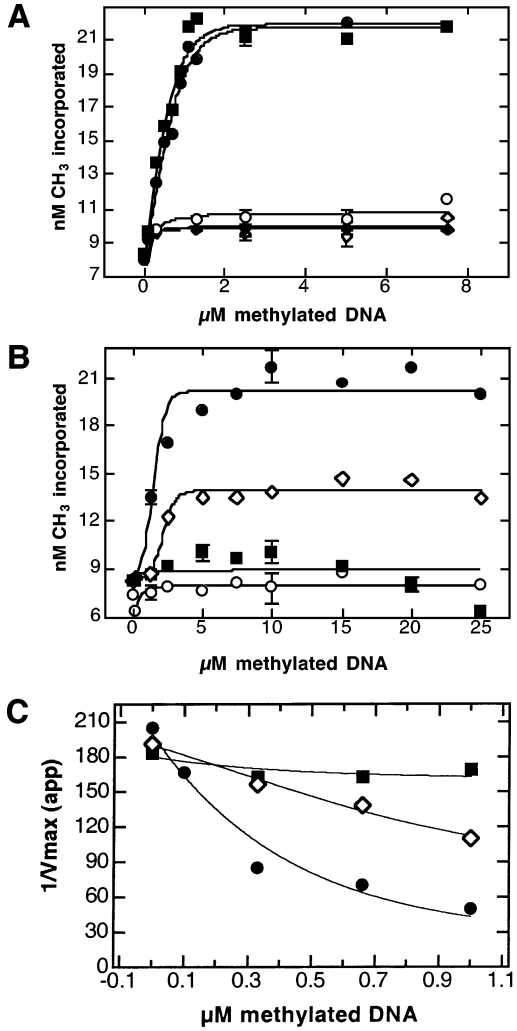


FIGURE 3: Allosteric activation of DNA methylation by methylated DNA. (A) Methyltransferase reaction of hDNMT1 and its deletion mutant as a function of increasing DNA concentration. Methylated (SNRPN exon 1), unmethylated (SNRPN exon 1), and poly(dA-dT)·poly(dA-dT) DNAs were used as potential allosteric activators. Full-length hDNMT1 (filled circles), $hDNMT1^{\Delta 121}$ (filled squares), and $hDNMT1^{\Delta 580}$ (open circles) are shown in the presence of methylated DNA. Data points for $hDNMT1^{\Delta 501}$ were similar to $hDNMT1^{\Delta 580}$ and thus were not included. Substitution of poly(dA-dT)·poly(dA-dT) (open diamonds) or unmethylated DNA (filled diamonds) instead of methylated DNA for the full-length hDNMT1 is shown. (B) Comparison of DNA-mediated methylation activation as a function of increasing amounts of methylated DNA using wild-type and point mutant enzymes. Wild type (filled circles), point mutant D12 (open diamonds), and point mutant D25 (filled squares) are shown. Open circle data points show methylation of wild-type hDNMT1 in the presence of unmethylated DNA as activator. Each time point depicts the mean values of nM [$^3$H]methyl group incorporation from two independent experiments. (C) Methylation activation by fully CG-methylated SNRPN exon 1 DNA. Replot of *y*-axis intercepts from double reciprocal plots of initial velocities obtained with unmethylated SNRPN exon 1 DNA as the variable substrate (0.25, 0.325, 0.42, 0.6, 1.25, 2.5 $\mu$M CG) in the presence of changing fixed concentration of methylated CG from 0 to 1 $\mu$M methyl-CG. AdoMet concentration was kept constant at 7 $\mu$M in all the reactions. Data points for wt DNMT1 (filled circle), mutant D12 (open diamond), and D25 (filled square) are shown.

Table 2: DNA Binding Regions of the hDNMT1[a]

| GST fusion clones | amino acids | DNA binding activity (cpm) |
|---|---|---|
| pGEXD1 | 1−148 | ND |
| pGEXD2 | 1−217 | ND |
| pGEXD3 | 1−261 | ND |
| pGEXD4 | 1−322 | 12000 ± 3000 |
| pGEXD5 | 1−336 | 9300 ± 1100 |
| pGEXD6 | 1−446 | 9200 ± 800 |
| pGEXD7 | 1−510 | 8500 ± 500 |
| pGEXD8 | 356−510 | 200 ± 50 |
| pGEXD9 | 432−836 | 3400 ± 500 |
| pGEXD10 | 789−1109 | ND |
| pGEXD11 | 1102−1616 | 1700 ± 700 |
| pGEXD12 | 1374−1616 | 800 ± 200 |

[a] The background DNA binding value for GST protein alone was ∼1100 cpm. The values in the table have this background already subtracted out. DNA binding values below 150 cpm were discarded (ND, not detected).

Table 3: Mutational Analysis of the DNA Binding Region between Amino Acids 261 and 356 of hDNMT1[a]

| mutation | DNA binding activity (cpm) |
|---|---|
| wild type | 11500 ± 2700 |
| 281(K-A) | 10500 ± 1500 |
| 284(K-A) | 2500 ± 500 |
| 291(K-A) | 11800 ± 3000 |
| 326(K-E) | 10800 ± 2200 |
| 335(K-N) | 17000 ± 3000 |
| 284(K-A), 285(K-A) | ND |
| 286(H-A), 287 (R-A) | ND |
| 326(K-E), 335(K-N) | 15500 ± 2700 |

[a] Wild-type fusion protein is from clone pGEXD7. The values in the table have background already subtracted out. DNA binding values below 150 cpm were discarded (ND, not detected).

covering the catalytic domain of hDNMT1 may be an intrinsic property of this enzyme, or alteration of DNA binding and selectivity may be due to large deletion or differential protein folding.

