The DNMT1 Target Recognition Domain Resides in the N Terminus*

Felipe D. Araujo‡, Sylvie Croteau§, Andrew D. Slack§, Snezana Milutinovic§, Pascal Bigey§, Gerald B. Price, Maria Zannis-Hajopoulos‡, and Moshe Szyf§

From the Departments of ‡Pharmacology and Therapeutics, §Biochemistry, and ¶Experimental Medicine, McGill University, Montreal, PQ, H3G 1Y6, Canada

DNA-cytosine-5-methyltransferase 1 (DNMT1) is the enzyme believed to be responsible for maintaining the epigenetic information encoded by DNA methylation patterns. The target recognition domain of DNMT1, the domain responsible for recognizing hemimethylated CGs, is unknown. However, based on homology with bacterial cytosine DNA methyltransferases it has been postulated that the entire catalytic domain, including the target recognition domain, is localized to 500 amino acids at the C terminus of the protein. The N-terminal domain has been postulated to have a regulatory role, and it has been suggested that the mammalian DNMT1 is a fusion of a prokaryotic methyltransferase and a mammalian DNA-binding protein. Using a combination of in vitro translation of different DNMT1 deletion mutant peptides and a solid-state hemimethylated substrate, we show that the target recognition domain of DNMT1 resides in the N terminus (amino acids 122–417) in proximity to the proliferating cell nuclear antigen binding site. Hemimethylated CGs were not recognized specifically by the postulated catalytic domain. We have previously shown that the hemimethylated substrates utilized here act as DNMT1 antagonists and inhibit DNA replication. Our results now indicate that the DNMT1-PCNA interaction can be disrupted by substrate binding to the DNMT1 N terminus. These results point toward new directions in our understanding of the structure-function of DNMT1.

Vertebrate genomes are modified by methylation of ~60–80% of the cytosines residing at the CG dinucleotide sequence (1). The distribution of methylated cytosines is not random, resulting in gene- and tissue-specific patterns of methylation (2). A large body of evidence supports the hypothesis that both methylation patterns and activity of DNA methyltransferases (DNMTs) play critical roles in development and in controlling genome functions such as differential gene expression, chromosome imprinting, and X-chromosome inactivation (3–5). It has also been suggested that DNMT1 is a downstream effector of many oncogenic pathways and a potential target for anticancer therapy (6–11). We have previously demonstrated that inhibition of DNMT1 leads to an inhibition of DNA replication (12). Recently, it has been shown that DNMT1 is able to form a complex with Rb, E2F, and HDAC1 and repress E2F-responsive expression (13, 14). Furthermore, it has been shown that DNMT1 can establish a transcriptional repressive complex with HDAC2 and DMAP1 at replication foci (15). These data suggest that DNMT1 has multiple functions in the cell. However, because the DNMT1 target recognition domain is unknown it is not possible to determine how these multiple functions and protein-protein interactions relate to its target specificity.

If DNA methylation patterns contain significant information, there must be a mechanism that ensures its proper inheritance in cell lineages. Razin and Riggs (16) have proposed that patterns of methylation are inherited, because DNMT1 is more proficient in methylating hemimethylated DNA than nonmethylated DNA. This hypothesis has been verified by a number of experiments (17, 18). Another level of specificity is the ability of DNMT1 to recognize CG sequences almost exclusively (19, 20). Thus, DNMT1 exhibits both substrate and sequence specificity.

The mammalian DNMT1 is a protein postulated to be composed, based on its similarity to other cytosine DNA methyltransferases, of at least three structural components (21–26). These domains are as follows: a catalytic domain at the C terminus, an N-terminal domain that is responsible for localization of the protein to the nucleus and replication foci, and another poorly characterized central domain. It is unclear as yet which segment is responsible for determining its specificity for hemimethylated CG sequences. Previous reports have shown that when the N-terminal domain is cleaved by proteolysis, the enzyme loses its ability to discriminate between hemimethylated and unmethylated DNA (23). It has been suggested that the N-terminal domain performs a regulatory role by inhibiting the de novo methylation activity of the C-terminal domain (23). Recently, it has been demonstrated that a mouse prokaryotic methyltransferase hybrid, DNMT1-HhaI, containing the intact N terminus of DNMT1 and most of the coding sequence of HhaI, has a 2.5-fold preference for hemimethylated DNA, whereas HhaI by itself has preference for unmethylated DNA (27). This finding indicates that the N terminus is responsible for binding specificity to hemimethylated DNA. DNA binding activity has been shown to reside within the N-terminal domain, but this binding activity does not show sequence specificity, and it has been suggested that it determines the distance traversed between replication and methylation (28).

