Influence of Pre-existing Methylation on the de Novo Activity of Eukaryotic DNA Methyltransferase†

Daniela Carotti,*[,‡,⊥] Salvatore Funiciello,[‡,||] Franco Palitti,[‡,⊥] and Roberto Strom[§,⊥]

Departments of Biochemical Sciences 'A. Rossi Fanelli' and of Cellular Biotechnology and Hematology, University of Rome 'La Sapienza', Research Institute 'S. Pietro' Hospital—Fatebenefratelli, Rome, and CNR Center of Molecular Biology, Rome, Italy

Received May 2, 1997; Revised Manuscript Received November 10, 1997

ABSTRACT: Aberrant de novo methylation of CpG island DNA sequences has been observed in cultured cell lines or upon malignant transformation, but the mechanisms underlying this phenomenon are poorly understood. Using eukaryotic DNA (cytosine-5)-methyltransferase (of both human and murine origin), we have studied the in vitro methylation pattern of three CpG islands. Such sequences are intrinsically poor substrates of the enzyme, yet are efficiently methylated when a small amount of 5-methylcytosine is randomly introduced by the M.*Sss*I prokaryotic DNA (cytosine-5)-methyltransferase prior to in vitro methylation by the eukaryotic enzyme. A stimulation was also found with several other double-stranded DNA substrates, either natural or of synthetic origin, such as poly(dG-dC)·poly(dG-dC). An A+T-rich plasmid, pHbβ1S, showed an initial stimulation, followed by a severe inhibition of the activity of DNA (cytosine-5)-methyltransferase. Methylation of poly(dI-dC)·poly(dI-dC) was instead inhibited by pre-existing 5-methylcytosines. The extent of stimulation observed with poly(dG-dC)·poly(dG-dC) depends on both the number and the distribution of the 5-methylcytosine residues, which probably must not be too closely spaced for the stimulatory effect to be exerted. The activity of the M.*Sss*I prokaryotic DNA methyltransferase was not stimulated, but was inhibited by pre-methylation on either poly(dG-dC)·poly-(dG-dC) or poly(dI-dC)·poly(dI-dC). The prokaryotic and eukaryotic DNA methyltransferases also differed in sensitivity to poly(dG-m^{5}dC)·poly(dG-m^{5}dC), which is highly inhibitory for eukaryotic enzymes and almost ineffective on prokaryotic enzymes.

De novo methylation of CpG dinucleotides in eukaryotic cells occurs during several physiological and pathological processes. It is involved in the establishment of the somatic methylation pattern during development (Monk, 1995), in genomic imprinting (Li et al., 1993), and in X chromosome inactivation (Singer-Sam & Riggs, 1993). Tumor progression (Szyf, 1996), viral integration (Toth et al., 1990), and fragile X syndrome (Knight et al., 1993) are all characterized by aberrant methylation patterns. CpG islands are particularly sensitive to the effects of de novo methylation. These 1-2 kb-long DNA sequences (Bird, 1986), found mainly at the 5′ end of housekeeping genes, are found in the unmethylated form in somatic cells, except for those localized on the inactive X chromosome or those associated with Alu sequences or L1 genes. In contrast, a widespread methylation may occur on CpG islands associated with nonessential genes in cultured cell lines (Antequera et al., 1990) or upon malignant transformation, in which case it is likely to act as a means to inactivate tumor suppressor genes. CpG island methylation has indeed been shown for the *VHL* and *RB* genes in renal carcinoma cell (Herman et al., 1994) and retinoblastomas (Greger et al., 1989), respectively, and for the gene encoding the cyclin-dependent kinase inhibitor p16 in a variety of human tumors and tumor-derived cell lines (Gonzales-Zulueta et al., 1995). In all cases, methylation correlates with loss of expression of the associated genes.