*Point Mutation in the N-Terminal DNA Binding Region of hDNMT1 Preserves de Novo Methylation and Reduces AdoMet-Dependent Positive Cooperativity and Methylated DNA-Dependent Allosteric Activation*. DNA binding analysis showed that amino acids 261−356 were crucial for DNA−protein association, perhaps playing a role in the allosteric activation of hDNMT1. To identify the ligand binding site for allosteric activation, PCR-mediated mutagenesis was performed on clone pGEX7. Mutant clones expressing GST fusions that did not bind methylated DNA were identified, and methylated DNA binding activity was measured for each fusion (Table 3). While mutation at amino acids 281 and 291 did not affect methylated DNA binding activity, a clone containing mutation at amino acid 284 showed ∼80% loss of binding. Furthermore, mutation at amino acid 326 or 335 did not show reduced methylated DNA binding activity, suggesting that the region containing amino acid 284 may be involved in methylated DNA binding. Thus a series of alanine substitutions of this region were performed to validate the above observation. Indeed, GST fusions with alanine substitution at positions 284 and 285 or 286 and 287 reduced methyl DNA binding activity to background but not in the double mutant 326 and 335. Thus these mutatations were incorporated into the full-length hDNMT1 cDNA to generate full-length mutant hDNMT1s. Two full-length hDNMT1 mutants, D12 (H286A and R287A) and D25 (K284A and K285A), exhibited reduced allosteric activation in the presence of methylated DNA and cosubstrate AdoMet (Figure 3B). To confirm these observations, a detailed product activation study with methylated DNA was performed with wild-type (wt) and point mutant enzymes. A concentration range of 0−1 $\mu$M methylated CGs (methylated SNRPN exon 1) was added to SNRPN exon 1 unmethylated substrate DNA containing 0.25−2.5 $\mu$M CGs. The AdoMet concentration in all of the reactions was kept constant at 7 $\mu$M. If methylated DNA acts as an activator, *y*-axis intercepts were expected to decrease with increasing concentration of methylated CGs reflecting methylated DNA-mediated activation. For the wt hDNMT1 the *y*-axis intercepts and $1/V_{max(app)}$ decreased as more methylated DNA was added to the reaction, confirming enzyme activation as observed previously with randomly methylated CG-containing DNA (*24*). Intercepts for D25 converged on the *y*-axis and did not change in the presence of increasing amounts of methylated DNA, suggesting competitive inhibition rather than activation (data not shown). A plot of $1/V_{max(app)}$ vs methyl-CG in Figure 3C shows highest, moderate, and very little activation for wt hDNMT1, point mutant D12, and point mutant D25, respectively. Overall, these data demonstrate that fully methylated DNA acts as an activator for wt hDNMT1 and point mutant D12 but not for point mutant D25. The decrease in allosteric activation by the mutations in D12 and D25 suggests that they increased dramatically the dissociation constant for methylated DNA and/or altered the ability of the enzyme to undergo necessary conformational changes.

The steady-state kinetic properties along with the kinetic constants of both D12 and D25 mutant enzymes were analyzed and compared with those of the full-length hDNMT1. The plot of $1/v$ vs 1/[S] for an allosteric enzyme is curved and approaches a horizontal line that intersects the $1/v$ axis at $1/V_{max}$ (*26*). A series of double reciprocal plots were made at five different AdoMet concentrations or six different CI (DNA) concentrations, using poly(dI-dC)·poly(dI-dC) with an average length of ∼7000 bp. Poly(dI-dC)·poly(dI-dC) concentration was maintained between 0.1 and 1 $\mu$M to avoid DNA substrate inhibition (*7*). The reciprocal of the amount of $[^3H]CH_3$ transferred to the DNA in 1 min by 1 nM DNMT1 ($1/v$) is plotted as a function of 1/substrate concentration (either 1/[CI] or 1/[AdoMet]) in the presence of fixed concentrations of the cosubstrate. For bireactant enzymes that use two substrates and give two products such as hDNMT1, these double reciprocal plots should yield straight lines, and their slopes and intercepts are then replotted to derive all of the steady-state kinetic constants. In methylation reactions using D12 and poly(dI-dC)·poly(dI-dC) as substrate, the double reciprocal plots, with 1/[CI] as the variable substrate, were linear (Figure 4A), while those with 1/[AdoMet] as the variable substrate were nonlinear (Figure 4B). Velocities were unchanged, and plots were parallel to the *x*-axis at AdoMet concentrations above 2.5 $\mu$M. This reflects a combination of linear and nonlinear behavior of mutant D12, suggesting positive cooperativity due to enzyme activation at higher AdoMet concentration as observed before for wt hDNMT1 (*7*). However, under identical reaction conditions and at the same substrate concentrations, the $1/v$ vs 1/[CI] and 1/[AdoMet] plots were linear for D25 (Figure 4C,D) compared to D12. The linearity

