Recombinant Human DNA (Cytosine-5) Methyltransferase

I. EXPRESSION, PURIFICATION, AND COMPARISON OF DE NOVO AND MAINTENANCE METHYLATION*

(Received for publication, May 12, 1999, and in revised form, August 16, 1999)

Sriharsa Pradhan, Albino Bacolla‡, Robert D. Wells‡, and Richard J. Roberts§

From New England Biolabs, Beverly, Massachusetts 01915 and the ‡Center for Genome Research, Institute of Biosciences and Technology, Texas A & M University, Texas Medical Center, Houston, Texas 77030-3303

A method is described to express and purify human DNA (cytosine-5) methyltransferase (human DNMT1) using a protein splicing (intein) fusion partner in a baculovirus expression vector. The system produces 1 mg of intact recombinant enzyme >95% pure per 1.5 × 10⁹ insect cells. The protein lacks any affinity tag and is identical to the native enzyme except for the two C-terminal amino acids, proline and glycine, that were substituted for lysine and aspartic acid for optimal cleavage from the intein affinity tag. Human DNMT1 was used for steady-state kinetic analysis with poly(dI-dC)-poly(dI-dC) and unmethylated and hemimethylated 36- and 75-mer oligonucleotides. The turnover number \( k_{\text{cat}} \) was 131–237 h⁻¹ on poly(dI-dC)-poly(dI-dC), 1.2–2.3 h⁻¹ on unmethylated DNA, and 8.3–49 h⁻¹ on hemimethylated DNA. The Michaelis constants for DNA \( K_m \) and S-adenosyl-l-methionine (AdoMet) \( K_m^\text{AdoMet} \) ranged from 0.39–1.32 and 2.6–7.2 \( \mu \)M, respectively, whereas the ratio of \( k_{\text{cat}}/K_m^\text{AdoMet} \) ranged from 3.9 to 44 (237–336 for poly(dI-dC)-poly(dI-dC)) \( 10^6 \text{M}^{-1} \text{h}^{-1} \). The preference of the enzyme for hemimethylated, over unmethylated, DNA was 7–21-fold. The values of \( k_{\text{cat}}/k_z \) on poly(dI-dC) DNA showed a 2–3-fold difference, depending upon which strand was pre-methylated. Furthermore, human DNMT1 formed covalent complexes with substrates containing 5-fluoro-CNG, indicating that substrate specificity extended beyond the canonical CG dinucleotide. These results show that, in addition to maintenance methylation, human DNMT1 may also carry out de novo and non-CG methyltransferase activities in vivo.

Methylated cytosine is found in the genome of organisms ranging from prokaryotes to mammals (1). Methylation of DNA in eukaryotes is implicated in various biological and developmental processes, such as gene regulation (2), DNA replication (3), genomic imprinting (4), embryonic development (5), carcinogenesis (6), and genetic diseases (7). The bulk of the methylation takes place during DNA replication in the S-phase of the cell cycle (8). The maintenance methylation ensures the propagation of tissue-specific methylation patterns established during mammalian development. The methyl transfer reaction proceeds via nonspecific binding of the enzyme to DNA, recognition of the methyl group donor S-adenosyl-l-methionine (AdoMet) to the active site of the enzyme. DNA (cytosine-5) methyltransferases (m⁵C MTase) introduce a methyl group onto carbon 5 of the target cytosine through a covalent intermediate between the protein and the target cytosine (9). During this process, the cytosine is flipped 180° out of the DNA backbone into an active site pocket of the enzyme (10). After completion of the methyl transfer reaction, the products, methylated DNA and S-adenosyl-l-homocysteine (AdoHcy), are released. Previous studies on the mechanism of methylation were mainly limited to prokaryotic m⁵C MTases although some limited kinetic studies have also been reported with mouse and human DNMT1 (11, 12).

Eukaryotic m⁵C MTases have been cloned from various organisms such as mouse (13), human (14), Xenopus (15), sea urchin (16), chicken (17), Arabidopsis (18), and pea (19). The human DNMT1 has been localized to the chromosomal site 19 p13.2-p13.3 by fluorescence in situ hybridization (14). The enzyme consists of a large N-terminal region and a smaller C-terminal region linked by a run of Gly-Lys dipeptide repeats (14). The large N-terminal domain is unique to the eukaryotic m⁵C MTases. The smaller C-terminal domain has strong homology with prokaryotic m⁵C MTases and contains the elements necessary for catalysis (20). The function of the N terminus is poorly understood, and there is little sequence homology between plant and animal enzymes. Sequences within this domain have been implicated in Zn²⁺ binding (21), interaction with proliferating cell nuclear antigen (22), and targeting to replication foci (8). The full-length mouse m⁵C MTase is about 1700 amino acids in length and has a molecular mass of 183.5 kDa (23).

Prokaryotic m⁵C MTases are usually part of restriction-modification systems and rarely discriminate between unmethylated or hemimethylated DNA substrates. In contrast, mammalian m⁵C MTases show a much higher reaction velocity on a hemimethylated DNA substrate (maintenance methylation), the product of semiconservative DNA replication, than on an unmethylated DNA substrate (de novo methylation). Methylation of single-stranded DNA by the mammalian m⁵C MTase is strongly stimulated by the presence of nearby 5-methylcytosines (24). This rate of methylation was comparable with that of hemimethylated DNA (24). It has been demonstrated that the human enzyme can also methylate non-B-DNA structures such as hairpins and mismatches (12). Recently, a family of mammalian m⁵C MTases has been identified and hypothesized to serve as de novo methyltransferases (25).

The abbreviations used are: AdoMet, S-adenosyl-l-methionine; AdoHcy, S-adenosyl-l-homocysteine; m⁵C MTase, DNA (cytosine-5) methyltransferases; kbp, kilobase pair; AcNPV, Autographa californica nuclear polyhedrosis virus; DTT, dithiothreitol; PAGE, polyacrylamide gel electrophoresis; bp, base pair(s); FdC, 5-fluoro-2'-deoxyctydine; SNRPN, small nuclear riboprotein-associated peptide N; FMR, fragile X mental retardation.