The mechanism(s) capable of maintaining CpG islands free of methylation have only been partially elucidated in the last years. These CpG-rich sequences have been shown to be intrinsically poor in vitro substrates (Carotti et al., 1989; Bestor et al., 1992) of the DNA (cytosine-5)-methyltransferase (DNA MTase)[1] that catalyzes the methyl transfer from *S*-adenosyl-L-methionine to the 5′ position of cytosines included in CpG dinucleotides (Smith, 1994); most of them contain Sp1 elements whose removal or alteration appears to induce methylation of the associated island in embryonic cells, as shown for the *aprt* gene (Brandeis et al., 1994; Macleod et al., 1994).

What has received little attention, so far, is the opposite question: how do these sequences get methylated in particular circumstances if their methylation-free status is so strictly controlled? In other words, how can DNA MTase be forced, under certain conditions, to catalyze methylation

† This work was supported by grants from Ministero per l'Università e la Ricerca Scientifica e Tecnologica and from Istituto Pasteur-Fondazione Cenci Bolognetti.

* Author to whom correspondence should be addressed at Dipartimento di Scienze Biochimiche 'A. Rossi Fanelli', Piazzale A. Moro 5, 00185 Roma, Italy. Telephone: +39-6-49910969. Fax: +39-6-4440062. E-mail: dcarotti@axcasp.caspur.it.

‡ Department of Biochemical Sciences 'A. Rossi Fanelli', University of Rome 'La Sapienza'.

§ Department of Cellular Biotechnology and Hematology, University of Rome 'La Sapienza'.

|| Research Institute 'S. Pietro' Hospital—Fatebenefratelli.

⊥ CNR Center of Molecular Biology.

[1] Abbreviations: DNA MTase, DNA (cytosine-5)-methyltransferase.

S0006-2960(97)01031-3 CCC: $15.00
Published on Web 01/13/1998

of CpG island DNA as efficiently as any other DNA?

It has been noticed by several authors that a high level of in vitro methylation of specific DNA sequences by the eukaryotic DNA MTase is favored by a pre-methylation step performed by a prokaryotic methylase, such as M.*Sss*I, M.*Hpa*II, or M.*Hha*I (Rachal et al., 1989). Moreover, studies with synthetic polymers have shown that the presence of a 5-methylcytosine (even if not inserted in a CpG dinucleotide) stimulates in vitro methylation of a neighboring CpG pair by the eukaryotic enzyme, in both single-stranded (Carotti et al., 1986; Christman et al., 1995; Lindsay & Adams, 1996) and double-stranded forms (Tollefsbol & Hutchinson, 1995).

We took advantage of these observations to investigate whether the presence of methyl groups could specifically stimulate the activity of a human placenta DNA MTase when methylating a plasmid containing the CpG island associated to the mouse *Htf9*-A*/RanBP1* and *Htf9*-C genes (Bressan et al., 1991). With this construct, we have previously demonstrated the reduced methyl-accepting ability of CpG-rich sequences (Carotti et al., 1989).

MATERIALS AND METHODS

Materials. Poly(dG-dC)·poly(dG-dC), poly(dI-dC)·poly(dI-dC), and poly(dG-m^{5}dC)·poly(dG-m^{5}dC) were purchased from Pharmacia Biotech, and *Escherichia coli* DNA was from Serva. *Micrococcus luteus* DNA was isolated from lyophylized bacteria (Sigma). M.*Sss*I, M.*Alu*, M.*Hae*III, and M.*Hha*I were from New England Biolabs. Eukaryotic DNA MTases (EC 2.1.1.37) were partially purified from human placenta and murine liver nuclei, and enzymatic assays were performed as previously described (Carotti et al., 1986, 1989). *S*-Adenosyl-L-[*methyl*-^{3}H]methionine (55−85 Ci/mmol) was from New England Nuclear, and *S*-adenosyl-L-methionine was from Sigma.

Plasmids. pL9.2, a kind gift from Dr. P. Lavia, is a pUC subclone of the mouse CpG-rich sequence *Htf9* (Lavia et al., 1987) (accession number X50830). The 3.8 kb-long *Eco*RI fragment includes CpG-depleted (genomic) regions on either side of the island. pEB9, a kind gift from Dr. D. Toniolo, is a 5.9 kb-long genomic *Eco*RI-*Bam*HI fragment, derived from the human X chromosome and cloned in pBR322, which contains two CpG islands, respectively, associated to the *GdX* and *P3* genes (Toniolo et al., 1988). pHbβ1S is a 3.4 kb-long *Pst*I fragment containing the human β globin locus cloned in pBR322 [sequence in Collins and Weissman (1984)].