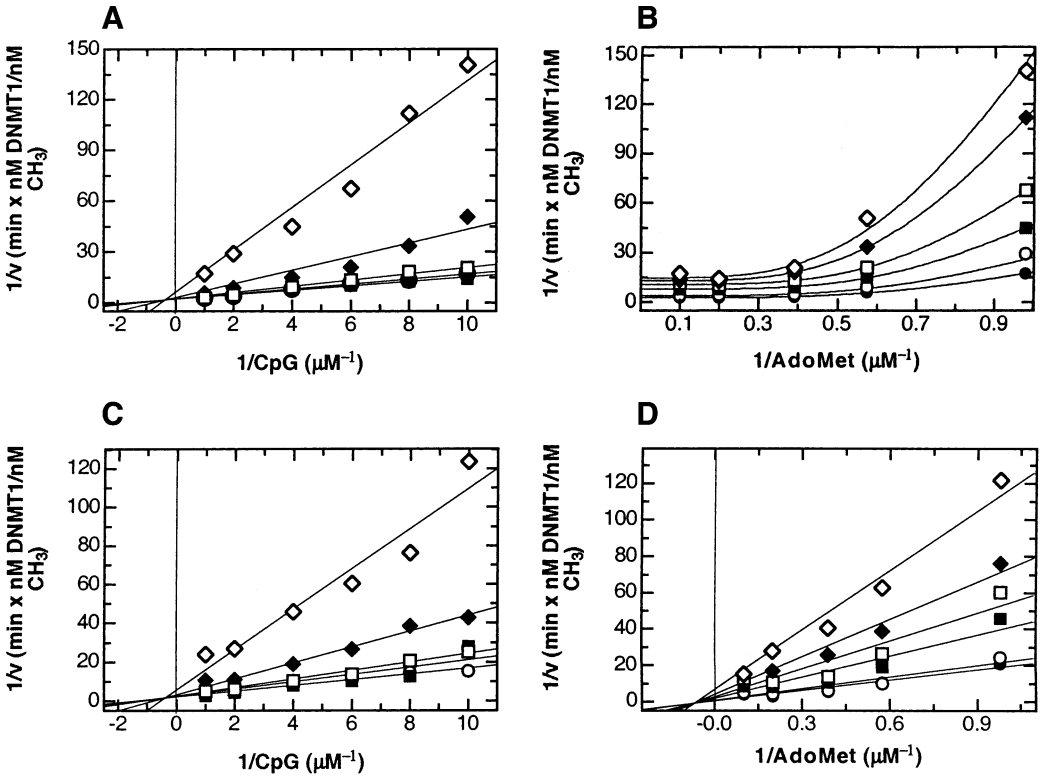


FIGURE 4: Double reciprocal plots for poly(dI-dC)·poly(dI-dC) for D12 and D25 mutants under identical conditions. (A, C) The AdoMet concentration was fixed at 1 $\mu$M (open diamonds) 1.7 $\mu$M (filled diamonds), 2.6 $\mu$M (open squares), 5 $\mu$M (open circles), and 10 $\mu$M (filled circles). $1/v$ is plotted against 1/[CI]. (B, D) The DNA (CI) concentration was fixed at 0.1 $\mu$M (open diamonds), 0.12 $\mu$M (filled diamonds), 0.17 $\mu$M (open squares), 0.25 $\mu$M (filled squares), 0.5 $\mu$M (open circles), and 1.0 $\mu$M (filled circles). $1/v$ is plotted against 1/[AdoMet] using a nonlinear regression. (A) and (B) represent D12 and (C) and (D) represent D25 data, respectively. The intercepts on the *y*-axis and the slopes were replotted to obtain the second set of kinetic constants. The average values of the kinetic constants are shown in Table 4.

Table 4: Comparison of Steady-State Kinetic Parameters of Wild-Type and Mutant hDNMT1 Using Poly(dI-dC)·Poly(dI-dC) as Substrate

| hDNMT1 | $k_{cat}$ ($h^{-1}$) | $K_m^{AdoMet}$ ($\mu$M) | $K_m^{CI}$ ($\mu$M) | $k_{cat}/K_m^{CI}$ ($\times 10^6$ $M^{-1}$ $h^{-1}$) |
|---|---|---|---|---|
| wild type | $210 \pm 26$ | $7.20 \pm 2.0$ | $0.56 \pm 0.1$ | 375 |
| mutant D12 | $108 \pm 08$ | $9.20 \pm 0.53$ | $0.50 \pm 0.1$ | 216 |
| mutant D25 | $48 \pm 04$ | $3.40 \pm 0.2$ | $0.82 \pm 0.02$ | 58 |

reflects reduction of AdoMet-dependent positive cooperativity for mutant D25. Thus, mutation at the amino-terminal regulatory region, which is situated ~1000 amino acids away from the AdoMet binding motif of DNMT1, may effect cosubstrate AdoMet binding, suggesting functional cooperation between both domains. Comparative parameters, such as $K_m^{AdoMet}$, $K_m^{CI}$, $k_{cat}$, and $k_{cat}/K_m^{CI}$, are listed in Table 4. The values of $k_{cat}$ for the recombinant human enzyme mutants D12 and D25 were 108 and 48 $h^{-1}$, respectively, at least 2−4-fold lower than the value for wt hDNMT1 (210 $h^{-1}$). The $K_m$ values of CI and AdoMet differed by 2−3-fold between enzymes.

To confirm the above observation on a representative human DNA sequence, mutant enzymes were examined with FMR-1 repeat substrates for de novo and maintenance methylation. FMR-1 contains highly polymorphic CGG repeats. In fragile X patients these repeats are methylated and the gene is silenced. For unmethylated FMR-1, linear velocity responses were observed for both D12 and D25, similar to those for wt hDNMT1 (*7*), where $1/v$ was plotted as a function of 1/[substrate] (data not shown). Remarkably, $K_m^{CG}$, $k_{cat}$, and $k_{cat}/K_m^{CG}$ values were similar (Table 5), suggesting that de novo methylation properties were not affected despite mutation in the allosteric activation pocket. $K_m^{AdoMet}$ was 2-fold higher for the D25 mutant.

When FMR-1 lower strand methylated substrate was used to determine maintenance methylation rates, D12 showed linear velocity responses, as DNA was the variable substrate (Figure 5A). However, in the presence of variable AdoMet concentrations, the plots were not linear (Figure 5B). Velocity remained unchanged, and the plots were parallel to the *x*-axis as AdoMet concentration rose above ~2 $\mu$M. These plots were similar to those for the wt hDNMT1 (*24*), suggesting that mutation on D12 cannot significantly reduce AdoMet-mediated cooperativity. However, mutant enzyme D25 exhibited linear velocity responses for both variable DNA (Figure 5C) and for variable AdoMet concentrations (Figure 5D), showing reduced cooperativity. For D25, $k_{cat}$ and preference for hemimethylation was 3-fold lower (Tables 5 and 6), despite a similar catalytic efficiency, $k_{cat}/K_m^{CG}$. The curved velocity responses at fixed AdoMet concentrations were dependent on DNA concentration, since $1/V_{max(app)}$ (*y*-axis intercepts) decreased with increasing CI or CG concentration. Thus, the kinetic behavior of the human DNMT1 and its mutant depends on the substrate DNA, its methylation status, and AdoMet-dependent cooperativity. This suggests that mutations in the allosteric pocket have direct consequences on the performance of hDNMT1. Mixed linear and nonlinear responses confirm a steady-state mechanism, where DNA can act as a substrate and an activator for methylation reaction as in the D12 mutant. Similar responses were observed for wt hDNMT1 (*24*). Linear

