Mammalian DNA (Cytosine-5-)-methyltransferase Expressed in Escherichia coli, Purified and Characterized*

(Received for publication, January 24, 1995, and in revised form, May 25, 1995)

Trygve O. Tollefsbol† and Clyde A. Hutchison III

From the Department of Microbiology and Immunology, University of North Carolina, Chapel Hill, North Carolina 27599

Besides modulating specific DNA-protein interactions, methylated cytosine, frequently referred to as the fifth base of the genome, also influences DNA structure, recombination, transcription, repair, transcription, imprinting, and mutagenesis. DNA (cytosine-5-)-methyltransferase catalyzes cytosine methylation in eukaryotes. We have cloned and expressed this enzyme in Escherichia coli, purified it to apparent homogeneity, characterized its properties, and we have shown that it hemimethylates DNA. The cDNA for murine maintenance methyltransferase was reconstructed and cloned for direct expression in native form. Immunoblotting revealed a unique protein (M, = 190,000) not present in control cells. The mostly soluble overexpressed protein was purified by DEAE, Sephadex, and DNA cellulose chromatography. Peak methylating activity correlated with methyltransferase immunoblots. The purified enzyme preferentially transferred radioactive methyl moieties to hemimethylated DNA in assays and on autoradiograms. All of the examined properties of the purified recombinant DNA methyltransferase are consistent with the enzyme purified from mammalian cells. Further characterization revealed enhanced in vitro methylation of premethylated oligodeoxynucleotides. The cloning of hemimethyltransferase in E. coli should allow facilitated structure-function mutational analysis of this enzyme, studies of its biological effects in prokaryotes, and potential large scale methyltransferase production for crystallography, and it may have broad applications in maintaining the native methylated state of cloned DNA.

The chemistry of cytosine-5 methylation consists of transfer of a methyl moiety from S-adenosyl-l-methionine (AdoMet)1 to carbon 5 of the pyrimidine ring of cytosine. This one-carbon transfer, catalyzed by DNA (cytosine-5-)-methyltransferase (DNA MTase), is ubiquitous, affecting approximately 5 × 102 cystines/mammalian diploid nucleus (1). Cytosine methylation is the most common modification of DNA found in nature and has been implicated in the control of developmental processes (2), DNA repair (3–5), chromatin organization (6–8), transcription (9–11), X chromosome inactivation (12–13), transposition (14–15), recombination (16), mutagenesis (17–18), replication (19), and genomic imprinting (20). DNA MTase has been shown in mice to be essential for embryonic survival (21) and has been proposed to play a role in general biological processes such as cellular aging (22), carcinogenesis (23), human genetic diseases (24), and evolution (17, 25).

The recognition sequence for DNA MTase is highly specific with almost all cytosine methylation occurring in the duplex palindromic 5′-C-p-G-3′ (CpG). Over half of CpG dinucleotides palindromes are methylated in the mammalian genome (26). After semiconservative replication of DNA, both daughter duplexes are hemimethylated, and DNA MTase, which is localized to replication foci (27), fully methylates the duplex CpG dinucleotides. This process, termed maintenance methylation, restores the parental genomic methylation pattern and is consistent with the in vitro propensity of the DNA MTase for hemimethylated sequences (28–29).

DNA MTase can also methylate certain CpGs that are not in a hemimethylated configuration, a process referred to as de novo methylation. Although the mechanisms for de novo methylation are not completely understood, a number of studies have reported the appearance of newly methylated CpG dinucleotides in the genome (29–32). Only one gene encoding mammalian DNA MTase has been found, and maintenance methylation and de novo methylation are generally believed to be catalyzed by a single enzyme (33–34). Several studies have noted the appearance of de novo methylated cytosines in genomic regions containing preexisting methylated cytosines (i.e. methylation spreading) such as occurs in newly integrated viral DNA in the genome (31, 35–37). Since cytosine methylation can affect the DNA binding of certain transcriptional regulatory factors, the introduction of additional methylated cytosines within gene regulatory sequences may influence gene expression (35). This spreading of cytosine methylation in gene regulatory sequences has been implicated in the gene silencing characteristic of fragile X syndrome (38–39), cellular senescence (22), and X chromosome inactivation (13).