In Vitro Methylation. DNAs (60 μg/mL) were pre-methylated at 37 °C for various times in 50 mM Tris-HCl, pH 7.8, with M.*Sss*I (25−50 U/mL) [or when required, with M.*Hha*I(160−530 U/mL, or with placenta DNA MTase], in the presence of 16 μM *S*-adenosyl-L-methionine, 5 mM EDTA, 0.5 mM DTT, and 10% (v/v) glycerol. Methylases were thereafter inactivated by heating at 55 °C for 15 min. Mock pre-methylation was carried out in the same conditions in the absence of enzyme. Parallel tests were run in the presence of *S*-adenosyl-L-[*methyl*-^{3}H]methionine (20 μCi/mL, 2700 dpm/pmol) to measure the amount of methylation reached. In order to determine their susceptibility to further methylation, pre-methylated (or mock pre-methylated) DNAs were diluted to 20 μg/mL and incubated for 2 h at 37 °C in the presence of placenta DNA MTase (5−10 units/mL) and fresh 16 μM *S*-adenosyl-L-[*methyl*-^{3}H]methionine (15 μCi/mL, 2000 dpm/pmol). The plasmid methylation pattern was instead determined after overnight incubation under the same conditions. DNA was thereafter purified and digested with suitable restriction enzymes, and fragments were separated by agarose gel electrophoresis. Gel lanes were then cut in 2 mm-wide slices and radioactivity was measured for each gel slice after redissolution in 2 mL of water for 10 min at 100 °C, as previously described (Carotti et al., 1989).

RESULTS

Effect of Pre-methylation on in Vitro Methylation of CpG Islands. M.*Sss*I is a bacterial enzyme with the same 5′-CpG-3′ recognition sequence as eukaryotic DNA MTase. Since its purification, this enzyme has been widely used to artificially methylate DNA to study the effect of CpG methylation on gene expression, and it has been preferred to other cytosine methylases such as M.*Hpa*II or M.*Hha*I, whose recognition sequences (CCGG and GCGC, respectively) are only a subset of all CpGs present in genomic DNA. We used this enzyme to introduce, in a quantitatively controlled manner, a small amount of randomly distributed methyl groups in the mouse *Htf9* CpG island contained in the plasmid pL9.2 (54% G+C, 349 CpGs = 5.4% CpG, on a single-strand basis). The plasmid was methylated in vitro at about 26% of its possible sites, and the distribution of radioactive methyl groups was analyzed by restriction digestion and agarose gel electrophoresis, as previously described (Carotti et al., 1989). The extent of methylation catalyzed by M.*Sss*I was found to be strictly dependent on the number of CpGs present in each restriction fragment, as expected (data not shown), therefore meaning that M.*Sss*I does not show any preference for particular CpG sites, and that every CpG has the same probability of being methylated.

When the partially methylated plasmid was used as substrate of human placenta DNA MTase, the overall methylation level catalyzed by the enzyme was approximately 30% higher than that of the unmethylated plasmid. We next analyzed how the incorporated radioactive methyl groups were distributed along the DNA molecule. We previously showed (Carotti et al., 1989) that the enzyme is not able to methylate the actual CpG island sequence (1.6 kb-long *Dra*I-*Pvu*II fragment) as efficiently as other DNA fragments derived from the same plasmid. Results in Figure 1 show that, while [*methyl*-^{3}H] incorporation is only about 60% of the expected value on the basis of the number of CpGs present in the CpG island (Figure 1A, fragment with 149 CpGs), this difference was abolished when about 17% of CpG sites pL9.2 were pre-methylated with M.*Sss*I. As shown in Figure 1B, the methylation level of the region corresponding to the island in the pre-methylated plasmid increased to the same value (per CpG) seen for the remaining regions of the plasmid, whose relative position was not modified by the M.*Sss*I treatment. Therefore, a partially methylated island becomes a good substrate for the eukaryotic enzyme that methylates it as efficiently as any other DNA fragment, behaving in the same way as M.*Sss*I. This can be regarded as a general phenomenon because very similar results were obtained when other CpG islands were used. Plasmid pEB9 (58% G+C, 634 CpGs = 6.2% CpG) contains two CpG islands associated to two genes, *GdX* and *P3*, present on the human X chromosome. Both islands are