Table 5: Steady-State Kinetic Parameters for de Novo and Maintenance Methylation between Wild-Type and Mutant Recombinant hDNMT1 Using FMR-1 Repeat DNA as Substrate[a]

| hDNMT1 | DNA | $k_{cat}$ ($h^{-1}$) | $K_m^{AdoMet}$ ($\mu$M) | $K_m^{CG}$ ($\mu$M) | $k_{cat}/K_m^{CG}$ ($\times 10^6$ $M^{-1}$ $h^{-1}$) |
|---|---|---|---|---|---|
| wild type | UM | $1.23 \pm 0.17$ | $3.81 \pm 1.24$ | $0.33 \pm 0.11$ | $3.9 \pm 0.9$ |
| mutant D12 | UM | $0.9 \pm 0.1$ | $8.20 \pm 2.0$ | $0.5 \pm 0.1$ | $1.80 \pm 0.2$ |
| mutant D25 | UM | $1.1 \pm 0.1$ | $13.6 \pm 4$ | $0.5 \pm 0.1$ | $2.2 \pm 0.08$ |
| wild type | HM | $22.3 \pm 2$ | $4.7 \pm 0.5$ | $0.5 \pm 0.1$ | $44.7 \pm 1.2$ |
| mutant D12 | HM | $23 \pm 4$ | $1.96 \pm 0.3$ | $2.5 \pm 0.5$ | $9.2 \pm 1.6$ |
| mutant D25 | HM | $7 \pm 0.5$ | $0.9 \pm 0.7$ | $0.6 \pm 0.4$ | $20 \pm 12$ |

[a] HM = hemimethylated substrate [$(CGG/GMC)_{12}$], UM = unmethylated substrate [$(CGG/GCC)_{12}$], and M = 5-methylcytosine.

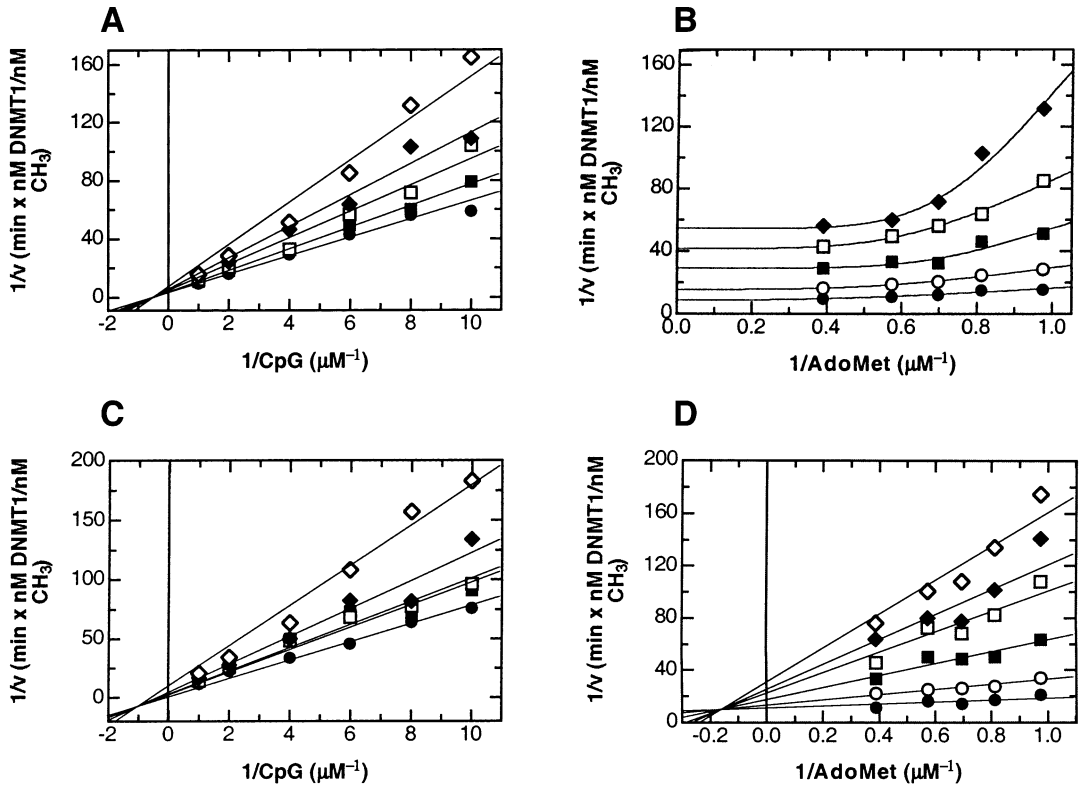


FIGURE 5: Double reciprocal plots for the hemimethylated FMR-1 for D12 and D25 mutants under identical conditions. (A, C) Fixed AdoMet plots. The AdoMet concentration was 1 $\mu$M (open diamonds), 1.2 $\mu$M (filled diamonds), 1.4 $\mu$M (open squares), 1.7 $\mu$M (filled squares), and 2.6 $\mu$M (filled circles). (B, D) Fixed DNA plots. The CG concentration 0.12 $\mu$M (filled diamonds), 0.17 $\mu$M (open squares), 0.25 $\mu$M (filled squares), 0.5 $\mu$M (open circles), and 1 $\mu$M (filled circles). $1/v$ is plotted against 1/[AdoMet] using a nonlinear regression. (A) and (B) represent D12 and (C) and (D) represent D25 data, respectively. The intercepts on the *y*-axis and the slopes were replotted to obtain the second set of kinetic constants. The average values of the kinetic constants are shown in Table 5.

Table 6: Preference for Hemimethylated DNA by Wild-Type and Mutant Recombinant hDNMT1

| hDNMT1 | relative $k_{cat}$ (HM/UM) | relative $k_{cat}/K_m^{CG}$ (HM/UM) |
|---|---|---|
| wild type | $18.5 \pm 1.0$ | $12.0 \pm 2.0$ |
| mutant D12 | $25.4 \pm 1.6$ | $5 \pm 0.15$ |
| mutant D25 | $6.4 \pm 0.15$ | $8.8 \pm 0.5$ |

behavior of the D25 mutant with both poly(dI-dC)·poly-(dI-dC) and hemimethylated FMR-1(Figures 4D and 5E) confirms that mutation at K284A and K285A indeed leads to reduction of AdoMet-dependent cooperativity and allosteric activation for hDNMT1 (Figure 3B,C).