To dissect the role that DNMT1 plays in different biological functions using structure-function analysis, one has to determine which domain of the DNMT1 is responsible for target recognition. Identifying the target recognition domain of the enzyme is also critical for developing direct inhibitors of this...
enzyme as potential anticancer agents. To map the target recognition domain, we utilized a solid-state hemimethylated DNMT1 substrate to test the binding affinities of in vitro-translated DNMT1 deletion mutant peptides. This method enabled us to determine which segment of DNMT1 per se is responsible for target recognition. It has been previously demonstrated that these modified hairpin oligonucleotide substrates are able to form a stable complex with DNMT1 and inhibit its activity (29). Furthermore, inhibition of DNMT1 with these modified hairpin oligonucleotides results in both inhibition of DNA replication (12) and transcriptional up-regulation of the tumor suppressor p21 (30). However, the mechanism by which inhibition of DNA replication occurs remains unclear. In this manuscript, we show that the target recognition domain of DNMT1 does not reside in the previously defined catalytic domain but in the N-terminal region of DNMT1, between amino acids 122 and 417. This result demonstrates the fundamental disparity between mammalian and bacterial DNA methyltransferases and the inadequacy of amino acid sequence-sequencing similarities per se in predicting the functional role of protein domains. The identification of the target recognition domain in the N terminus points toward new directions in our understanding of the structure-function relationship of mammalian DNA methyltransferases.

**EXPERIMENTAL PROCEDURES**

**Generating DNMT1 Deletion Mutant Constructs—**DNMT1 cDNA deletion mutants were generated by RT-PCR (31) from 1 µg of total RNA prepared from the human small cell lung carcinoma cell line H446 (ATCC number HTB-171) using the following set of primers: 5′-cctcgggtggcaggaa 3′ (sense) and 5′-gtatcgcggcttc 3′ (antisense) for codons 1 to 125; 5′-gcagcaggaataaag 3′ (sense) and 5′-gagtsgtggggtgctgg 3′ (antisense) for codons 122 to 170; 5′-gcagccccagagagagggg 3′ (sense) and 5′-gtcctgcagcttc 3′ (antisense) for codons 1113 to 1616; 5′-tatcggagggctacatt 3′ (sense) and 5′-cccttccctttgtttccagggc 3′ (antisense) for codons 122 to 1112; 5′-tatcggagggctacatt 3′ (sense) and 5′-cccttccctttgtttccagggc 3′ (antisense) for codons 122 to 1112. Amplification conditions were as follows: 95°C for 0.5 min, 60°C for 0.5 min, and 68°C for 5 min for 30 cycles using Promega Taq polymerase. The PCR products were cloned in a pCR3.1 vector (Invitrogen), and the sequence of the DNA was verified through the dye-deoxy chain termination method (32) using a T7 DNA sequencing kit (Amersham Pharmacia Biotech). To generate construct M, we cleaved EcoRI bearing codons 1113–1616 with EcoRI, and the fragment was blunted and ligated to a pCR3.1 vector bearing codons 122 to 1212, which was cleaved at the 3′ EcoRI site. Construct FTR bears the RT-PCR product encoding codons 122–652 in the EcoRI site of pCR3.1. Constructs 5′- and 3′-FTR were generated by digestion of construct FTR with BstEI and NotI. The resulting fragments were blunted and ligated separately into pCR3.1. Construct N-FTR bears an RT-PCR product encoding codons 122 to 1112 inserted in the EcoRI site of pCR3.1. To generate construct N-CAT, the blunt fragment bearing codons 1113–1616 was inserted into the EcoRV site of a pCR3.1 bearing codons 122–652. To generate construct CAT, we inserted the 1113–1616 coding fragment into the EcoRV site of a pCR3.1 bearing codons 122 to 1652. To assay binding to the hairpin-bound beads, 3 µl of in vitro-translated polypeptides were preincubated in a 30-µl reaction mixture including nonspecific or specific competitor oligonucleotides (100 µM), 10 mM Tris-HCl, 1 mM EDTA (pH 7.4), 0.1% SDS solution and boiled for 5 min. The supernatant (fraction S, corresponding to unbound proteins) and boiled fraction (fraction B, corresponding to bound proteins) were loaded onto an SDS-PAGE gel, dried, and exposed to autoradiography. The autoradiograms were scanned, and the amount of proteins in each fraction was quantified, and the percentage of protein stably bound to the beads was calculated as boiled (boiled/supernatant) × 100.