The importance of cytosine methylation in general and the DNA MTase in particular has led us to express this enzyme in Escherichia coli and to further study its mechanisms. Although the cloned cDNA for murine DNA MTase (33) has been expressed in mammalian COS cells (40), we report the first successful expression and purification of catalytically active mammalian DNA MTase in E. coli, providing a potential means for preserving native methylation patterns of cloned DNA in this widely used and simplified system. The purification to apparent homogeneity of DNA maintenance methyltransferase overexpressed in E. coli will facilitate mutational analysis of this enzyme and may allow its large scale production for crystallography. Studies of the effects of the recombinant methyltransferase on the prokaryotic genome and cellular processes will be useful in further elucidating the biological significance of DNA methylation.

18543

This is an Open Access article under the CC BY license.
Mammalian DNA-Methyltransferase Expressed in E. coli

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids—For standard transformations E. coli Sure cells (mcrA-, merCB-, mrr-, hsdR, recB-, lacIq-, Stragetrage) were used routinely. The prokaryotic expression vector pKK223-3 was obtained from Pharmacia Biotech Inc. The cDNA for murine DNA MTase (EMBL accession 14805 (corrected version)) was kindly provided by Timothy Bestor (Columbia University) as overlapping cDNA sequences (pMG and pR2K) cloned into pBluescript SK M13+ (33). pMG contains all of the sequence from the EcoRI linker at the 3' terminus of the cDNA clone to a BglII site near the 3' end of coding (33). pR2K contains the sequence between the unique XhoI site at nucleotide 3138 and an Eco47III site just downstream of the AATAAA polyadenylation signal.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.

Plasmid Construction—The identity of pMG and pR2K was verified by endonuclease digestion. Each plasmid contained an internal XhoI site in the DNA MTase coding sequence as well as a 3' XhoI site in the pBluescript sequence (33). Both plasmids were digested with XhoI and gel-purified, and the XhoI-XhoI sequence of pR2K was ligated into the digested gel-purified pMG plasmid lacking this segment. This fused the coding sequences at the XhoI site (nucleotide 3138) without alteration of the original sequence as confirmed with extensive restriction digests. The newly formed plasmid containing the entire DNA MTase coding sequence in pBluescript was used as template for PCR amplification (20 cycles) of the coding sequence and 3'-untranslated region using (New England Biolabs) and Perkin Elmer Scanning densitometry was performed on these bands to confirm successful cloning.
buffer). The dialyzed S1 solution was diluted with an equal volume of column buffer and loaded onto a DEAE-Sephael column (2.5 × 12 cm bed volume). The eluted column was washed with 2 bed volumes of column buffer to remove unbound protein and further eluted with a 200-ml 0–400 mM NaCl gradient. Fractions were collected and assayed for methyltransferase activity as indicated above. Active pooled fractions were stored at −70 °C in 50% glycerol, 5 mM EDTA.

For ammonium sulfate precipitation, pooled active fractions from DEAE chromatography were diluted with 1 volume of column buffer and brought to 30% ammonium sulfate with gentle stirring over 10 min followed by continued stirring for 20 min on ice. The mixture was centrifuged for 20 min at 10,000 × g, and the supernatant was brought to 60% ammonium sulfate, stirred, and recentrifuged. The 60% ammonium sulfate pellet containing the DNA MTase (30) was resuspended in 2 ml of column buffer, loaded onto a Sephadex G-150 column (2 × 70 cm bed volume) and eluted. Active fractions were pooled and stored as above. DNA cellulose (4 mg of double-stranded DNA/sol of solid, Sigma) chromatography was performed in a 1 × 3-cm bed volume and eluted with a 0–400 mM NaCl gradient. Active fractions were pooled and stored at −70 °C in 50% glycerol, 5 mM EDTA.