*Binding of Rb at the Allosteric Site Regulates Methylation Spreading*. We have previously shown that B and C pockets of Rb (amino acids 701−872) bind to the regulatory domain of hDNMT1 and modulate its activity (*18*). Overexpressing Rb in human cells using an adenovirus construct resulted in a hypomethylated genome (*18*). Rb was also able to inhibit maintenance methyltransferase activity in vitro. To understand the mechanism of modulation and its relationship with the allosteric DNA binding site, a direct binding assay of Rb2 and GST-DNMT1 fusion protein containing the allosteric activation domain (amino acids 1−510) was performed. Furthermore, the DNA binding activity of the binary complex was measured with methylated FMR-1 and SNRPN exon 1 DNA. Both DNA sequences mimic genes that are methylated during development or disease progression. FMR-1 DNA contains high CG density as typically found in CpG islands of tumor suppressor genes that are silenced in cancer cells via DNA methylation.

In a GST pull-down assay, a fusion protein containing the allosteric activation domain of hDNMT1, GST-DNMT7 (pGEXD7), was able to pull-down Rb2 (Figure 6A, lane 5). However, GST alone with or without methylated DNA (F = FMR-1 or S = SNRPN exon 1) was not able to pull down Rb2 (Figure 6A, lanes 1−3). These results show a direct

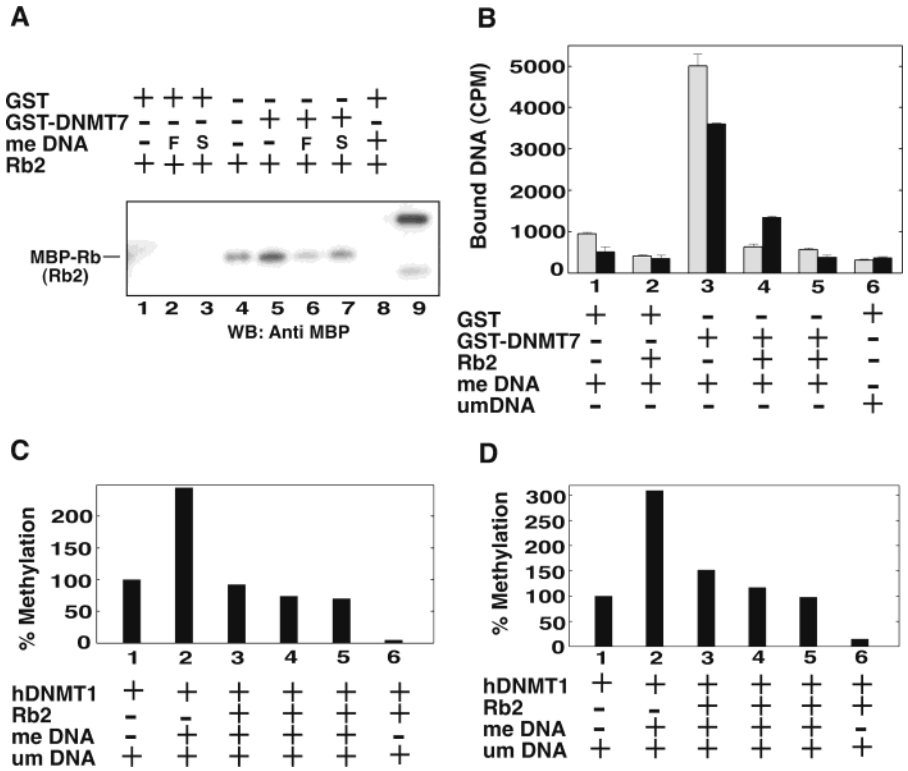


FIGURE 6: Rb negatively modulates de novo methylation spreading. (A) Western blot analysis of binding between Rb2 and the allosteric activation subdomain of human DNMT1 and the effect of methylated DNA on protein−protein interaction. The reaction components are indicated above in the order added: GST, glutathione *S*-transferase bound to glutathione−Sepharose beads; GST-DNMT7, GST fusion protein containing amino acids 1−510 (pGEXD7) of hDNMT1 bound to glutathione−Sepharose beads; Rb2, MBP fusion Rb gene product containing amino acids 701−872; me DNA, methylated FMR-1 (F) or SNRPN exon 1 (S). Lane 4 shows one-fifth of input Rb2. Lane 9 is the biotinylated marker. (B) Regulation of methylated DNA binding at the allosteric activation subdomain of hDNMT1. The reaction components are indicated. Gray and black bars represent the radioactive FMR-1 and SNRPN exon 1 DNA fragment bound to various protein complexes. Rb2 was added to GST-DNMT1 for binary complex formation followed by addition of me DNA in bars 4 and 5. The relative concentrations of Rb2 for construction of the binary complexes were 2× (bar 4) and 4× (bar 5) molar concentrated as compared to GST-DNMT fusion protein. Unmethylated DNA controls are in bar 6. (C) Methyltransferase assay with FMR-1 substrate and (D) with SNRPN exon 1 DNA. The reaction components are indicated below in the order added. A fixed concentration of hDNMT1 (40 nM, the first reactant) and an increasing concentration of Rb2 (0, 80, 160, and 240 nM, bars 2−5) were added and incubated for 10 min at room temperature followed by addition of DNA. Methyl group transfer shown in bars 6 and 4 represents with or without allosteric activator methylated DNA in the reaction. Radioactivity measured without the presence of hDNMT1 in the reaction mix was considered background value, and the final methylation values presented in the figures are after subtraction of the background value. De novo methylation was calculated as 100%. Each bar represents an average of at least two independent assays.

interaction between hDNMT1 (amino acids 1−510) and Rb2. Preincubation of GST-DNMT1 with methylated DNA before addition of Rb2 in a similar pull-down assay resulted in reduction of DNMT1 and Rb interaction (Figure 6A, lane 3 vs lane 4 or 5), suggesting that methylated DNA binding to the allosteric activation domain will reduce Rb binding. To examine whether the hDNMT1-Rb binary complex could bind methylated DNA in the allosteric subdomain, a direct DNA binding assay was performed. As expected, both FMR-1 and SNRPN exon 1 methylated DNA bound to the allosteric activation domain in comparison to control GST protein plus BSA or Rb2 (Figure 6B, column 3 vs column 1 or 2). Addition of equal or a 4 molar excess amount of methylated DNA to preformed DNMT1-Rb binary complexes resulted in reduced DNA binding (Figure 6B, column 4 or 5 vs column 3), suggesting that binding of Rb to the allosteric activation domain of hDNMT1 interferes with the access of methylated DNA.