**Cell Culture and Transient Transfections—**HEK 293 cells, a human adenosine type 5 transformed human embryonal kidney cell line (35) (ATCC number CRL 1573), were grown in Dulbecco's modified Eagle's medium (high glucose) supplemented with 10% fetal calf serum and 2 mM glutamine. For transient transfection experiments, HEK 293 cells were plated 18 h prior to transfection at a density of 7 × 10^5 cells/100-mm tissue culture dish. The cells were preincubated with 100 nM final concentration of competitor oligonucleotides (l00 µM) for 1 h before addition of the PCNA antibody. After the overnight incubation, the beads were spun and washed three times with 500 µl of the same buffer. The final concentration of the oligonucleotide was 150 pmol per mg. To assay binding to the hairpin-bound beads, 3 µl of in vitro-translated polypeptides were preincubated in a 30-µl reaction mixture including nonspecific or specific competitor oligonucleotides (100 µM), 10 mM Tris-HCl, 1 mM EDTA (pH 7.4), 0.1% SDS solution and boiled for 5 min. The supernatant (fraction S, corresponding to unbound proteins) and boiled fraction (fraction B, corresponding to bound proteins) were loaded onto an SDS-PAGE gel, dried, and exposed to autoradiography. The autoradiograms were scanned, and the amount of proteins in each fraction was quantified, and the percentage of protein stably bound to the beads was calculated as boiled (boiled/supernatant) × 100.

**Immunoprecipitation and Western Blot Analysis**—Nuclear extracts were isolated as described previously (37). For the immunoprecipitation assay, 400 µl of nuclear extracts were incubated with 10 µl of the agarose-conjugated PCNA antibody (PC16; Santa Cruz Biotechnology) at 4°C overnight. For the competition experiments the reactions were preincubated with 100 nM final concentration of competitor oligonucleotides (Sc and H1) for 1 h before addition of the PCNA antibody. After the overnight incubation, the beads were spun and washed three times with phosphate-buffered saline. The immune complexes were resolved by polyacrylamide gel electrophoresis. After transferring to a nitrocellulose membrane and blocking the nonspecific binding sites with 5% milk, Xpress-FTR protein was detected using mouse anti-Xpress antibody. The immune complexes were washed and resuspended in 500 µl of the same buffer. The final concentration of the oligonucleotide was 100 pmol per µl. After 30 min of preincubation with competitor oligonucleotide, the mixture was applied to 500 µg of hairpin-bound beads (mix 1/1 of oligonucleotides B1- and B2-coated beads) for 5 min. The beads and supernatant were separated by a magnet, and the beads were washed twice with 500 µl of 10 mM Tris-HCl, 1 mM EDTA (pH 7.4) buffer followed by one wash in a 1 M NaCl-containing buffer. The beads were resuspended in 30 µl of 10 mM Tris-HCl, 1 mM EDTA (pH 7.4), 0.1% SDS solution and boiled for 5 min. The supernatant (fraction S, corresponding to unbound proteins) and boiled fraction (fraction B, corresponding to bound proteins) were loaded onto an SDS-PAGE gel, dried, and exposed to autoradiography. The autoradiograms were scanned, and the amount of proteins in each fraction was quantified. The amount and the percentage of protein stably bound to the beads was calculated as boiled (boiled/supernatant) × 100.