Gel Scanning of SDS-PAGE—To estimate percentage of total cellular protein represented by the recombinant DNA MTase, Coomassie-stained 5% SDS-PAGE gels were scanned on an Apple OneScanner, plotted, and integrated for density using the Image 1.49 program on a Macintosh IIfx computer. A total of four different protein concentrations of the S1 and S2 lysates (representing a 4-fold difference in total protein loaded) from two independent DNA MTase purifications were resolved on SDS-PAGE gels, scanned, and plotted in duplicate. The M, 180,000 protein band (identified with molecular mass markers) was integrated for density in each lane and compared with the total integrated density of all proteins in the same lane to obtain percentage DNA MTase of total E. coli protein. For similar determinations in the mammalian system, a photomicrograph (kindly provided in reprint form by Steven Smith, City of Hope National Medical Center, Duarte, CA) of a SDS-PAGE gel resolving the crude lysate fraction of total human placental protein containing the identified DNA MTase (40) was also gel-scanned, plotted, and integrated for density as for the E. coli crude lysates.

RESULTS
We chose to express the maintenance DNA MTase in its native form to allow its use in future in vivo studies (e.g., preserving methylation patterns of cloned DNA) without potential interference with activity or DNA binding from a fusion product. Plasmid pTOTT1 was constructed to express the native DNA MTase from the strong inducible tac promoter (Fig. 1). Immunoblotting kinetic studies for DNA MTase indicated full expression of this enzyme within 3 h of IPTG induction (data not shown). We cloned pTOTT1 into mcr (modified cytosine restriction) cells to prevent potential DNA degradation by the mcr system (47). The lysed E. coli cells containing pTOTT1 (expression vector) revealed a unique protein (M, = 190,000) on immunoblots probed with the DNA MTase polyclonal antibody (Fig. 2). This protein was not present in lysates of cells containing pKK223–3 (control vector lacking the DNA MTase insert). The calculated molecular mass of the DNA MTase is 172,258 based on its coding sequence. However, this enzyme has been shown previously to resolve at an apparent relative molecular mass of 190,000 on SDS-PAGE gels, which is thought to be due to posttranslational modifications of the enzyme and/or its molecular shape (33).

E. coli cells expressing mammalian DNA MTase do not appear to grow as well as cells containing the control vector and typically require 3.5 h to reach an A600 of 0.5 at 37 °C, whereas control cells reach this stage of growth within 3 h. The pTOTT1 cells produce slightly smaller colonies on culture plates and less turbid overnight cultures compared with cells containing the control vector (data not shown). These differences in comparison to control cells became more pronounced as the cells were transferred to successive culture plates over a period of several months. To prevent progressive cellular proliferative retardation, we periodically transformed fresh mcr− cells with the pTOTT1 expression vector. We have not yet fully quantified the degree of apparent cellular proliferative and growth impairment. It seems possible that its cause may be related to...
effects of DNA MTase expression on the E. coli genome, although the large size of the novel protein product itself may also be a factor. Evidence for methylation of high molecular weight E. coli genomic DNA in vitro can be seen (see Fig. 6B), suggesting that a similar process may occur in vivo affecting the growth of these cells.

To assess enzymatic catalysis by the cloned DNA MTase, the DNA methylating activity (as measured by transfer of tritiated methyl groups from AdoMet to DNA) of DEAE-purified fractions was compared for cells transformed with pTOT1 and pKK223–3 (Fig. 3). The DEAE columns were simultaneously chromatographed and eluted with a salt gradient. Peak methylating activity for the fractions from the pTOT1-transformed cell lysates eluted in the range of 100–150 mM NaCl, consistent with results of the DNA MTase purified from mammalian cells (29–30, 48). No obvious methylating peak was seen for the DEAE-chromatographed lysates of pKK223–3-transformed control cells. Immunoblots performed on the pTOT1 DEAE fractions indicated a M, 190,000 protein correlating with peak methylating activity (fractions 32–52; pTOT1), which was not apparent below 100 mM salt (fractions 8–30; pTOT1) or above 150 mM salt (fraction 71; pTOT1). Fig. 3 also shows that the most intense M, 190,000 bands (fractions 38–42; pTOT1) correlated with fractions having the highest methylating activity. The control DEAE column showed no evidence of the M, 190,000 protein as indicated by the absence of this band at peak methylating activity for the pTOT1 column (fraction 38; pKK223–3).