To investigate the regulatory role of Rb on methylation spreading, methyltransferase assays were performed with unmethylated FMR-1 and SNRPN exon 1 oligonucleotide duplexes as substrates in the presence of Rb2. When hDNMT1 was incubated with unmethylated and methylated DNA, methyl group incorporation was ~2.5 and 3 times higher, respectively, as compared with unmethylated substrate DNA alone (Figure 6C,D, bar 1 vs bar 2). This observed de novo methylation stimulation was reduced when hDNMT1 was preincubated with Rb2 followed by methylated and unmethylated DNA substrate addition. A 2× molar concentration of Rb was able to bring down the methylation stimulation to the original de novo methylation level for FMR-1 DNA (Figure 6C,D, bar 3 vs bar 1), whereas a much higher concentration of Rb2 (6×) was required for the same effect with SNRPN exon 1 substrate (Figure 6C, bar 5 vs bar 1). This result suggests that Rb is more effective in suppressing methylation spreading in CpG islands than in regions where CG dinucleotides are dispersed. In the presence of methylated DNA, a 6× molar excess of Rb2 had little or no effect on de novo methylation activity (Figure 6C,D, bar 1 vs bar 5) although allosteric activation was abolished (Figure 6C,D, bar 2 vs bar 5). In the absence of methylated DNA, a 4× molar concentration of Rb in the reaction mixture inhibited the de novo methylation for both of the substrates (Figure 6C,D, bar 1 vs bar 6) suggesting

that Rb also could modulate de novo methyltransferase reaction effectively. The same molar concentration of Rb2 had little or no effect on de novo methylation in the presence of the allosteric activator, methylated DNA (Figure 6C,D, bar 1 vs bar 4).

## DISCUSSION

Mammalian DNMT1 has multiple functional domains. While a region of the N-terminus binds to PCNA and facilitates maintenance methylation in vivo, another region functions as a transcriptional repressor domain (*27*). In this study, deletions of large parts of the N-terminus resulted in catalytically competent mutant enzymes such as that of full-length wt hDNMT1. These enzymes have maintenance methylation properties and target CG specificity identical to that of full-length hDNMT1. Lack of DMAP1, PCNA, DB1, Zn binding, serine phosphorylation at SK (*28*), and the entire replication foci directing peptide region had no effect on the enzyme's ability to recognize and methylate hemimethylated DNA in vitro. Both target recognition and methylation occur in an identical fashion for hDNMT1$^{\Delta 580}$ and hDNMT1.

Using biochemical approaches for DNMT1, small regions of the large N-terminus have been identified as sites responsible for nuclear localization or for directing the enzyme to replication foci (*10*). Several other protein binding sites have been mapped for DNMT1 (Figure 1). Still the function of much of the N-terminus of hDNMT1 is poorly understood. Expression of either the N-terminal or the catalytic domain alone gives a catalytically inactive protein, even though all of the predicted conserved motifs for catalysis are present in the catalytic domain. Thus the interplay between the domains is crucial for catalysis, perhaps through domain−domain interactions or because essential catalytic residues reside in the N-terminal region of the protein. Results presented here suggest that the TRD lies between amino acids 580 and 1616 in hDNMT1 and is not localized between amino acids 121−417 as proposed earlier by Araujo et al. (*29*). Furthermore, the central region of the enzyme (amino acids 580−1110) is essential for enzyme activity. In most other DNA (cytosine-5) methyltransferases the TRD lies squarely in the middle of the enzyme between conserved motifs VIII and IX. The only exception is the phage-encoded multispecific methyltransferase M. (ø)*Bss*HII (*30*). Isoform I has a second TRD at a noncanonical position at the N-terminus along with a conventional TRD. The presence of a noncanonical TRD provides *Hae*II methylation specificity to isoform I.

Hyperbolic velocity curves for full-length hDNMT1 or hDNMT1$^{\Delta 121}$ but not for hDNMT1$^{\Delta 501}$ and hDNMT1$^{\Delta 580}$ led us to identify the first 121−500 amino acids of hDNMT1 as an allosteric activation domain that needs binding of a methylated DNA for catalytic activation. Amino acid sequence comparison of the allosteric activation domain region shows a lower degree of sequence similarity between vertebrate DNMTs, suggesting different evolutionary functions. For the murine enzyme, zinc ion is essential for catalytic activation in the presence of methylated DNA (*31*) but not for hDNMT1, suggesting a fundamental difference between the murine and human enzymes. We had previously shown that the complex pattern of methylation for hDNMT1 in vitro is the result of additional DNA binding sites at the amino-terminal region (*24*, *32*). Thus, we hypothesized that two DNA molecules may bind hDNMT1 simultaneously, a methylated substrate at the regulatory region and an unmethylated or hemimethylated DNA at the catalytic center. Binding at the regulatory site exposes the catalytic center at the C-terminal region and perhaps facilitates access to AdoMet and DNA. In vivo this is possible since PCNA binds to hDNMT1 (amino acids 163−174) and brings the allosteric DNA binding site of the enzyme close to the replication foci. Indeed, physical interaction between PCNA and DNMT1 facilitates methylation of hemimethylated (newly replicated) DNA (*33*). Steady-state kinetic analysis with hDNMT1 and mutants confirms the presence of an allosteric activation pocket. Mutation in this pocket affected hemimethylation, although the catalytic efficiency remains unchanged (Table 5). However, misfolding or conformational changes of point mutants cannot be ruled out. Thus, interactions between the allosteric activator pocket and the catalytic domain appear to be crucial for DNA methylation.