**Nuclear Extract Binding Assay—**Nuclear extracts were isolated as previously described (37). The agarose-conjugated PCNA antibody was blocked with 500 µg of hairpin-bound beads (mix 1/1 of oligonucleotides B1- and B2-coated beads) for 5 min. The beads and supernatant were separated by a magnet, and the beads were washed twice with 50 µl of 10 mM Tris-HCl, 1 mM EDTA (pH 7.4) buffer followed by one wash in a 1 M NaCl-containing buffer. The beads were resuspended in 30 µl of 10 mM Tris-HCl, 1 mM EDTA (pH 7.4), 0.1% SDS solution and boiled for 5 min. The supernatant (fraction S, corresponding to unbound proteins) and boiled fraction (fraction B, corresponding to bound proteins) were loaded onto an SDS-PAGE gel, dried, and exposed to autoradiography. The autoradiograms were scanned, and the amount of proteins in each fraction was quantified. The amount and the percentage of protein stably bound to the beads was calculated as boiled (boiled/supernatant) × 100.
RESULTS

Coupled in Vitro Transcription/Translation of Recombinant Human DNMT1 Deletion Mutants—To map the DNMT1 target recognition domain, DNMT1 deletion mutant peptides were created using a coupled in vitro transcription/translation rabbit reticulocyte system (Fig. 1, A and B). The constructs were designed as follows: DNMT1 bears the entire coding sequence of the enzyme starting at the upstream ATG initiation site (34), construct M bears the coding sequence of the DNMT1 starting at the downstream ATG initiation site (33), construct PS bears a single base mutation converting the previously proposed catalytic site proline-cysteine dipeptide to proline-serine (24), construct N-FTR bears the N-terminal fork-targeting region and the central domain region contained within amino acids 122–1112, construct N-CAT bears the N-terminal and catalytic domains with a deletion of the central domain from amino acid 652 to 1113, construct CAT encodes the entire catalytic domain as proposed by homology to bacterial cytosine methyltransferases plus the N-terminal portion responsible for nuclear localization (21, 23, 24), construct FTR bears the N-terminal region between amino acids 122 and 652, construct 5′-FTR is a deletion of FTR containing the region between amino acids 122 and 417, and construct 3′-FTR contains the region between amino acids 417 and 731 (Fig. 1A). We utilized a luciferase construct as a control (Fig. 1A). Following coupled in vitro transcription/translation in the presence of [35S]-methionine, the peptides were size-fractionated by SDS-PAGE and visualized by autoradiography (Fig. 1B). All the peptides migrated according to their expected sizes.

Oligonucleotide Design—Hemimethylated and nonmethylated CG-containing phosphorothioate hairpin oligonucleotides were designed as probes and competitors for our binding assays (Fig. 2). Because oligonucleotides B1 and B2 are biotinylated at the 5′ arm of the hairpin, they can be conjugated to streptavidin-coated magnetic beads and used as solid-state probes for peptide binding experiments. The remaining oligonucleotides (Sc, H1, and N1) were utilized as competitors. The 5′ arm of the hairpins containing methylated cytosines behaves as the methylation-guiding parental strand and the 3′ arm behaves as the methyl-acceptor nascent strand of replicating DNA. The hairpin substrates we utilized contained an inosine (I) instead of the methyl acceptor, cytosine. These substrates have been previously shown to form a high affinity complex with DNMT1 that could be dissociated only by boiling (29). Results from these experiments have also indicated that the inosine (I) group substitution, as well as synthesis of the backbone with a phosphorothioate modification, results in increased hairpin binding affinity to DNMT1 (29). Moreover, the CG sites present in the hairpins are separated by 4 bases to avoid tandem CGs, which have been demonstrated to be poor substrates for DNA methyltransferases (29).