The cloned DNA MTase was purified to apparent homogeneity by assaying for methylating activity in a three-column system based on protein charge (DEAE), size and shape (Sephadex), and DNA-affinity (DNA cellulose) (Fig. 4). Due to the presence of various inhibitory substances in crude fractions and lability of the enzyme (49), meaningful estimates of total purification factor could not be obtained consistent with reports by others (29–30, 50). Gel filtration yielded a single peak of methylating activity in the M, 180,000–205,000 range, consistent with polyacrylamide gel estimates. While size separation is efficient in this expression system due to the relatively large size of the mammalian DNA MTase compared with most E. coli proteins (Fig. 5C), some protein impurities remain in the Sephadex fraction, and a final purification based on the affinity of this enzyme for DNA is quite effective in producing a homogeneous purification as assessed by silver staining (Fig. 4D).

Although improvements of the purification procedure are expected to increase the yield of recombinant DNA MTase, we recovered almost a full milligram (887 µg) of apparently pure enzyme from about 10 liters of E. coli cells. The apparently homogeneous protein exhibiting peak methylating activity following DNA cellulose chromatography reacted with the DNA MTase antibody on immunoblots (Fig. 4E).

Partially purified recombinant DNA MTase was used for comparison of substrate preference with increasing DNA MTase purity, assessment of relative effectiveness of purification steps, estimates of solubility and degree of expression of the cloned DNA MTase in E. coli, and DNA substrate analysis studies (Fig. 5). A hemimethylated oligodeoxynucleotide was synthesized containing methyl moieties at approximately 15-base pair intervals for use as substrate in DNA MTase assays (see “Experimental Procedures” for chemical synthesis and Table I for structure of hemimethylated oligodeoxynucleotide). Preferential transfer of radioactive methyl moieties to the oligodeoxynucleotide substrate containing hemimethylated CpG sites over the control lacking substrate was apparent after DEAE purification (Fig. 5A), and this ratio improved with gel filtration (Fig. 5B). Ethidium bromide staining of agarose gels indicated minor amounts of large molecular weight E. coli genomic DNA present after DEAE purification (data not shown), accounting for the slight activity of control assays lacking oligonucleotide substrate (Fig. 5A). The chemically synthesized hemimethylated oligodeoxynucleotide underwent greater methylating activity in DEAE and gel filtration fractions than the highly methylatable de novo substrate, poly(dI·dC)·poly(dI·dC), indicating preferential hemimethylation by the recombinant DNA MTase.

The pooled active fractions as well as the crude lysates were assessed on polyacrylamide gels for protein content and purity (Fig. 5C). The soluble (S1) and insoluble (S2) SDS-PAGE crude lysate fractions were estimated for percent DNA MTase by scanning stained gels (see “Experimental Procedures”). The DNA MTase comprised approximately 2% (range of 1.0–3.0%) of total E. coli protein in the S1 fraction and about 0.3% (range of 0–0.53%) for the insoluble S2 fraction, indicating that approximately 85% of the enzyme is expressed in soluble form (see Fig. 5C for comparison of S1 and S2 fractions). The overall expression of DNA MTase in these cells is about 2.5% of total E. coli protein. By contrast, mammalian cells contain a mean of 0.05% DNA MTase of total human placental protein (see “Experimental Procedures” under “ Gel Scanning of SDS-PAGE” and Ref. 46).