Methylation spreading in mammalian genomes is common, although the mechanisms are poorly understood. For murine DNMT1, 5-methyl-2′-deoxycytidine in single-stranded DNA can act in cis to signal de novo DNA methylation (*34*). In this work, double-stranded methylated DNA bound at the amino-terminal allosteric activation pocket was shown to activate de novo methylation of CG sites at the catalytic center of hDNMT1. This phenomenon explains the most common transcriptional gene silencing mechanism of parasitic DNA sequences, such as IAP endogenous retroviruses, in the genome. A specific methylated sequence can activate methylation of identical repeat sequences. This is physiologically significant since host methylation appears to be involved in host defense against the activation of retroposon promoters (*35*), thus controlling their spreading. The other type of methylation spreading may involve cis activation, which in turn could lead to localize preferred methylation and, hence, spreading. This kind of cis activation requires that the previously methylated DNA be able to loop back and interact with the enzyme. In fact, unusual DNA substrates, such as hairpins of CCG repeats in cruciform structures, were better substrates than control Watson−Crick-paired duplexes by human DNMT1 (*36*). The above mechanisms may be possible due to the modular structure of DNMT1 where both regulatory and catalytic domains cooperate with each other. Indeed, the domains of DNMT1 could be separated by V8 protease digestion (*37*).

A typical human nuclear genome is $3.2 \times 10^9$ base pairs packed into 23 chromosomes (*38*), out of which methylated cytosine accounts for ~1% and therefore affects about 70−80% of total CpGs (*39*). It takes about 24 h for cell division, out of which S phase is about 8−10 h. During this time the nuclear genome is replicated, and maintenance methylation is ensured by hDNMT1. Thus, in normal cells the biological importance of the allosteric activation domain, which normally binds to the replication fork with the help of PCNA, may be to sense methylated DNA during DNA replication and to activate catalysis by a fast velocity state of the enzyme. During these processes chromatin modeling proteins are excluded and transcription factors are allowed to interact with the promoters. Insulator proteins ensure the unmethylated status of CpG islands. Possible involvement of regulator protein(s), such as Rb, for controlled maintenance methy-

lation cannot be ruled out. During somatic cell division, hDNMT1, an enzyme with low turnover numbers, allosteric activation would be beneficial for rapid DNA methylation.

The challenge of establishing and maintaining the methylation pattern is further complicated during early embryonic development in mammals, where the cell division time and S phase are very short, and several dynamic demethylation and remethylation events take place in the genome. In early development, during preimplantation, the genome becomes substantially demethylated (*40*). At the blastocyst stage, DNA methylation is restored (*41*). During the postimplantation stage DNMT1, DNMT3a, and DNMT3b are expressed. The de novo methylation activity of DNMT3 and its expression in early development are believed to be responsible for reestablishing the methylation pattern (*42*). However, in vitro analysis of DNMT3 shows that de novo activity is about 6-fold less than that of mammalian DNMT1 (*43*). This suggests a possible cooperation between de novo and maintenance enzymes for an initial establishment of methylation pattern. One such example may be DMR2 of Igf2, where all three methyltransferases participate in maintenance of methylation (*44*). Recently, several published reports show functional as well as physical cooperation between de novo and maintenance DNA (cytosine-5) methyltransferases in human cells (*19*, *45*). Thus, it would be safe to assume that de novo activity of DNMT3s marks the genome with a limited number of methylcytosines, which in turn act as a ligand for hDNMT1 activation via binding to the allosteric domain.

Overexpression of DNMT1 (*46*, *47*) or recruitment of oncogenic transcriptional factors such as PML-PAR (*48*) leads to hypermethylation of normally unmethylated CpG islands. In cancer cells a number of CpG islands are methylated (*49*), suggesting that allosteric activation may be responsible for methylation spreading in a way similar to FMR-1 substrate in the absence of insulator and modulator proteins. This in turn attracts methyl-CpG binding proteins (MECP) and other chromatin modeling factors such as histone deacetylase (HDAC) and CAF1. As a result, transcription factors are excluded and the gene is silenced. In many cancer cell lines inactivation of the genes coding for tumor suppressor proteins, such as Rb, via mutation deletion or methylation is observed. Rb is also capable of methyltransferase inhibition via binding to the allosteric activation domain of hDNMT1. Thus, disruption of methylated DNA binding and allosteric communication between the N- and C-terminal domains by protein factors, such as tumor suppressor p21 or Rb, might be a regulatory mechanism in a normal cell.

## ACKNOWLEDGMENT

We thank Drs. R. J. Roberts, W. Jack, T. Evans, and N. Nichols for suggestions and A. Bacolla (Texas A&M University) for advice and critical reading of the manuscript. Help and support from Dr. D. Comb is highly appreciated.