The DNMT1 Target Recognition Domain Resides in the N-terminal Region—Previous studies have shown that DNMT1 forms a low affinity complex with both hemimethylated and nonmethylated DNA (19). It has been proposed that the initial binding of DNA to DNMT1 with DNA does not discriminate between hemimethylated and nonmethylated DNA (19) and that discrimination between the two substrates occurs during catalysis (19). In our binding assays (see Figs. 3–5), the ability of a polypeptide to form a stable complex with the magnetic bead-conjugated substrate is measured by comparing the abundance of [35S]-labeled polypeptide in the supernatant fraction (indicated as S in Figs. 3–5) versus the abundance in the fraction eluted by boiling the hairpin-bound beads (indicated as B in Figs. 3–5). As observed (Fig. 3A), in the absence of any competitor, peptide M is present exclusively in the bound fraction (B), whereas a luciferase peptide, used as a negative control, is unable to form a complex with the hemimethylated solid-state
substrate and is therefore present exclusively in the supernatant fraction (S).

To test the binding specificity of the M peptide for the hemimethylated solid-state substrate, we incubated [35S]methionine-labeled in vitro-translated M peptide with the substrate in the presence of increasing concentrations of either a hemimethylated (H1) or a nonspecific (Sc) oligonucleotide. As observed (Fig. 3B), in the absence of any competitors (N), M is able to form a stable complex with the hemimethylated hairpin. This is indicated by the presence of M exclusively in fraction B (Fig. 3B). Complex formation is not challenged by an excess of 100 μM nonspecific oligonucleotide (Sc), because M remains exclusively in fraction B. Conversely, complex formation is partly abolished by a challenge with 10 μM of cold hemimethylated competitor (H1) and is completely abolished by a challenge with 100 μM cold competitor, as indicated by the disappearance of M from B and its presence in S.

To determine which domain of DNMT1 interacts specifically with the hemimethylated substrate, we incubated the different [35S]methionine-labeled recombinant polypeptides with the solid-state hemimethylated substrate in the presence of 100 μM of either hemimethylated (H1), nonmethylated (N1) CG-bearing hairpin oligonucleotides or a nonspecific competitor (Sc) (Fig. 2). As observed in the autoradiograms (Fig. 4A) and in the graphical representations (Fig. 4B), all the polypeptides form a
stable complex with the solid-state substrate in the absence of any competitors. This is indicated by their presence in fraction B, suggesting that both C-terminal and N-terminal domains can recognize and form a stable complex with the substrate. However, the binding of the catalytic domain peptide (CAT) to the hemimethylated substrate is completely abolished by nonspecific (Sc), as well as specific, competitors (H1 and N1). This is evident from the presence of the peptide exclusively in fraction S and its complete absence from fraction B. The binding of peptides bearing either the N-terminal domain (N-FTR, N-CAT, FTR, and 5’-FTR) or the complete protein (M and DNMT1) is not competed by the nonspecific competitor (Sc) (>80% remain bound) but is abolished (<40% remain bound) by the hemimethylated competitor (H1). The nonmethylated CG-bearing hairpin (N1) is a partial competitor as indicated by the distribution of the polypeptide in both the bound and unbound fractions. In all of these cases the hemimethylated hairpin is 2–4-fold more effective in competing out binding to the solid-state substrate than the nonmethylated homolog (Fig. 4B). PS also shows a preference for the hemimethylated substrate, but interestingly it is competed less efficiently by the H1/N1 competitor pair, indicating the formation of a tighter complex. The fact that the competition profiles of DNMT1 and N-CAT are similar indicates that the central region of the enzyme (amino acids 652–1113) is not required for target specificity. This is confirmed by the similarity of competition profile of these two peptides with the FTR peptide competition profile. Moreover, results indicating that binding of F-CAT to the hemimethylated substrate is competed as efficiently by the nonspecific competitor (Sc) as it is by the hemimethylated competitor (H1) further confirm that the central region is not essential for target recognition. The 5’-FTR peptide still retains hemimethylated binding specificity, which is indicated by the fact that H1 competes for binding more efficiently than N1 and Sc. Interestingly, the binding specificity of 3’-FTR peptide for the hemimethylated substrate is weaker than the more inclusive FTR peptides, because all competitors have a similar effect on binding. Taken together, these results support the hypothesis that the target recognition domain of DNMT1 resides in the N-terminal segment (amino acids 122–417) of the protein, overlapping with the fork-targeting region.