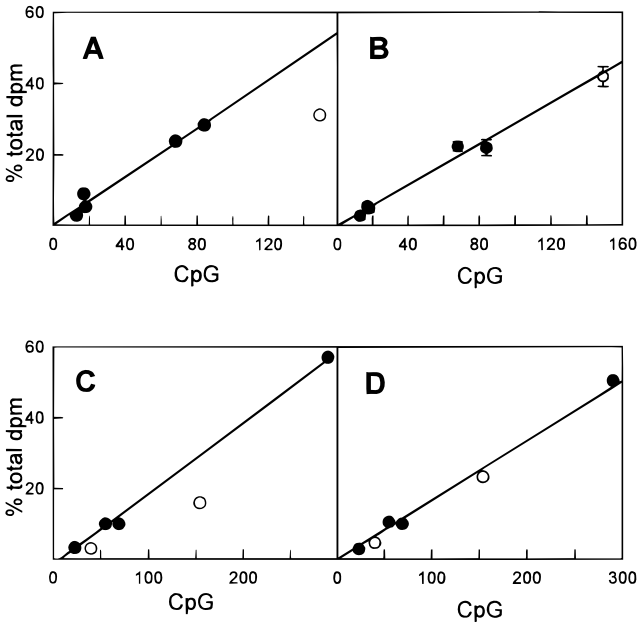

FIGURE 1: In vitro methylation pattern of unmethylated (A) and pre-methylated (B) pL9.2 plasmid and unmethylated (C) and pre-methylated (D) pEB9 plasmid. Plasmids were methylated with M.*Sss*I at 17% and 4% of all CpG sites, respectively, and then subjected to methylation by human placenta DNA MTase. Each point represents a restriction fragment or the sum of co-migrating fragments. Methylation is expressed, for each fragment, as the percentage of the total counts recovered from the gel. Open symbols indicate CpG island-derived fragments. CpGs are recorded on a single-strand basis.

poorly methylated by placenta DNA MTase [Figure 1C; also see Carotti et al. (1989)]. When M.*Sss*I-pre-methylated pEB9 (4% of the CpG sites) was used, the overall methylation level of the entire plasmid increased by approximately 40%, and the methylation pattern specifically changed only for the two restriction fragments derived from the CpG-rich DNA regions (Figure 1D).

The actual positions of the pre-methylated cytosines are difficult to assess due to the processivity of M.*Sss*I. After the pre-methylation step, heterogeneous DNA molecules will be present in the reaction mixture, including completely methylated DNA molecules (which will not be susceptible to further methylation by eukaryotic methylase and will therefore be undetectable in our radioactive assay), partially methylated molecules, and completely unmethylated molecules. Methyl groups added by M.*Sss*I or DNA MTase on unmethylated DNA molecules would not be distinguished from those added by DNA MTase on partially methylated molecules, by either sequencing or restriction enzyme analysis of the plasmid.

Since the placenta enzyme used in these experiments was only partially purified, the possibility could exist that some other protein be responsible for the observed effect. However, gel retardation analyses indicate that our enzyme preparation was free from *Htf*9-specific or Sp1 binding activities [data not shown; also see Carotti et al. (1989, 1996)].