To demonstrate that the recombinant DNA MTase is indeed active with a preference for hemimethylated DNA, we reacted the partially purified enzyme with oligodeoxynucleotides in the
Mammalian DNA-Methyltransferase Expressed in E. coli

![Graph](image)

**TABLE I**

| DNA structurea | Type            | Meanb (pmol/h) |
|----------------|-----------------|----------------|
| Complete 60-mer sequence |                | 213 ± 9 | 92 ± 19 | 73 ± 40 | 16 ± 19 |

**Complete 60-mer sequence**

5'-GTTAGATTCATATGTTAGTTAGATATCGCCCCATCCTCG-3' 3'-CAGTTAATACATGTTAGTTAGATATCGCCCCATCCTCG-5'  

Oligonucleotide

| Oligonucleotide | Mean (pmol/h) |
|-----------------|---------------|
| 5' ... CGGATGTCGACTATCGATATTGCCCATCCTCG ... 3' | 213 ± 9 |
| 3' ... GCCTACAGCTGTATGCTATATAGCCGTTAGAGG ... 5' | 92 ± 19 |
| 5' ... CGGATGTCGACTATCGATATTGCCCATCCTCG ... 3' | 73 ± 40 |
| 3' ... GCCTACAGCTGTATGCTATATAGCCGTTAGAGG ... 5' | 16 ± 19 |

**Footnotes**

a All oligodeoxynucleotides are 60 base pairs in length (full sequence shown at top of table) except where indicated. Asterisks indicate position of methylated cytosines placement during chemical synthesis of oligodeoxynucleotide substrates.

b Assays were conducted with excess substrate (5 µg of DNA) for 3 h. Other assay conditions were as indicated under “Experimental Procedures.” Each value is the mean ± S.E. of three independent determinations. Control values (i.e. samples otherwise identical to and assayed side-by-side with substrate-containing samples but lacking added DNA substrate; mean = 144.1 ± 6 pmol/h) were subtracted from each sample value in each individual experiment before determination of the indicated means and S.E. All assays utilized 40 µg of partially purified DNA MTase (SD fraction, Fig. 5C).

---

The presence of radioactive AdoMet, resolved the samples on agarose gels, and subjected the gels to autoradiography (Fig. 6). The gel-isolated hemimethylated oligonucleotide produced the most intense band on autoradiography, demonstrating preferential transfer of methyl moieties to hemimethylated CpGs. Some radioactivity was apparent in the otherwise identical
nonmethylated oligodeoxynucleotide (i.e. de novo methylation), and this activity was greater than that for the identical fully methylated oligonucleotide containing no methylatable CpGs. Thus it is apparent that the recombinant DNA MTase transfers methyl moieties directly to these oligodeoxynucleotides with a preference for hemimethylated CpG sites and with a much lower propensity for nonmethylated CpG sites. Very little methylation appears to occur at sites other than CpG (Fig. 6B, lane 4).

To further characterize the enzymatic activity of the DNA MTase purified from E. coli, we quantitated in assays the methyl receptivity of otherwise identical oligonucleotides differing only in placement of methyl moieties (Table I). These analyses utilized the more purified gel filtration fraction (Fig. 5C) containing no evidence of contaminating E. coli DNA. Table I shows that the hemimethylated oligodeoxynucleotide substrate received the most radioactive methyl transfer catalyzed by the recombinant DNA MTase consistent with the DNA MTase partially purified from mammalian cells (51). Also similar to the mammalian cell enzyme, nonmethylated oligonucleotides can undergo de novo methylation, and sequences containing no methylatable CpGs (i.e. pre-methylated at all CpG sites) are poor templates for the DNA MTase (Table I), demonstrating its strong preference for cytosine methylation specifically in CpG dinucleotides (29–30, 51–52). A duplex trimethylated oligodeoxynucleotide containing only two de novo methylatable CpGs on each stand (Table I) is more receptive to de novo methylation (22.9 pmol/h/CpG) than an otherwise identical nonmethylated oligonucleotide containing five de novo methylatable CpGs on each stand (7.3 pmol/h/CpG), indicating enhanced de novo methylation of a pre-methylated oligodeoxynucleotide containing methylatable CpGs.