We then tested whether the FTR peptide still binds specifically to hemimethylated DNA when both CAT and FTR peptides are present in the same reaction, as is the case with the natural product. We tested their affinity to hemimethylated DNA under two conditions, in the absence or in the presence of competitor DNA. The presence of competitor DNA mimics the in vivo scenario, where DNMT1 has to discriminate between its specific target and the bulk of non-CG DNA. The results (Fig. 5A) demonstrate that in the absence of any competitors both CAT and FTR peptides can bind the target, in agreement with results from Fig. 4, showing that both peptides have high affinity to DNA. However, in the presence of a nonspecific competitor (Fig. 5B), as is the case when DNMT1 interacts with the genome in vivo, only FTR is bound to hemimethylated DNA. Having both CAT and FTR peptides concurrently also verifies that the differences observed between distinct peptides in the competition experiments depicted in Fig. 4 are not the result of impurities in the in vitro translation reaction of one peptide versus another. For clarity only the bound fractions were loaded.

PCNA/FTR Complex Can Be Disrupted by a Hairpin Oligonucleotide—Previous reports have shown that modified hairpin oligonucleotides can work as bona fide antagonists of DNMT1 and inhibit DNA replication (12, 29). Because the target recognition of DNMT1 may overlap with its PCNA binding site, we tested the hypothesis that the DNMT1/PCNA complex could be disrupted by FTR binding to the modified hairpin oligonucleotide. We transfected HEK 293 cells with an Xpress-tagged FTR expression vector followed by PCNA immunoprecipitations in the presence of a nonspecific oligonucleotide (Sc), a hemimethylated hairpin (H1), or no competitor (N). The results (Fig. 6A, top panel) indicate that the PCNA/DNMT1 complex can be disrupted by H1 but not by Sc. The presence of IgG and PCNA in all the immunoprecipitations was also tested (Fig. 6A, middle and bottom panels). We also immunoblotted the supernatants of these immunoprecipitations to confirm FTR expression (Fig. 6A, bottom panel). Because the previous experiments

![Fig. 5. Differential interaction of the catalytic and N terminus regions of the DNMT1 with the hemimethylated substrate when present concurrently. A, FTR and CAT peptides were in vitro-transcribed and translated, and different ratios of both peptides were incubated at room temperature with avidin-coated magnetic beads bound to biotinylated hemimethylated hairpins (Bl and B2) in the absence of any competitors, as described under “Experimental Procedures.” The hairpin-bound (B) and unbound (S) fractions were separated on SDS-PAGE and exposed to autoradiography. B, a similar experiment was performed but with the addition of nonspecific competitor (Sc) where indicated. The bound and unbound fractions were separated as above. The arrows indicate the expected positions of FTR and CAT peptides. The plus sign (+) indicates presence of the indicated peptide, the minus sign (−) indicates absence of the indicated peptide, and the triangle indicates increasing concentrations of the indicated peptides.](image1)