In order to verify whether the stimulating effect due to 5-methylcytosine residues was dependent on their inclusion within CpG dinucleotides, we used, in the pre-methylation step with pL9.2, M.*Alu*I instead of M.*Sss*I. M.*Alu*I recognizes the AGCT sequence, and the resulting 5-methylcytosine is therefore not included in a CpG doublet. Subsequent methylation by the placenta enzyme and methylation pattern analyses were performed in the same conditions as above. The only difference was at restriction: *Pvu*II was omitted because the *Pvu*II recognition sequence contains the *Alu* site, and *Pvu*II cleavage is inhibited when the cytosine is methylated. The methylation pattern obtained with pre-methylated pL9.2 in these experiments was indistinguishable from that obtained with unmethylated pL9.2 (data not shown), indicating that, at least as far as CpG island in vitro methylation is concerned, the occurrence of 5-methylcytosine in a CpG doublet is strictly required to exert a stimulatory effect on eukaryotic DNA MTase.

Effect of Pre-methylation on in Vitro Methylation of DNAs with Varying Base Composition. The pre-methylation stimulating effect was further investigated on various DNAs with widely different base composition (listed in Table 1) ranging

Table 1: Effect of Pre-methylation on the Activity of Placenta DNA MTase with Various DNAs

pre-methylating enzyme	DNA	pre-methylation[a]	unmethylated DNA[b]	pre-methylated DNA[c]	-fold induction
M.*Sss*I	poly(dG-dC)	25.0 (1.7)	1.01	3.93	3.9
		43.2 (2.9)	1.20	8.65	7.2
	poly(dI-dC)	39.2 (2.6)	2.55	1.65	0.6
		67.7 (4.5)	2.70	1.33	0.5
	E. coli	25.9 (12.8)	0.37	1.67	4.5
		55.1 (27.1)	0.58	2.27	3.9
	M. luteus	30.9 (7.3)	0.92	1.56	1.7
		50.8 (12.0)	0.35	0.68	1.9
	pHbβ1S	15.8 (13.4)	0.10	0.37	3.7
placenta	poly(dG-dC)	3.4 (0.2)	1.24	2.75	2.2
	poly(dI-dC)	2.6 (0.2)	3.95	4.56	1.2
M.*Hae*III	pHbβ1S	1.8 (16.1)	0.31	0.38	1.2
		3.5 (32.2)	0.52	0.70	1.3
	pL9.2	4.8 (20.8)	0.39	0.42	1.1
		8.8 (38.3)	0.49	0.61	1.3
M.*Hha*I	poly(dG-dC)	15.1 (1.0)	0.67	2.82	4.2
	poly(dI-dC)	33.9 (2.2)	1.28	0.88	0.7

[a] pmol of [*methyl*-^{3}H] incorporated/μg of DNA during pre-methylation by the indicated enzyme; numbers in parentheses indicate the corresponding percentages of methylation of all possible sites. [b] pmol of [*methyl*-3H] incorporated/μg of DNA using the placenta DNA MTase. [c] pmol of [*methyl*-3H] incorporated/μg of DNA using the placenta DNA MTase, after pre-methylation as indicated in footnote a. Activities were measured as described in Materials and Methods, and values are the average of three determinations.