**DISCUSSION**

The widely-used techniques of DNA cloning and PCR amplification strip mammalian genomic DNA of its original cytosine methylation. DNA that lacks its native cytosine methylation pattern may give different results in mobility shift analysis, endonuclease digestions, and other procedures analyzing its properties and behavior. We developed the idea that the methylation pattern of cloned DNA could be preserved in host bacteria expressing the maintenance DNA MTase. However, the cDNA for this enzyme has previously been expressed only in mammalian cells (COS-1) (40). Whereas this may be of use in studying the effects of variations in DNA MTase levels in mammalian cells, we chose to clone and express DNA MTase in E. coli. We developed this system not only for its possible use in maintaining methylation patterns of cloned DNA in bacteria but also because of the widespread use of E. coli as a protein expression system, the simplification of cell culture and purification processes, the potential of large scale production of the enzyme for crystallography, and the facilitation of mutagenesis studies of this enzyme.

The known potential for de novo methylation and methylation spreading by the DNA MTase (29–30, 35) could be a factor in preserving methylation patterns of genomic DNA in this system; however, both of these processes occur in proportion to greater DNA MTase levels (48, 51) and number of cell genera-
Mammalian DNA-Methyltransferase Expressed in E. coli

18549

FIG. 6. Methyl transfer by the recombinant DNA MTase. A, ethidium bromide-stained agarose gel. A 100-µl assay mix containing 1.66 µCi [methyl-2H]AdoMet (85 Ci/mmol) and 200 µg of partially purified recombinant DNA MTase (fraction D, Fig. 5C) was incubated for 3 h at 37 °C with each separate reaction either containing 5 µg of DNA or lacking added DNA (control). Phenol-extracted reaction mixtures were resolved on a 4.5% agarose gel and stained with ethidium bromide. Lane M, molecular weight marker; lane 1, control lacking DNA; lane 2, nonmethylated duplex; lane 3, trimethylmethylated duplex; lane 4, pentamethylated (i.e. fully methylated at CpG sites) duplex (see Table I for DNA structures). Each lane consisted of 1 µl of a 10-µl solution (500 ng of DNA where indicated). B, autoradiogram of 3H-labeled methyl transfer to DNA. A different freshly prepared 4.5% agarose gel was loaded with 7 µl of each of the same phenol-extracted samples depicted in A above (3.5 µg of DNA where indicated) and impregnated with En3Hance for 3 h, soaked in 5% acetic acid for 1 h, and dried on filter paper. The film for autoradiography was preflashed and exposed to the gel for 14 days at −70 °C before developing. Lane 1, control lacking added DNA; lane 2, nonmethylated duplex as substrate; lane 3, trimethylmethylated duplex; lane 4, pentamethylated (fully methylated) duplex (see Table I for DNA structures). Minor diffusion of oligodeoxynucleotide bands due to processing in En3Hance and acetic acid is apparent. High molecular weight DNA in wells is considered to be in vitro methylated E. coli genomic DNA (see also gel A). Each oligodeoxynucleotide is identical except for its methylated state.

Contributions (31, 35). Modulating the DNA MTase expression by limiting IPTG induction and minimizing cell culturing times may be useful approaches for reducing the possibility of de novo methylation. Analysis of the cloned product with methylating-sensitive isoschizomers (44) or methylation sequencing (53–54) would be prudent to assess the possibility of ectopic methylation.

Previously it was thought that the mammalian DNA MTase might be toxic to E. coli since de novo methylation of the E. coli genome may activate the mer system leading to DNA degradation (47), even though the mammalian DNA MTase is primarily a maintenance methyltransferase and appears to de novo methyleate only as a secondary function (29–30). In order to circumvent this potential problem, we cloned the reconstituted murine MTase cDNA in mer− cells. The mer− cells expressing DNA MTase are slightly less proliferative than control mer− cells (i.e. containing the cloning vector alone), and others relating to de novo methylation of the E. coli genome. Transformation of the vector into fresh mer− cells appears to improve cellular proliferation to near control levels. In spite of this minor growth impairment, these cells are able to overexpress the DNA MTase to relatively high levels compared with the levels of this enzyme in mammalian cells.