* This work was supported by National Institutes of Health Grants GM46127 (to R. J. R.), NS37554, and GM52982 (to R. D. W.) and by the Robert A. Welch Foundation (to R. D. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This advertisement must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: New England Biolabs, 32 Tozer Rd., Beverly, MA 01915. Tel.: 978-927-3382; Fax: 978-921-327; E-mail: roberts@neb.com.

1 The abbreviations used are: AdoMet, S-adenosyl-l-methionine; AdoHcy, S-adenosyl-l-homocysteine; m⁵C MTase, DNA (cytosine-5) methyltransferases; kbp, kilobase pair; AcNPV, Autographa californica nuclear polyhedrosis virus; DTT, dithiothreitol; PAGE, polyacrylamide gel electrophoresis; bp, base pair(s); FdC, 5-fluoro-2'-deoxyctydine; SNRPN, small nuclear riboprotein-associated peptide N; FMR, fragile X mental retardation.
The diverse role played by DNA methylation in mammals has motivated several groups to express mammalian mC MTases in Escherichia coli (26), baculovirus (23, 27), and mammalian (COS) cells (28). In the E. coli and COS cell expression systems, the amount of protein produced was low, whereas protein expression was high using the baculovirus system. Insect cells are rich in proteases, and long purification protocols often result in partial proteolytic degradation of proteins. To overcome the problem of degradation, we have constructed and expressed the human DNMT1 as a fusion protein with a Saccharomyces cerevisiae VMA1 intein gene (29). The C terminus of the human DNMT1 was fused to the N terminus of the intein. A small chitin-binding domain from Bacillus circulans has been added to the C terminus of the intein for affinity purification. Thus, the three-part fused human DNMT1 protein has been added to the C terminus of the intein for affinity purification. Therefore, the three-part fused human DNMT1 can be immobilized using chitin beads as an affinity matrix.\(^2\) Addition of dithiothreitol (DTT) at low temperature initiates cleavage on the column between the intein and the recombinant human DNMT1.\(^2\) Whereas the intein and chitin-binding domain remain attached to the column, the purified recombinant enzyme is eluted. This purified recombinant human DNMT1 is identical with the native enzyme except for the two C-terminal amino acids. In this report, we have studied its steady-state kinetic parameters on double-stranded, unmethylated, and hemimethylated DNA substrates, which are representative of human genes. We also assayed methylation of non-CG sequences by human DNMT1.

**EXPERIMENTAL PROCEDURES**

Human DNA (Cytosine-5) Methyltransferase Transfer Vector—Human DNMT1 expression constructs were derived from pBKSHMT5.0 (b) (gift of Prof. S. Baylin, The Johns Hopkins University). This plasmid had the full-length human DNMT1 cDNA based on the previously published sequence (14). A polymerase chain reaction of the 3' end of the cDNA was used to incorporate an Smal restriction endonuclease site in place of the stop codon of the cDNA and to provide the optimal amino acids for protein cleavage (sense primer, 5'-GGAAATTCATAT-GCATGCAGGAGAAGCCGCGCAGCAG-3' and antisense primer, 5'-TTAGGCCCCGAGCCTTCTCTCTTTATTTAAC-3'). The restriction sites are underlined. This change causes the C-terminal amino acids lysine and aspartic acid to become proline and glycine. Substitution of these two amino acids at the C terminus of the mammalian DNMT1 had no effect on the methyl transfer reaction as compared with the native sequence (data not shown). This polymerase chain reaction product of about 500 bp was digested with NdeI and Smal and ligated to a baculovirus vector pVIC1 to give pVIC1MT4. This construct contains the 3' end of the human DNMT1 cDNA in frame with the intein chitin-binding domain of the transfer vector pVIC1.\(^2\) pVIC1MT4 was digested with NraI and EagI and a 4.1-kbp EcoRV-EagI fragment from pBKSHMT5.0(b) was ligated to give the functional human DNMT1 transfer vector pVIC1MT (Fig. 1). The ligated junctions in the final pVIC1MT construct were verified by DNA sequencing.

**Insect Cell Culture, Viral Transfection, and Recombinant Protein Expression**—A pupal ovarian cell line (SF9) from the worm Autographa californica nuclear polyhedrosis virus (AcNPV) DNA (PharMingen) and the transfer vector pVIC1MT. Four recombinant clones from individual plaque were isolated previously (23) and were used for fitting the data points and to obtain the kinetic constants, \(V_{max}\), \(K_{M}\), and \(K_{AdoMet}\), were reported elsewhere (32). These figures represent estimated values.

**Southern Blot Analysis**—Genomic DNA was isolated from SF9 cells 48 h post-infection with recombinant virus using the Easy DNA Kit (Invitrogen). The purified DNA was digested either by AatII, BsiEI, or SnaI, and the digested DNA fragments were separated on a 1% agarose gel in TBE buffer gel. The DNA in the gel was blotted onto Hybond-N+ nylon membranes with random-banded human DNMT1 clone. The blot was washed as described (33) and autoradiographed.

**Western Blot Analysis**—Cell extracts were mixed with SDS loading dye with 2 \(\mu\)M Tris-(2-carboxyethyl) phosphine in place of DTT, boiled at 95 °C for 5 min, and loaded on a 4–20% Tris/glycine/SDS gradient gel. Purified protein was quantitated using the Bradford assay with bovine serum albumin as a standard.