![Fig. 6. The FTR/PCNA complex can be disrupted by hairpin oligonucleotides. A, an Xpress-tagged FTR construct was transiently transfected into HEK 293 cells and immunoprecipitated utilizing a PCNA antibody in the presence of no competitor (N), a hemimethylated competitor (H1), or a nonspecific competitor (Sc). Western blot analysis of FTR, IgG, and PCNA of immunoprecipitated fractions was performed as indicated. Western blot analysis of the FTR present in supernatant fractions (Sup.) was also performed. B, nuclear extracts from Xpress-tagged FTR-transfected HEK 293 cells were incubated with avidin-coated magnetic beads and bound to biotinylated hemimethylated hairpins Bl and B2. The hairpin-bound (B) and unbound (S) fractions were separated on SDS-PAGE and analyzed by Western blot with an anti-Xpress antibody.](image2)
had been performed using in vitro-translated peptides, we
tested the ability of the FTR peptide to bind the hairpin solid-
state substrate in HEK 293 nuclear extracts. The results (Fig.
6B) demonstrate that the FTR peptide does indeed bind the
solid substrate under these conditions, as indicated by the FTR
presence in the bound fraction (B). These results suggest that
the disruption of the PCNA/DNMT1 complex might be a pos-
sible mechanism by which these modified hairpin oligonucleo-
tides inhibit DNA replication.

DISCUSSION

Recognition of hemimethylated CGs, the DNMT1 target se-
quence, is a critical step in the replication of the DNA methy-
lation pattern and the epigenetic information that it encodes.
Although the first mammalian dnmt has been cloned and its
cDNA sequenced more than a decade ago (21, 22), the domain
responsible for its target recognition has not yet been clearly
defined. Based on sequence homology between mammalian and
bacterial cytosine methyltransferases the entire catalytic do-
main, including the target recognition domain, has been pro-
posed to reside in the C terminus of the protein (24). The
additional N-terminal and central domains of the protein were
proposed to perform regulatory roles (23, 28). This hypothesis
is based on the assumption that the structure of the mamma-
lain DNMT1 follows the rules laid out in bacteria and that
DNMT1 is an evolutionary hybrid of a primordial DNMT plus
at least two additional regulatory protein modules (23).
Our results are in agreement with recent target specificity studies
utilizing a mouse prokaryotic hybrid DNMT, DNMT1-HhaI,
containing the intact N terminus of DNMT1 and most of the
coding sequence of HhaI, demonstrating that it has a 2.5-fold
preference for hemimethylated DNA, whereas HhaI by itself
has preference for unmethylated DNA (27). Moreover, struc-
tural analysis of de novo DNMTs, DNMT3a and DNMT3b (4,
38), which lack the DNMT1 N-terminal region and recognize
nonmethylated CGs, further supports the hypothesis that the
hemimethylated target recognition domain resides within the
DNMT1 N terminus.

To delineate the DNMT1 target recognition domain, we per-
formed experiments utilizing a solid-state hemimethylated
substrate and in vitro-translated deletion mutant peptides.
Competition assays with hemimethylated and nonmethylated
substrates revealed that the N-terminal domain specifically
recognizes a hemimethylated CG substrate. The previously
described catalytic domain, which has been proposed to contain
all the elements required for catalytic activity including target
recognition (24), binds DNA but does not exhibit specificity to
hemimethylated CG duplex DNA as determined by competition
assays. Moreover, when the catalytic domain and the N-termi-
nal domain are concurrently present in the reaction with non-
specific DNA, the substrate binds to the N-terminal domain
suggesting that it bears the target recognition domain. Inter-
estingly, the PS mutant is competed less efficiently by the
HI/N1 competitor pair, indicating the formation of a tighter
complex. This result is supported by a previous study indicat-
ing that EcoRII DNMT mutants with substitution of the con-
served cysteine for serine bind tightly to DNA (39). These
mutants resemble the wild-type enzyme in that their binding
to substrate is not eliminated by the presence of nonspecific DNA
in the reaction (39).

Deciphering whether the catalytic components of DNMT1
reside entirely in the C-terminal domain or whether they reside
elsewhere is critical for both structure-function and crystal
structure analysis of the protein, as well as for drug design.
Identifying the location of the different components of the cat-
alytic domains of DNMTs is also critical for interpretation of
knockout experiments by homologous recombination, which
target putative catalytic domains (5). If these alleles, which are
inactivated at the C terminus, still maintain the sequence-
specific target recognition domain some of the previously de-
scribed biological effects of this knockout could be attributed to
the DNA binding effects of the truncated protein rather than in-
hibition of methylation.