Studies of the effects of expression of the cloned mammalian DNA MTase on the E. coli genome, on the control of cellular processes in E. coli, and on replication rates as well as cell viability may contribute to understanding the control mechanisms of this enzyme and its biological significance. A number of prokaryotic cellular control processes could be affected by expression of this recombinant enzyme in these cells such as the transcription of key regulatory genes, DNA repair, replication, and recombination. We have previously reported several theoretical molecular mechanisms of cellular senescence (22), a hallmark of which is reduced replicative capacity, and have suggested that de novo methylation by the DNA MTase may contribute to this phenomenon in aging eukaryotic cells (22). Prokaryotic cells do not senesce (22), and studies are in progress analyzing the E. coli cells now expressing this protein for evidence suggestive of senescing cells (e.g. morphological changes, slowing of cell replication) and the DNA MTase as one of the putative “mortality gene” products.

The maintenance DNA MTase, purified from mammalian cells, is highly susceptible to proteolytic degradation (34) and loss of enzyme activity due to its lability (29–30). Moreover, the DNA MTase is present in very limited quantities in mammalian cells (46). The expression of the DNA MTase in E. coli and purification of this enzyme to apparent homogeneity may help overcome some of these problems. It is generally known that the use of E. coli allows rapid, easy growth of large numbers of cells with less endogenous protein heterogeneity and bypassing of nuclear isolation protocols. In the case of this specific enzyme, its purification from E. coli may also be facilitated by its relatively large size compared with most E. coli proteins, allowing more effective size separation and reducing the risk of proteolytic degradation and loss of enzyme activity. Whatever the choice of purification protocol, the expression of mammalian DNA MTase in E. coli should allow greater availability of purified enzyme. All of the properties of the purified recombinant DNA MTase examined in this study including relative molecular mass, elution in salt gradients, affinity for DNA, immunoreactivity, and substrate preference are consistent with the known properties of the enzyme purified from mammalian cells (29–30, 33–34, 49, 51–52).

Although it is generally thought that the eukaryotic DNA MTase is capable of maintenance and de novo methylation without assistance from associated mammalian proteins or factors, this important question is still not fully resolved (55). Purification of the mammalian enzyme has helped address this issue, but minor contaminants that assist the DNA MTase could still be present in apparently pure fractions. Our studies indicate that the enzyme is indeed capable of both types of DNA methylation. The expressed product in E. coli was originally derived from a single mammalian gene (33), and when this cDNA is expressed in E. coli and purified, it can perform both maintenance and de novo methylation of DNA. Whatever other proteins may be involved in the eukaryotic methylation process, it is clear that the essential features of maintenance and de novo methylation are not dependent upon associated proteins unique to the mammalian replication apparatus.

Similar to the mammalian DNA MTase isolated from mammalian cells (56–58), the recombinant enzyme purified from E. coli has a preference for hemimethylated CpG dinucleotides, has a tendency to de novo methylate DNA, and transfers methyl moieties at very low levels in substrates not containing methyleatable CpG dinucleotides. Although some cytosine methylation can occur in other dinucleotides in the mammalian genome containing cytosine in the 5′ position (31, 59–60), and such activity has occasionally been reported to be at relatively high levels (59), our studies with the recombinant enzyme indicate that this occurs only very rarely in oligodeoxynucleotides containing these dinucleotides.

The mechanisms for the propensity of the enzyme to methylate in regions already containing methyl moieties (i.e. genomic methylation spreading) are not fully understood (35). These studies indicate enhanced de novo methylation of oligodeoxynucleotides containing preexisting methyl moieties, which suggests in vitro methylation spreading. A more detailed study of methylation spreading in vitro will be reported elsewhere.