## REFERENCES

1. Wu, J. C., and Santi, D. V. (1987) *J. Biol. Chem. 262*, 4778−4786.
2. Bestor, T. H., and Verdine, G. L. (1994) *Curr. Opin. Cell Biol. 9*, 380−389.
3. Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) *Cell 76*, 357−369.
4. Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988) *J. Mol. Biol. 203*, 971−983.
5. Mertineit, C., Yoder, J. A., Taketo, T., Laird, D. W., Trasler, J. M., and Bestor, T. H. (1998) *Development 125*, 889−897.
6. Bonfils, C., Beaulieu, N., Chan, E., Cotton-Montpetit, J., and MacLeod, A. R (2000) *J. Biol. Chem. 275*, 10754−10760.
7. Pradhan, S., Bacolla, A., Wells, R. D., and Roberts, R. J. (1999) *J. Biol. Chem. 274*, 33002−33010.
8. Okano, M., Xie, S., and Li, E. (1998) *Nat. Genet. 19*, 219−220.
9. Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R., and Li, E. (1996) *Development 122*, 3195−3205.
10. Leonhardt, H., Page, A. W., Weier, H. U., and Bestor, T. H. (1992) *Cell 71*, 865−873.
11. Yen, R. W., Vertino, P. M., Nelkin, B. D., Yu, J. J., el-Deiry, W., Cumaraswamy, A., Lennon, G. G., Trask, B. J., Celano, P., and Baylin, S. B. (1992) *Nucleic Acids Res. 20*, 2287−2291.
12. Finnegan, E. J., and Dennis, E. S. (1993) *Nucleic Acids Res. 21*, 2383−2388.
13. Pradhan, S., Cummings, M., Roberts, R. J., and Adams, R. L. (1998) *Nucleic Acids Res. 26*, 1214−1222.
14. Pradhan, S., Talbot, D., Sha, M., Benner, J., Hornstra, L., Li, E., Jaenisch, R., and Roberts, R. J. (1997) *Nucleic Acids Res. 25*, 4666−4673.
15. Rountree, M. R., Bachman, K. E., and Baylin, S. B. (2000) *Nat. Genet. 25*, 269−277.
16. Chuang, L. S., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G., and Li, B. F. (1997) *Science 277*, 1996−2000.
17. Robertson, K. D., Ait-Si-Ali, S., Yokochi, T., Wade, P. A., Jones, P. L., and Wolffe, A. P. (2000) *Nat. Genet. 25*, 338−342.
18. Pradhan, S., and Kim, G.-D. (2002) *EMBO J. 21*, 779−788.
19. Kim, G. D., Ni, J., Kelesoglu, N., Roberts, R. J., and Pradhan, S. (2002) *EMBO J. 21*, 4183−4195.
20. Kumar, S., Cheng, X. D., Klimasauskas, S., Sha, M., Posfai, J., Roberts, R. J., and Wilson, G. G. (1994) *Nucleic Acids Res. 22*, 1−10.
21. Margot, J. B., Aguirre-Arteta, A. M., Di Giacco, B. V., Pradhan, S., Roberts, R J., Cardoso, M. C., and Leonhardt, H. (2000) *J. Mol. Biol. 297*, 293−300.
22. Zimmermann, C., Guhl, E., and Graessmann, A. (1997) *Biol. Chem. 378*, 393−405.
23. Pradhan, S., and Roberts, R. J. (2000) *EMBO J. 19*, 2103−2114.
24. Bacolla, A., Pradhan, S., Roberts, R. J., and Wells, R. D. (1999) *J. Biol. Chem. 274*, 33011−33019.
25. Dubey, A. K., and Roberts, R. J. (1992) *Nucleic Acids Res. 20*, 3167−3173.
26. Segel, I. H. (1975) in *Enzyme Kinetics: Behaviour and Analysis of Rapid Equlibrium and Steady-State Enzyme System*, pp 346−385, John Wiley & Sons, New York.
27. Fuks, F., Burgers, W. A., Brehm, A., Hughes-Davies, L., and Kouzarides, T. (2000) *Nat. Genet. 24*, 88−91.
28. Glickman, J. F., Pavlovich, J. G., and Reich, N. O. (1997) *J. Biol. Chem. 272*, 17851−12857.
29. Araujo, F. D., Croteau, S., Slack, A. D., Milutinovic, S., Bigey, P., Price, G. B, Zannis-Hadjopoulos, M., and Szyf, M. (2000) *J. Biol. Chem. 276*, 6930−6939.
30. Sethmann, S., Ceglowski, P., Willert, J., Iwanicka-Nowicka, R., Trautner, T. A., and Walter, J. (1999) *EMBO J. 18*, 3502−3508.
31. Fatemi, M., Hermann, A., Pradhan, S., and Jeltsch, A. (2001) *J. Mol. Biol. 309*, 1189−1199.
32. Bacolla, A., Pradhan, S., Larson, J. E., Roberts, R. J., and Wells, R. D. (2001) *J. Biol. Chem. 276*, 18605−18603.
33. Iida, T., Suetake, I., Tajima, S., Morioka, H., Ohta, S., Obuse, C., and Tsurimoto, T. (2002) *Genes Cells 10*, 997−1007.
34. Christman, J. K., Sheikhnejad, G., Marasco, C. J., and Sufrin, J. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7347−7351.
35. Walsh, C. P., Chaillet, J. R., and Bestor, T. H. (1998) *Nat. Genet. 2*, 116−117.
36. Kho, M. R., Baker, D. J., Laayoun, A., and Smith, S. S. (1998) *J. Mol. Biol. 275*, 67−79.
37. Bestor, T. H. (1992) *EMBO J. 11*, 2611−2617.
38. Makalowski, W. (2001) *Acta Biochim. Pol. 48*, 587−598.
39. Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., McCune, R. A., and Gehrke, C. (1982) *Nucleic Acids Res. 10*, 2709−2721.
40. Santos, F., Hendrich, B., Reik, W., and Dean, W. (2001) *Dev. Biol. 241*, 172−182.

41. Reik, W., Dean, W., and Walter, J. (2001) *Science 293*, 1089−1093.
42. Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P., and Jaenisch, R. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 5237−5242.
43. Yokochi, T., and Robertson, K. D. (2002) *J. Biol. Chem. 277*, 11735−11745.
44. Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) *Cell 99*, 247−257.
45. Rhee, I., Bachman, K. E., Park, B. H., Jair, K. W., Yen, R. W., Schuebel, K. E., Cui, H., Feinberg, A. P., Lengauer, C., Kinzler, K. W., Baylin, S. B., and Vogelstein, B. (2002) *Nature 416*, 552−556.
46. Belinsky, S. A., Nikula, K. J., Baylin, S. B., and Issa, J. P. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 4045−4050.
47. Vertino, P. M., Yen, R. W., Gao, J., and Baylin, S. B. (1996) *Mol. Cell. Biol. 16*, 4555−4565.
48. 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.
49. Costello, J. F., Fruhwald, M. C., Smiraglia, D. J., Rush, L. J., Robertson, G. P., Gao, X., Wright, F. A., Feramisco, J. D., Peltomaki, P., Lang, J. C., Schuller, D. E., Yu, L., Bloomfield, C. D., Caligiuri, M. A., Yates, A., Nishikawa, R., Su Huang, H., Petrelli, N. J., Zhang, X., O'Dorisio, M. S., Held, W. A., Cavenee, W. K., and Plass, C. (2000) *Nat. Genet. 25*, 132−138.

BI034160+