from poly(dG-dC)·poly(dG-dC) (G+C only) to the A+T-rich plasmid pHbβ1S (46% G+C, 3.9% CpG) which contains the 5 kb-long human β globin gene cloned in pBR322. The results of these experiments are shown in Table 1. In the first set of experiments, DNAs were pre-methylated by M.*Sss*I and then subjected to the action of placenta DNA MTase. The amount of methyl groups incorporated in the pre-methylation step depends on the number of CpGs present in the various DNAs, on the amount of M.*Sss*I, and on the length of the incubation time. Poly(dG-dC)·poly(dG-dC) was found to be most sensitive to the effect of pre-methylation, showing a 7-fold increase in its methyl-accepting ability following pre-methylation of about 2.8% of the CpG sites. Surprisingly, no pre-methylation-induced stimulation of poly(dI-dC)·poly(dI-dC) methylation was detected; on the contrary, we found a 50% inhibition when about 4% of the CpI sites were pre-methylated. Methylation of *E. coli* (50% C+G, 6.7% CpG) and of *M. luteus* (74% G+C, 14% CpG) DNAs showed intermediate levels of stimulation. Similar effects were also seen when placenta DNA MTase was used in both pre-methylation and methylation steps. Methylation of poly(dG-dC)·poly(dG-dC) was stimulated while that of poly(dI-dC)·poly(dI-dC) was only slightly affected; however, both effects were much less evident than those obtained with M.*Sss*I, probably because of the lower amount of methyl group incorporation achievable with placenta DNA MTase (Table 1). M.*Sss*I processively methylates DNA on each strand separately (Renbaum & Razin, 1992), probably yielding transiently hemimethylated sites. The presence of such hemimethylated sites, which would be very efficiently methylated by the eukaryotic DNA MTase, might in principle account for the stimulating effect of pre-incubation with M.*Sss*I. However the results with poly(dI-dC)·poly(dI-dC) ruled out this possibility since pre-methylation had an inhibitory, rather than a stimulatory, effect. In addition, we used M.*Hha*I in the pre-methylation step. This enzyme adds a methyl group to the internal cytosines in the GCGC or ICIC sequences, on both DNA strands. Poly(dG-dC)·poly(dG-dC) and poly(dI-dC)·poly(dI-dC) were chosen as model substrates in order to avoid misleading effects due to a preferential distribution of the methylation sites. The experiment confirmed the results obtained with M.*Sss*I (Table 1); namely that the presence of 5-methylcytosines stimulates the activity of eukaryotic DNA MTase only on poly(dG-dC)·poly(dG-dC) and exerts an opposite effect on poly(dI-dC)·poly(dI-dC). Thus, poly(dI-dC)·poly(dI-dC), in the unmethylated form, is a far better substrate for eukaryotic DNA MTase than poly(dG-dC)·poly(dG-dC) (Pedrali-Noy & Weissbach, 1986), while the situation is reversed after pre-methylation.

Then, we finally tested the effect of M.*Hae*III, another cytosine methylase whose recognition sequence, GGCC, contains no CpG site. The pL9.2 and pHbβ1S plasmids contain 49 and 32 *Hae*III sites, respectively. In either case, pre-methylation with M.*Hae*III had very little, if any, effect on the activity of eukaryotic DNA MTase (Table 1), as previously seen with M.*Alu*I, indicating that the extra 5-methylcytosines must be present in a CpG dinucleotide in order to affect the activity of the eukaryotic enzyme.

Effect of the Levels of Pre-methylation. We investigated the effect of the pre-methylation level on the subsequent activity of eukaryotic DNA MTase with three different substrates, chosen because of their different behavior: poly(dG-dC)·poly(dG-dC), plasmid pHbβ1S, and poly(dI-dC)·poly(dI-dC). Upon pre-methylation with M.*Sss*I, the stimulating effect seen with poly(dG-dC)·poly(dG-dC) became apparent only above a threshold level of 30 pmol/μg of DNA of pre-methylation (Figure 2A), with a nonlinear correlation to the percentage of CpG methylation reached in the pre-methylation step, possibly suggesting some cooperative effect. When the same experiment was performed with plasmid pHbβ1S, there was an initial stimulation, followed by a severe inhibition of the activity of DNA MTase, which dropped to 5% of the values obtained for control DNA, when 37% of the CpGs were methylated (Figure 2B). Methylation of poly(dI-dC)·poly(dI-dC) showed a still different behavior: the inhibition observed with this polymer was directly, though not linearly, related to the extent of pre-methylation and never dropped below 50% of the value obtained with the unmethylated sequence (Figure 2C). In the experiments with synthetic polymers, we controlled whether the mild heat treatment used to destroy M.*Sss*I activity could induce per se a stimulating effect, for example by favoring unusual DNA