DNA (Cytosine-5) Methyltransferase Assay and Data Analysis—mC MTase assays were carried out at 37 °C for 30 min in duplicate with a total volume of 25 \(\mu\)l of reaction mix. A typical reaction contained 15 Ci/mmol, Amersham Pharmacia Biotech), substrate DNA, and enzyme in assay buffer (50 mm Tris-HCl, pH 7.8, 1 mm Na2EDTA, pH 8.0, 1 mm DTT, 7 mm phenylmethylsulfonyl fluoride, 100 mm NaCl, 50% glycerol (v/v), and 1 mm DTT. The purified protein was stored at –20 °C. The purity of the protein was checked by SDSPAGE (4–20% Tris/glycine/SDS gradient gel). Purified protein was quantitated using the Bradford assay with bovine serum albumin as a standard.

**Methylation by Recombinant Human DNA Methyltransferase**

**EXPERIMENTAL PROCEDURES**

Human DNA (Cytosine-5) Methyltransferase Transfer Vector—Human DNMT1 expression constructs were derived from pBKSHMT5.0 (b) (gift of Prof. S. Baylin, The Johns Hopkins University). This plasmid had the full-length human DNMT1 cDNA based on the previously published sequence (14). A polymerase chain reaction of the 3' end of the cDNA was used to incorporate an Smal restriction endonuclease site in place of the stop codon of the cDNA and to provide the optimal amino acids for protein cleavage (sense primer, 5'-GGAAATTCATAT-GCATGCAGGAGAAGCCGCGCAGCAG-3' and antisense primer, 5'-TTAGGCCCCGAGCCTTCTCTCTTTATTTAAC-3'). The restriction sites are underlined. This change causes the C-terminal amino acids lysine and aspartic acid to become proline and glycine. Substitution of these two amino acids at the C terminus of the mammalian DNMT1 had no effect on the methyl transfer reaction as compared with the native sequence (data not shown). This polymerase chain reaction product of about 500 bp was digested with NdeI and Smal and ligated to a baculovirus vector pVIC1 to give pVIC1MT4. This construct contains the 3' end of the human DNMT1 cDNA in frame with the intein chitin-binding domain of the transfer vector pVIC1. pVIC1MT4 was digested with NraI and EagI and a 4.1-kbp EcoRV-EagI fragment from pBKSHMT5.0(b) was ligated to give the functional human DNMT1 transfer vector pVIC1MT (Fig. 1). The ligated junctions in the final pVIC1MT construct were verified by DNA sequencing.

Insect Cell Culture, Viral Transfection, and Recombinant Protein Expression—A pupal ovarian cell line (SF9) from the worm Spodoptera frugiperda was used for co-transfection and expression of the human DNMT1. As described previously (23) with the following modification, SF9 cells were maintained as a suspension culture in TNM-FH media (JRH Biosciences) supplemented with fetal calf serum 10% (v/v) and an antibiotic/antimyocyte solution at a final concentration of 25 \(\mu\)l of reaction mix. A typical reaction contained S-adenosyl-L-[methyl-\(^3\)H]methionine (AdoMet) (specific activity 15 Ci/mmol, Amersham Pharmacia Biotech), substrate DNA, and enzyme in assay buffer (50 mm Tris-HCl, pH 7.8, 1 mm Na2EDTA, pH 8.0, 1 mm DTT, 7 mm phenylmethylsulfonyl fluoride, 5% glycerol, and 100 mm NaCl, 70% amberlite) were used for fitting the data points and to obtain the kinetic constants, \(V_{max}\), \(K_{M}\), and \(K_{AdoMet}\), were reported elsewhere (32). These figures represent estimated values.

Southern Blot Analysis—Genomic DNA was isolated from SF9 cells 48 h post-infection with recombinant virus using the Easy DNA Kit (Invitrogen). The purified DNA was digested either by AatII, BsiEI, or SnaI, and the digested DNA fragments were separated on a 1% agarose gel in TBE buffer gel. The DNA in the gel was blotted onto Hybond-N+ nylon membranes with random-banded human DNMT1 clone. The blot was washed as described (33) and autoradiographed.

Western Blot Analysis—Cell extracts were mixed with SDS loading dye with 2 \(\mu\)M Tris-(2-carboxyethyl) phosphine in place of DTT, boiled at 95 °C for 5 min, and loaded on a 4–20% Tris/glycine/SDS gradient gel. Tris-(2-carboxyethyl) phosphine is a strong reducing agent and does not take part in the cleavage process (34). The proteins were blotted on an Immobilon-P membrane and probed with an anti-intein antibody

\(^2\) M. E. Scott, S. Pradhan, and M. Q. Xu, submitted for publication.
Methylation by Recombinant Human DNA Methyltransferase

5-Fluoro-2'-deoxycytidine Assay—Duplex oligonucleotides containing 5-fluoro-2'-deoxycytidine (FdC) were synthesized at New England Biolabs. The full-length duplexes were purified on a 15% polyacrylamide gel, stained with ethidium bromide, and excised. The resulting bands were excised and purified by ethanol precipitation. The purified oligonucleotides were then digested with restriction endonucleases and analyzed by electrophoresis on a 1% agarose gel. The gel was stained with ethidium bromide and visualized under UV light. The band corresponding to the 36 bp product was cut out and eluted from the gel. The eluted product was then analyzed by sequencing to confirm the presence of the FdC modification.