The catalytic domain should, by definition, include the sub-
strate recognition domain, otherwise there is no catalysis. Our
results shed doubt on the previous designation of the DNMT1
C terminus as the exclusive catalytic domain. The data pre-
SENTED in this paper suggest that this conclusion might be
revisited. In contrast to many bacterial DNMTs (40), the target
recognition domain of the mammalian DNMT1 resides at the N
terminus, at a significant distance from the conserved catalytic
and AdoMet binding domains. Our data is consistent with the
hypothesis that the N-terminal is not just a regulatory domain,
as has been previously suggested (23, 41), but it is the sub-
strate recognition module of the enzyme. Therefore, the cata-
lytic component of the vertebrate enzyme is composed of two
modular domains, a target recognition domain in the N termi-
nus, in addition to the previously described AdoMet binding
and methyl-transfer domains in the C terminus of the enzyme.
The fact that the target recognition domain and methyl-trans-
fer domain are located at such a distance on the linear amino
acid sequence must not necessarily translate to a similar dis-
tance in the tertiary structure of the enzyme. The results
presented here might explain the failure of previous attempts
to demonstrate DNA methylation activity in the previously
postulated catalytic domain of the enzyme (41). Despite the
striking structural homology of the C terminus of the verte-
brate to bacterial DNMTs, its expression in bacterial or mam-
malian cells does not result in DNA methylation activity (26).
These results illustrate the risk in predicting biochemical func-
tions exclusively from sequence homology.

We had previously demonstrated that the hairpin oligonu-
cleotides utilized here can form a stable complex with DNMT1
(29), inhibit its activity, and inhibit DNA replication in parallel
with a transcriptional up-regulation of the tumor suppressor p21
(12, 30). With the knowledge that the PCNA interaction do-
main of DNMT1 and its target recognition domain are over-
lapping, we were able to show that the PCNA/DNMT1 complex
could be disrupted by the hemimethylated hairpins. These
results might provide an alternative mechanism by which
DNMT1 inhibition can supress oncogenic transformation (42).
Because DNMT1 is a multifunctional protein, identifying the
domains important for transformation is critical for DNMT1
anti-oncogenesis drug design. Moreover, the identification of
the target recognition domain allows a more complete struc-
ture-function analysis of DNMT1, and it will help the develop-
ment of novel direct inhibitors of this enzyme (7).

Acknowledgments—We are grateful to Methylgene Inc. for kindly
synthesizing the hairpin oligonucleotides utilized here.

REFERENCES

1. Razin, A., and Shyf, M. (1984) Biochim. Biophys. Acta 782, 331–342
2. Yisraeli, J., and Shyf, M. (1984) in Methylation: Biochemistry and Biological
Significance (Razin, A., Cedar, H., and Riggs, A. D., eds.), pp. 353–378,
Springer-Verlag New York Inc., New York
3. Siegfried, Z., and Cedar, H. (1997) Curr. Biol. 7, R305–7
4. Bird, A. P., and Wolffe, A. P. (1999) Cell 99, 451–454
5. Li, E., Beard, C., and Jaenisch, R. (1993) Nature 366, 362–365
6. Shyf, M. (1996) Pharmacol. Ther. 76, 1–37
7. Shyf, M. (1998) Cancer Metastasis Rev. 17, 219–231
8. MacLeod, A. R., Rouleau, J., and Shyf, M. (1995) J. Biol. Chem. 270,
11327–11337
9. MacLeod, A. R., and Shyf, M. (1995) J. Biol. Chem. 270, 8037–8043
10. Ramchandani, S., MacLeod, A. R., Pinard, M., von Hofe, E., and Shyf, M. (1997)
Proc. Natl. Acad. Sci. U. S. A. 94, 684–689
11. Slack, A., Cervoni, N., Pinard, M., and Shyf, M. (1999) J. Biol. Chem. 274,
10105–10112
12. Knox, J. D., Araujo, F. D., Bigey, P., Slack, A. D., Price, G. B., Zannis-
Hadjopoulos, M., and Shyf, M. (2000) J. Biol. Chem. 275, 17986–17999