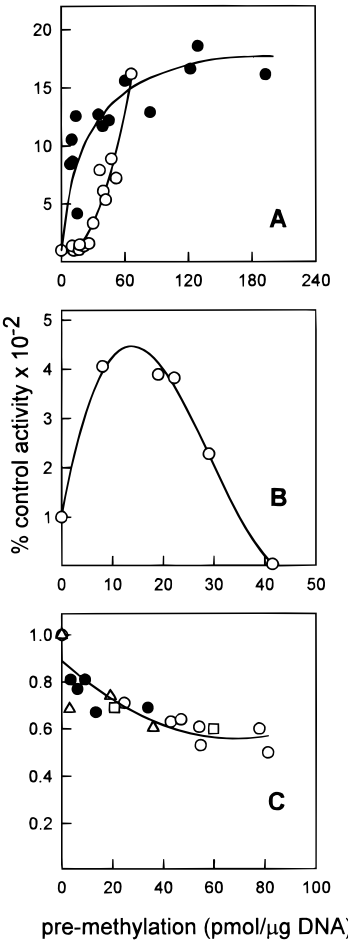

FIGURE 2: Effect of the pre-methylation levels on the activity of human placenta DNA MTase with poly(dG-dC)·poly(dG-dC) (A), with plasmid pHbβ1S (B), and with poly(dI-dC)·poly(dI-dC) (C); open circles refer to M.*Sss*I-pre-methylated polymers, and closed circles refer to M.*Hha*I-pre-methylated ones. Each curve derives from the combination of more than one experiment, and each point is the mean of three determinations. In Figure 2C the effect of various levels of M.*Sss*I pre-methylation on the activity of M.*Sss*I on poly(dG-dC)·poly(dG-dC) (open triangles) and on poly(dI-dC)·poly(dI-dC) (open squares) is also shown. Pre-methylation is expressed as pmol of [*methyl*-^{3}H] incorporated/μg of DNA.

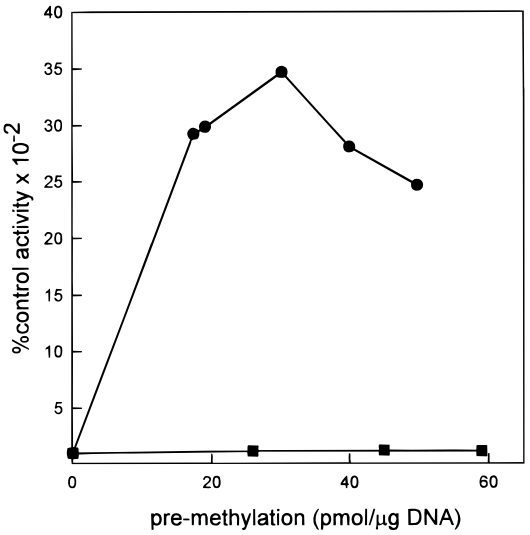

FIGURE 3: Effect of the pre-methylation levels on the activity of murine liver DNA MTase with poly(dG-dC)·poly(dG-dC) (●) and with poly(dI-dC)·poly(dI-dC) (■). Both polymers were pre-methylated with M.*Hha*I. Each point is the mean of two determinations. Pre-methylation is expressed as pmol of [*methyl*-3H] incorporated/ μg of DNA.

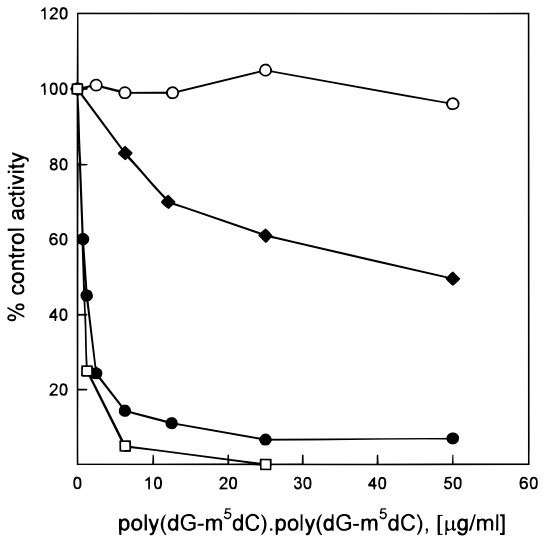

FIGURE 4: Effect of poly(dG-m^5dC)·poly(dG-m^5dC) on the activity of human placenta DNA MTase (●), of murine liver DNA MTase (□), of M.*Hha*I (○), and of M.*Sss*I (◆) with 20 μg/mL poly(dG-dC)·poly(dG-dC) as substrate. Each point is the mean of two determinations.

structures that are known to be preferred sites of methylation by DNA MTase (Smith et al., 1992). We have found no difference in the methylation when the mild heat treatment was omitted (data not shown).