RESULTS

Expression and Purification of the Biologically Active Human DNA (Cytosine-5) Methyltransferase—Co-transfection of the human DNMT1 transfer vector, pVICHMT, with linear AcNPV DNA (BaculoGold DNA, PharMingen) resulted in homologous human DNMT1 transfer vector, pVICHMT, with linear AcNPV genome. pVICHMT is shown schematically. A 4.85-kbp cDNA containing human DNMT1, indicated by the gray shading, was fused to the N-terminal of the VMA1 intein-chitin binding domain, as shown by the hatched line. Two flanking black arrows corresponding to AcNPV open reading frame (ORF) 603 and 1629 are also indicated. The Amp and ori loci are indicated with black shading. The recipient AcNPV genome has deletions of the following sequences, C-terminal ORF1629, polyhedrin promoter (pPH), and the polyhedrin open reading frame. The N-terminal intein cleavage site APG/C, which is about 10 bp, is also shown. The regions available for homologous recombination are indicated. B, a schematic of VICHMT8, a recombinant virus following homologous recombination. The polyhedrin promoter lies upstream of the DNMT1-intein-CBD fusion genes. The full-length cDNA was used as a probe to identify the correct homologous recombinants. The restriction sites are indicated as EcoRI (S), SmaI (S), and BstEI (E). Southern blot analysis of the DNA from the recombinant clone VICHMT8. The DNA band released by SmaI contains the whole cDNA plus the intein-CBD, which is about 10 kbp. The other two enzymes release the expected fragments of 3.5 and 8.0 kbp for EcoRI and 2.5 and 6.5 kbp for BstEI. Minor bands represent defective viral DNAs resulting from illegitimate recombination (30). D, time course of human DNMT1 expression using VICHMT8. Infected cell extracts were analyzed on an SDS gel, blotted, and probed with anti-intein antibody. A 235-kDa band appears in all lanes. The day of harvest is indicated at the top of each lane. Control lanes with non-infected SF9 cell extract are labeled C. M indicates the biotinylated protein molecular weight markers. BSA, bovine serum albumin.

been reported to express between 5- and 10-fold higher protein levels than SF9 cells (35), but for human DNMT1 the protein expression did not improve (data not shown). All further experiments used the SF9 cell line for expression. To follow expression and find optimal conditions for production, extracts of the
human DNMT1 clones were assayed by Western blots using anti-intein antibody. The optimal expression of the enzyme was 48 h post-infection (Fig. 1D), and the human DNMT1-specific band starts decreasing after 54 h post-infection (data not shown).

Mouse DNMT1 is highly unstable and susceptible to proteolysis (13). The protease-susceptible domains in mouse and human DNMT1 are similar (14). To minimize degradation, all steps of the purification were carried out at 4 °C using buffers containing a mixture of protease inhibitors supplemented with additional E64, a strong inhibitor of the cysteine protease, which is abundant in insect cells. During sonication, the presence of 500 mM NaCl ensures that most of the enzyme, which is normally tightly bound to DNA, remains soluble (36). Most of the enzyme was bound to the chitin-bead affinity matrix. Overnight incubation at 4 °C with 50 mM DTT induced cleavage on the enzyme was bound to the chitin-bead affinity matrix. Over-night incubation at 4 °C with 50 mM DTT induced cleavage on which was at least 95% pure as determined by SDS-PAGE on a 4–20% Tris/glycine gradient gel and Coomassie staining. The dialyzed protein was stored at 4 °C for 4 weeks with little loss of activity. However, significant loss of activity appeared upon prolonged storage (>6 weeks) at either -20 or -80 °C.

**Steady-state Kinetic Properties of Recombinant Human DNA (Cytosine-5) Methyltransferase on Poly(dI-dC)-Poly(dI-dC) Substrate**—Historically, poly(dI-dC)-poly(dI-dC) has been used to define the biochemical properties of mammalian m^5^C MTases (36, 37). Each molecule of poly(dI-dC)-poly(dI-dC) provides a large number of potential (C1) dinucleotide sites for methylation. It has also been claimed that this polymer can undergo methylation at rates comparable to a hemimethylated DNA substrate (38). Thus, we used this substrate to define the biochemical properties of the recombinant human DNMT1. The time course of methylation for poly(dI-dC)-poly(dI-dC) was examined at constant DNA and enzyme concentrations. The reaction was found to be linear for the first 30 min (Fig. 3A), during which time each enzyme molecule undergoes several rounds of catalysis. This is evident from the fact that 1 nmol of human DNMT1 incorporated several nanomoles of methyl groups onto the substrate DNA (Fig. 3, A and B). Similar reactions were carried out with fixed concentrations of DNA and AdoMet but variable amounts of enzyme. The rate of methylation was linear at enzyme concentrations between 1 and 4 nM (Fig. 3B). The differing curves (Fig. 3, A and B) were used to determine the optimal enzyme and substrate concentrations under which linear methylation reaction was obtained. Based on the above observations, a series of double-reciprocal plots was made at six different AdoMet concentrations or six different CI (DNA) concentrations, with poly(dI-dC)-poly(dI-dC) having an average length of 7000 bp. With this procedure the reciprocal of the amount of [3H]CH3 transferred to the DNA in 1 min by 1 nM DNMT1 (1/v) is plotted as a function of DNA concentration (either 1/C1 or 1/AdoMet) in the presence of fixed concentrations of the co-substrate. For bireactant enzymes, which use two substrates and give two products such as DNMT1, these double-reciprocal plots usually yield linear regressions, and their slopes and intercepts are then replotted to derive all the kinetic constants. However, for poly(dI-dC)-poly(dI-dC), both double-reciprocal plots, i.e. those with 1/C1 as the variable substrate (Fig. 4A) as well as those with 1/AdoMet as the variable substrate (Fig. 4B), were not linear. In particular, the increases in 1/v at low 1/C1 (i.e. at high DNA concentration) indicated that the DNA substrate inhibited the reaction, especially in the presence of low amounts of AdoMet (Fig. 4A). We conclude that poly(dI-dC)-poly(dI-dC) acts both as a substrate and as an inhibitor of the methyl transfer reaction with human DNMT1. The regressions from the plots of 1/v versus 1/AdoMet (Fig. 4B) were concave, suggesting some positive cooperativity due to enzyme activation at high AdoMet concentrations. Overall, these data indicate that the methyl transfer reaction with poly(dI-dC)-poly(dI-dC) by human DNMT1 does not follow simple kinetics.