Currently, work is aimed toward preserving methylation patterns of cloned DNA using our expression system. Other
intended studies are the effect of expression of the mammalian MTase on control of biological processes in prokaryotic cells, further delineation of the functional domains of the main­

enance methyltransferase in mutagenesis studies, and large scale production of this enzyme for crystallography. Finally, studies are in progress focusing on a more extensive analysis of the molecular mechanisms of methylation spreading using the defined in vitro oligodeoxynucleotide system reported in this initial study.

Acknowledgments—We thank Timothy Bestor, Lucy Andrews, Steven Smith, Gerd Pfeifer, and David Klapper for helpful discussions and Lucy Andrews, Mike Conrad, Bradford Coffee, David Klapper, and Devon Byrd for critical comments on the manuscript. We also thank Susan Elmore for oligodeoxynucleotide synthesis.

REFERENCES

1. Antequera, F., and Bird, A. (1993) in DNA Methylation: Molecular Biology and Biological Significance (Jost, J. P., and Saluz, H. P., eds.) pp. 169–185, Bruckhauser Verlag, Basel, Switzerland.
2. Monk, M. (1991) Philos. Trans. R. Soc. Lond. B 326, 299–312.
3. Kato, M., and Taylor, H. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 7350–7354.
4. Brown, T., and Jiricny, J. (1987) EMBO J. 6, 945–950.
5. Brown, T. C., and Jiricny, J. (1989) EMBO J. 8, 705–711.
6. Gruenbaum, Y., Cedar, H., and Razin, H. (1982) Nature 305, 262–268.
7. Neske, M., and Inouye, K. (1983) Cell 34, 197–206.
8. Williams, J., and Bird, A. (1991) FEBS Lett. 285, 155–159.
9. Busslinger, M., Hurst, J., and Flavell, R. A. (1983) Cell 34, 197–206.
10. Cedar, H. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 21–25.
11. Barletta, J., and Greer, S. B. (1992) Anticancer Res. 12, 1–5.
12. Morrissey, J. H. (1981) Annu. Rev. Genet. 15, 563–587.
13. Johne, R., and Jaenisch, R. (1982) EMBO J. 1, 519–526.
14. de Boer, H. A., Comstock, L. J., and Vasser, M. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 21–25.
15. Sano, H., Noguchi, H., and Sager, R. (1983) Cell 34, 197–206.
16. Bolden, A., Ward, C. A., and Weissbach, A. (1984) J. Biol. Chem. 259, 12437–12443.
17. Selker, E. (1992) Mol. Biol. 23, 805–807.
18. Toth, M., Muller, K., and Jaenisch, R. (1981) EMBO J. 1, 519–526.
19. Bolden, A. H., Nalin, C. M., Ward, C. A., Poonian, M. S., and Weissbach, A. (1984) Mol. Cell. Biochem. 29, 113–114.
20. Rijssen, K. E., Riggs, A. D., and Smith, S. S. (1985) Cell 40, 261, 1097–1102.
21. Bolden, A. H., Nalin, C. M., Ward, C. A., Poonian, M. S., and Weissbach, A. (1984) Mol. Cell. Biochem. 29, 113–114.
22. Rijssen, K. E., Riggs, A. D., and Smith, S. S. (1985) Cell Biochem. 29, 337–349.
23. Raleich, E. A., and Wilson, G. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 9070–9074.
24. Oberle, I., Rousseau, F., Heitz, D., Ketz, C., Devys, D., Hanauer, A., Boue, J., Bertheas, M. F., and Mandel, J. L. (1991) Science 252, 1097–1102.
25. Czank, A., Hauselmann, R., Page, A. W., Leonhardt, H., Bestor, T. H., Schaffner, W., and Herzelberg, M. (1991) Gene (Amst.) 106, 269–283.
26. de Boer, H. A., Comstock, L. J., and Vasser, M. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 21–25.