Since the determination of the Michaelis constants and Vmax requires an estimate of the slopes and intercepts from the lines in Fig. 4, the two data points at low 1/C1 that deviated from linearity at the four lowest AdoMet concentrations (Fig. 4A) were excluded from the regressions (32, 39). Fig. 5, A and B, shows that, with the exception of 1 μM AdoMet, such slopes and
De Novo and Maintenance Methylation Catalyzed by the Recombinant Human DNA (Cytosine-5\') Methyltransferase—Finding a 30-fold difference in the $k_{\text{cat}}$ value of poly(dI-dC)/poly(dI-dC) by the recombinant human and mouse DNMT1 prompted us to address the question of the de novo and maintenance methylation properties of the enzyme. Two sets of oligonucleotides were evaluated kinetically. One set of oligonucleotides corresponds to the imprinted locus SNRPN (small nuclear riboprotein-associated peptide N) exon-1, and the other set was from a short stretch of the FMR-1 locus (fragile X mental retardation syndrome). Both DNA sequences mimic human genes and are methylated during development or disease progression. Three sets of duplex DNAs were made with either both strands unmethylated or one strand methylated (hemimethylated). Two hemimethylated duplexes contained either the coding (upper) or the complementary strand (lower) in methylated form. The FMR-1 locus contained either 24 (un-
methylated duplex) or 12 (hemimethylated duplex) cytosine residues available for methylation. Similarly, the SNRPN exon-1 locus contained either 16 or 8 target cytosines (assuming that the CG dinucleotide is the only target site for methylation).

A series of double-reciprocal plots were made with six different AdoMet or six different DNA concentrations following the same principles described for poly(dI-dC)-poly(dI-dC). For the unmethylated duplex, the CG concentration ranged between 0.5 and 5.0 μM, and for the hemimethylated substrate it was between 0.1 and 1.0 μM. The AdoMet concentration was varied from 4.0 to 15.1 μM for unmethylated DNA and 2.0 to 12.1 μM for hemimethylated DNA. At each given DNA and AdoMet concentration, less than 5% of the cytosines were converted to 5-methylcytosine, and the AdoMet concentration was in vast excess over that of the end product, AdoHcy. Double-reciprocal plots were obtained using the SNRPN exon-1 locus either as an unmethylated or hemimethylated substrate. With increasing concentrations of the co-substrate, the slope of the plots decreased, as expected. Representative double-reciprocal plots of SNRPN exon-1 at fixed AdoMet and at fixed DNA are shown in Fig. 6, A and B. Replots of 1/Vmax(app) versus 1/AdoMet (from Fig. 6B), slopes of 1/v versus 1/AdoMet (from Fig. 6B), 1/Vmax(app) versus 1/C (from Fig. 6A), and slopes of 1/v versus 1/C (from Fig. 6A) were all linear (data not shown). kcat for the unmethylated imprinted SNRPN exon-1 locus was about 2.3 h⁻¹. However, the kcat value for hemimethylated SNRPN exon-1 was 23 and 49 h⁻¹ for the lower and upper methylated strand duplexes, respectively (Table II). The values obtained for Kcat were similar for both hemimethylated duplexes.

Similar experiments were carried out with the FMR-1 locus substrate. In this case, the double-reciprocal plots and replots were linear for (CGG-CGG)12 and (m5CGG-CGG)12 but were curved for (CGG-CGG)12. Together with the data from poly(dI-dC)-poly(dI-dC), these curved velocity patterns reveal that the sequence composition and methylation status of the DNA template may dramatically alter the kinetics of the reaction. The kinetic constants measured for the FMR-1 locus are shown in Table III. Unmethylated trinucleotide repeats had a very low kcat (1.2 h⁻¹). The kcat of hemimethylated triplet repeat DNA increased about 10–20-fold, depending on the methylated strand (Table IV), as observed for the SNRPN exon-1 locus. Furthermore, since kcat measures the rate of dissociation of the products (methylated DNA and AdoHcy) from the enzyme, the increases in kcat for the hemimethylated DNA substrates are consistent with human DNMT1 having a greater affinity for hemimethylated DNA (41) than for fully methylated DNA. Finally, the finding that the kcat values varied 2–3-fold between the upper and lower hemimethylated strands suggests that DNA product release is additionally influenced by the sequence composition flanking the CG substrate site.

Some of the hemimethylated templates, such as SNRPN exon-1 and FMR-1 substrates (methylated lower strand, Tables II and III), showed a lower Kcat/AdoMet than their unmethylated counterpart. According to the reaction scheme used to derive the velocity equations for the methyl transfer reaction by DNMT1 (32), Kcat/AdoMet is defined as the ratio between kcat and the forward rate constant for the binding of AdoMet to the catalytic center. For all the DNA templates tested, kcat was greater for the hemimethylated substrates than for the unmethylated ones, which increases Kcat/AdoMet in the former molecules. However, since Kcat/AdoMet values were lower for hemimethylated duplexes, we believe that such low Kcat/AdoMet values are due to substantially higher rates of AdoMet binding. In the accompanying paper (32), we provide evidence that hemimethylated DNA activates the reaction rates by binding to an allosteric site in DNMT1. We propose that the consequence of such allosteric binding is to increase the accessibility of AdoMet to the catalytic center.

The second-order rate constant defined by the ratio of kcat/Km is also a good measure of the catalytic efficiency of an enzyme. For human DNMT1 the catalytic efficiency for hemimethylated DNA was 3–10-fold higher than the corresponding unmethylated DNA (Table IV). Thus we conclude that human
Methylation by Recombinant Human DNA Methyltransferase

**TABLE II**

| Oligonucleotide duplex | \(k_{cat}\) | \(k_{cat}/K_{cat}\text{CG}\) | \(k_{cat}/K_{cat}\text{CG}\) | \(k_{cat}/K_{cat}\text{CG}\) |
|------------------------|-----------|----------------|----------------|----------------|
|                        | \(h^{-1}\) | \(\mu M\)     | \(10^{-6}\times h^{-1}\) |
| Unmethylated           | 2.3 ± 0.1 | 7.2 ± 1.7     | 0.5 ± 0.1      | 4.5 ± 1.1     |
| Methylated upper strand| 49 ± 12.7 | 7.1 ± 2.8     | 1.3 ± 0.3      | 37.6 ± 0.3    |
| Methylated lower strand| 23 ± 3.3  | 1.6 ± 0.6     | 1.3 ± 0.2      | 17.3 ± 0.1    |

**TABLE III**

| Oligonucleotide duplex | \(k_{cat}\) | \(k_{cat}/K_{cat}\text{CG}\) | \(k_{cat}/K_{cat}\text{CG}\) |
|------------------------|-----------|----------------|----------------|
|                        | \(h^{-1}\) | \(\mu M\)     | \(10^{-6}\times h^{-1}\) |
| Unmethylated           | 1.2 ± 0.2 | 3.8 ± 1.2     | 0.3 ± 0.1      |
| Methylated upper strand| 8.3 ± 1.5 | 2.6 ± 1.5     | 0.2 ± 0.06     |
| Methylated lower strand| 22 ± 3.6  | 1.7 ± 0.3     | 0.5 ± 0.1      |

**TABLE IV**

| Preference for hemimethylated DNA by human DNMT1 |
|-----------------------------------------------|
| Oligonucleotide duplexes | Relative \(k_{cat}\) (HM/UM) | Relative \(k_{cat}\) (UM/HM) |
|--------------------------|-----------------|-----------------|
| FMR-methylated upper strand | 6.8 ± 0.7 | 7.1 ± 0.2      |
| FMR-methylated lower strand | 18.3 ± 1.5 | 11.8 ± 2.2    |
| SNRPN-methylated upper strand | 21.0 ± 4.9 | 8.8 ± 2.1    |
| SNRPN-methylated lower strand | 9.9 ± 1.0 | 3.2 ± 0.1     |

\(^a\) HM indicates hemimethylated, and UM indicates unmethylated.

DNMT1 is catalytically more efficient on hemimethylated than on unmethylated DNA.

Non-CG Methylation in Vitro by the Recombinant Human DNA (Cytosine-5) Methyltransferase—The above data clearly demonstrate that the full-length human DNMT1 prefers hemimethylated DNA as a substrate. However, there is methylation in mammals at non-CG sequences (42, 43). To test whether human DNMT1 can methylate non-CG sequences, a series of FdC-containing oligonucleotide duplexes were made with non-CG sites such as FXG, where X was either C, A, or T. DNA-protein cross-linking bands are above 175 kDa and are indicated by an arrow. Pre-stained protein molecular mass markers are indicated on the right.

FIG. 7. Asymmetric methylations by human DNMT1. Radioactive duplex oligonucleotides were covalently cross-linked with human DNMT1. The presence or absence of DNMT1 and/or 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-fluoro-2-deoxyctydine, and W is either A or T. DNA-protein cross-linked bands are above 175 kDa and are indicated by an arrow. Pre-stained protein molecular mass markers are indicated on the right.

FIG. 8. Human DNMT1 DNA complex.

as placenta, lung, brain, and heart (14). Recently, recombinant expression systems using COS cells (28) and baculovirus (23) have been reported for the murine DNMT1. In COS cells the expression was very low, but the baculovirus system led to a much higher level of protein expression. However, even in the baculovirus system, purification involved several chromatographic steps leading to a reduced yield of full-length protein. The addition of a hexa-histidine tag to the N terminus of the DNMT1 and nickel column affinity purification was attempted (23). However, in this purification method several other proteins were present after the affinity elution, and further chromatographic purification was required. Thus we switched to an intein-based purification system (29).2

The full-length DNMT1 gene was placed in a vector such that it was fused in frame to an intein and a chitin-binding domain. The fusion precursor is immobilized on a chitin matrix where the intein can be activated to catalyze its own cleavage and release the full-length human DNMT1 protein. Although the intein-chitin domain remains bound to the affinity beads, the target protein is eluted. We have been successful in obtaining 1 mg of protein/1.5 × 10^9 cells (a liter of infected SF9 cells). This purification protocol yields intact protein at over 95% purity. Often a secondary band close to 55 kDa was observed which was later established to be the viral (AcNPV)-encoded chitinase (data not shown). To facilitate cleavage from the purification tag, the two C-terminal amino acids were changed from lysine and aspartic acid to proline and glycine. This had no detectable effect on activity.

Poly(dI-dC) · poly(dI-dC) as a substrate has been shown to have either comparable or greater methyl group acceptance ability than other DNAs (38). In this work, we demonstrate a complex behavior of the human DNMT1 with the poly(dI-

dC) · poly(dI-dC) substrate. At higher concentrations of CI, substrate inactivation was apparent. It was suggested that mouse DNMT1 can form reversible, multimeric complexes at a higher protein concentration (46), and this may be the cause of the inactivation. Kinetic observations by Hitt et al. (47) suggest an irreversible binding of murine DNMT1 to DNA and aggregation at high molar excess of the enzyme. However, our kinetic measurements were carried out at 2 nM human DNMT1. In this concentration range, protein aggregation was not detected in a
gel shift assay (data not shown). Thus aggregation seems unlikely. Several other DNA substrates that we tested using the above conditions also do not support the aggregation hypothesis. However, an alternative explanation of enzyme inactivation at high poly(dI-dC):poly(dI-dC) concentrations via the formation of a ternary complex consisting of DNA-enzyme-DNA may explain our observations. This hypothesis is supported by the presence of a DNA-binding peptide motif, B1, at the N terminus of the human DNMT1 protein (48). Motif B1, as well as a deleted portion of it—DB1, interacts and binds strongly to DNA. DB1 can bind oligonucleotide duplexes as small as 30 bp through the lysine residues (48). Thus, it is possible that poly(dI-dC):poly(dI-dC) may have formed catalytically inactive complexes with human DNMT1 at high CI concentrations. It should be noted that proteolysis of the N terminus of murine DNMT1 leads to its catalytic activation (21), suggesting that this domain may be involved in repression of methyl transfer.

One potentially significant observation with recombinant human DNMT1 is the high turnover number compared with previous reports for poly(dI-dC):poly(dI-dC) (Table I). This difference could be due to the inherent nature of the human DNMT1 or could reflect a greater percentage of active protein because of our purification protocol. This could also be caused by the substrate, which has a limited search space for methylation. Our results are consistent with the \( k_{\text{cat}} \) values of other AdoMet-dependent methyltransferases such as the bacterial m\(^5\)C \( \text{MTases} \) M.HhaI, 78 h\(^{-1} \) on poly(dG-dC):poly(dG-dC) (9), and tRNA (m5U54) methyltransferase, 108 h\(^{-1} \) (49).

The mammalian enzymes have been classified as maintenance m\(^5\)C \( \text{MTases} \). The mouse enzyme is found to be associated with replication foci where it likely methylates the newly synthesized DNA during the S-phase of the cell cycle (8). We analyzed two synthetic DNAs representing the whole exon-1 sequence of \( \text{SNRPN} \) and 12 CGG trinucleotide repeats from the \( \text{FMR1} \) gene. These are representative of human DNA sequences that undergo methylation during development (imprinting) or are implicated in disease (fragile X syndrome). Both sequences have multiple CGs, thus enabling the study of both de novo and maintenance methylation processes in well-defined, biologically relevant sequences.

The \( k_{\text{cat}} \) values for normal DNA substrates shown in Tables II and III are smaller than the turnover number for poly(dI-dC):poly(dI-dC) (Table I) but are comparable to other m\(^5\)C \( \text{MTases} \) (11, 49). However, the strand bias for methylation is evident for human DNMT1, once the \( k_{\text{cat}} \) values for complementary hemimethylated duplexes are compared (Table IV). This could be the result of an unusual topological structure of the DNA, or the sequence context of 5mC may act as a signal for enhanced methylation in adjacent DNA regions as observed previously (24). Kinetic studies with more substrates will be necessary to assess whether this is due to a specific sequence effect. The human enzyme has a typical turnover number between 7 and 49 h\(^{-1} \) for hemimethylated DNA as compared with 1.2–2.3 h\(^{-1} \) for unmethylated DNA. Our data show that the turnover number for human m\(^5\)C \( \text{MTase} \) and unmethylated DNA is at least 110-fold greater than previously reported (12). The preference for methylation between hemimethylated and unmethylated DNA is about 15-fold, as measured under optimal kinetic conditions (Table IV), whereas a 134-fold difference was observed by Kho et al. (12) for the human enzyme using oligonucleotide duplexes. A major reason for this discrepancy may be the quality of the enzyme since a much smaller human m\(^5\)C \( \text{MTase} \) was used for kinetic analysis by Kho et al. (12) based on the FdC data presented. Our data show clearly (Table IV) that the enzyme prefers hemimethylated DNA as a substrate and is capable of both de novo and maintenance methylation. However, within the nucleus, DNMT1 binds to the replication fork (8) along with many other factors, such as p21 and proliferating cell nuclear antigen (22). Thus, the enzyme is in a very different environment during DNA replication, and the rate of methylation in vitro may differ from that observed in vivo.

Apart from methylating the canonical CGs, the human DNMT1 can also methylate unusual structures such as slipped duplexes, snapbacks, and cruciforms (50). Clark et al. (42) have shown that mammalian cells were capable of maintaining methylation of cytosine at CNG sites within the transfected plasmids. There are also reports of non-CG methylation in the hypermethylated CG promoter region of the human L1 retrotransposon (43). Methylation at such sites and its maintenance were postulated to be mediated by a different enzyme(s) (25). We tested the ability of the recombinant human DNMT1 to carry out non-CG methylation by substituting the target cytosine with a fluorocytosine and allowing the reaction to proceed in the presence of AdoMet (44). Oligonucleotide duplexes containing FCG or FWG sequences formed covalent complexes with the human enzyme, indicative of a successful methyl transfer event. Thus, human DNMT1 can carry out de novo and maintenance methylation at the CG sites and can also maintain methylation of some non-CG sites.

Acknowledgments—We thank Prof. S. Baylin, The Johns Hopkins University, for the gift of the cDNA clones for human DNMT1 (pBH-SHMT5.0-b). We thank Drs. Ming Xu and Sanjay Kumar for suggestions.

REFERENCES

1. Shapiro, H. S. (1970) in CRC Handbook of Biochemistry Selected Data for Molecular Biology (Sober, H. S., ed) pp. 80–103, The Chemical Rubber Co., Ohio

2. Bestor, T. H. (1992) J. Biochem. 204, 204–208

3. Rein, T., Zorbas, H., and DePamphilis, M. L. (1997) Mol. Cell. Biol. 17, 416–426

4. Barlow, D. P. (1995) Science 270, 1610–1613

5. Li, E., Bestor, T. H., and Jaenisch, R. (1992) Cell 70, 915–920

6. Schmutte, C., and Jones, P. A. (1998) Biol. Chem. Hoppe-Seyler 379, 377–388

7. Imbert, G., Feng, Y., Nelson, D. L., Warren, S. T., and Mandel, J. L. (1998) in Genetic Instability and Hereditary Neurological Diseases (Wells, R. D., and Warren, S. T., ed) pp. 27–46, Academic Press, San Diego, CA

8. Leonard, H., Page, A. W., Weir, H. U., and Bestor, T. H. (1992) Cell 71, 865–873

9. Wu, J. C., and Santl, D. V. (1987) J. Biol. Chem. 262, 4778–4786

10. Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) Cell 74, 357–369

11. Flynn, J., Glickman, J. F., and Reich, N. O. (1996) Biochemistry 35, 7308–7315

12. Kho, M. R., Baker, D. J., Laayouni, A., and Smith, S. S. (1998) J. Biol. Chem. 273, 67–70

13. Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988) J. Biol. Chem. 263, 871–883

14. Yamada, M., Ohashi, C., Sugiyama, H., and Tsumori, Y. (1993) Nucleic Acids Res. 21, 2291–2294

15. Koseki, K., Ishihara, G., and Tsumori, Y. (1995) J. Biochem. (Tokyo) 118, 1182–1189

16. Aniello, F., Locascio, A., Fucci, L., Geraci, G., and Branno, M. (1996) Cell 84, 270–281

17. Bestor, T. H. (1992) EMBO J. 11, 67–70

18. Finnegan, E. J., and Dennis, E. S. (1993) Nucleic Acids Res. 21, 2383–2388

19. Pfeifer, M., Bhattacharya, S., Wilson, G. G. (1994) Nucleic Acids Res. 22, 1214–1222

20. Kumar, S., Cheng, X. D., Klimasauskas, S., Wu, J. J., El-Deiry, W., Cumarrasawamy, A., Lennon, G. G., Trask, B. J., Celano, P., and Baylin, S. B. (1992) Nucleic Acids Res. 20, 2287–2291

21. Kojima, H., Ishihara, G., and Tsumori, Y. (1996) J. Biochem. (Tokyo) 120, 1182–1189

22. Schaffner, W., and Hergersberg, M. (1991) Cell 67, 1182–1189

23. Barlow, D. P. (1995) Science 269, 270–273

24. Prath, D., Talbot, D., Sha, M., Benner, J., Hornstra, L., Li, K., Jaenisch, R., and Roberts, J. R. (1997) Nucleic Acids Res. 25, 4666–4673

25. Okano, M., Xie, S., and Li, E. (1998) Nat. Genet. 19, 219–220

26. Tollefsbol, T. O., and Hutchison C. A., III (1995) J. Biol. Chem. 270, 18543–18550

27. Glickman, J. F., and Reich, N. O. (1994) Biochem. Biophys. Res. Commun. 204, 1003–1008

28. Crank, A., Hauselmann, R., Page, A. W., Leonard, H., Bestor, T. H., Schaffner, W., and Hergersberg, M. (1991) Gene (Amst.) 98, 259–263
Methylation by Recombinant Human DNA Methyltransferase

29. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M.-Q. (1996) *J. Biol. Chem.* **271**, 22159–22168
30. Wickham, T. J., Davis, T., Granados, R. R., Hammer, D. A., Shuler, M. L., and Wood, H. A. (1991) *Biotechnol. Lett.* **13**, 483–488
31. O’Reilly, D. R., Miller, L. K., and Luckow, V. A. (1992) *Baculovirus Expression Vectors: A Laboratory Manual*, pp. 124–127, W. H. Freeman & Co., New York
32. Bacolla, A., Pradhan, S., Roberts, R. J., and Wells, R. D. (1999) *J. Biol. Chem.* **274**, 33011–33019
33. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, pp. 9.52–9.55, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
34. Burns, J. A., Butler, J. C., Morgan, J., and Whitesides, G. M. (1991) *J. Org. Chem.* **56**, 2648–2650
35. Davis, T. R., Trotter, K. M., Granados, R. R., and Wood, H. A. (1992) *Bio/Technology* **10**, 1148–1150
36. Turnbull, J. F., and Adams, R. L. (1976) *Nucleic Acids Res.* **3**, 677–695
37. Hitt, M. M., Wu, T. L., Cohen, G., and Linn, S. (1988) *J. Biol. Chem.* **263**, 4392–4399
38. Chuang, L. S., Ng, H. H., Chia, J. N., and Li, B. F. (1996) *J. Mol. Biol.* **257**, 935–948
39. Copeland, R. A. (1996) *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis*, pp. 93–120, Wiley-VCH, New York
40. Segel, I. H. (1975) *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-state Enzyme Systems*, pp. 813–826, John Wiley & Sons, Inc., New York
41. Flynn, J. A., Pradhan, S., Roberts, R. J., and Wells, R. D. (1999) *J. Biol. Chem.* **274**, 33011–33019
42. Clark, S. J., Harrison, J., and Frommer, M. (1995) *Nat. Genet.* **10**, 20–27
43. Woodcock, D. M., Lawler, C. B., Linsemmeyer, M. E., Doherty, J. P., and Warren, W. D. (1997) *J. Biol. Chem.* **272**, 7810–7816
44. Hanek, T., Schmidt, S., and Fritz, H.-J. (1995) *Nucleic Acids Res.* **21**, 303–309
45. Adams, R. L., Gardiner, K., Rinaldi, A., Bynans, M., McGarvey, M., and Burdon, R. H. (1986) *Biochim. Biophys. Acta* **868**, 9–16
46. Reale A., Lindsay, H., Saluz, H. P., Pradhan, S., Adams, R. L., Jost, J. P., and Strom, R. (1995) *Biochem. J.* **312**, 855–861
47. Hitt, M. M., Wu, T. L., Cohen, G., and Linn, S. (1988) *J. Biol. Chem.* **263**, 4392–4399
48. Chuang, L. S., Ng, H. H., Chia, J. N., and Li, B. F. (1996) *J. Mol. Biol.* **257**, 935–948
49. Kealey, J. T., Gu, X., and Santi, D. V. (1994) *Biochemie (Paris)* **76**, 1133–1144
50. Laayoun, A., and Smith, S. S. (1995) *Nucleic Acids Res.* **23**, 1584–1589
51. Glickman, J. F., Flynn, J., and Reich, N. O. (1997) *Biochem. Biophys. Res. Commun.* **230**, 280–284