When using M.*Hha*I in the pre-methylation step, the results were significantly different from those seen with M.*Sss*I, with poly(dG-dC)·poly(dG-dC), and no cooperative effect could be shown (Figure 2A, closed symbols).

The inhibitory effect exerted on poly(dI-dC)·poly(dI-dC) by pre-methylation with M.*Hha*I was instead indistinguishable from that obtained with M.*Sss*I. When M.*Sss*I was used in both steps, the activities of the enzyme were inhibited to a similar extent (Figure 2C) on poly(dG-dC)·poly(dG-dC) and on poly(dI-dC)·poly(dI-dC). M.*Sss*I methylates hemimethylated and unmethylated substrates with the same efficiency, and no effect should therefore be seen on premethylated DNA if it was present in a hemimethylated configuration. The lower activity seen on premethylated poly(dG-dC) is likely to reflect the lower affinity of M.*Sss*I for double-stranded methylated DNA (O'Gara et al., 1996) and can be seen as a further control for the absence of hemimethylated sites.

The differences between the effects of poly(dG-dC)·poly(dG-dC) pre-methylation by M.*Hha*I and M.*Sss*I can be tentatively ascribed to their different processivity—higher for M.*Sss*I and lower or absent for M.*Hha*I (Renbaum & Razin, 1992). At the low pre-methylation levels used in our experimental conditions, M.*Sss*I can be expected to yield both fully methylated (or patch-methylated) and unmethylated DNA molecules; when M.*Hha*I is used, only randomly methylated molecules should be present. As shown in Figure 2A, subsequent methylation by eukaryotic DNA MTase appears to be favored by a random distribution of 5-methylcytosines compared to processive methylation of adjacent residues. These experiments were in part repeated using the DNA MTase extracted from murine liver nuclei (Figure 3). Quite similar results were obtained: methylation of M.*Hha*I-premethylated poly(dG-dC)·poly(dG-dC) was highly stimulated, while that of M.*Hha*I-premethylated poly(dI-dC)·poly(dI-dC) was unaffected, suggesting that such an opposite behavior with the two oligodeoxyribonucleotides is a general feature of eukaryotic enzymes.

Poly(dG-m^5dC)·Poly(dG-m^5dC) Inhibits in Vitro Methylation. Given the different affinities shown by prokaryotic and eukaryotic methylases toward methylated DNA, the two classes of enzymes can be expected to be differentially sensitive to the presence of fully methylated polymers. Eukaryotic methylases were indeed shown to be highly inhibited by poly(dG-m^5dC)·poly(dG-m^5dC) (Carotti et al., 1986; Bolden et al., 1986), while no data are available so far on prokaryotic methylases. We therefore tested the effect of poly(dG-m^5dC)·poly(dG-m^5dC) on the activity of M.*Hha*I and of M.*Sss*I, compared to that of placenta and murine DNA MTases, using poly(dG-dC)·poly(dG-dC) as the substrate. As shown in Figure 4, the activities of placenta and murine DNA MTases were reduced to 15 and 5% of control, respectively, in the presence of 6 μg/mL poly(dG-m^5dC)·poly(dG-m^5dC), as expected; that of M.*Sss*I was much less affected, being 80% of the control at the same polymer concentration, while M.*Hha*I activity was not affected at all.

DISCUSSION

In the present work we have shown that the activity of eukaryotic DNA MTase can be modified following partial pre-methylation of the DNA substrate. The observed effects depend indeed on the DNA base composition. Poly(dI-dC)·poly(dI-dC) becomes, in the partially methylated form, a poor substrate for eukaryotic DNA MTase, while the activity of the enzyme on a (G+C)-rich DNA, such as the synthetic poly(dG-dC)·poly(dG-dC) polymer or *M. luteus* DNA, is positively influenced by partial pre-methylation. If pre-methylated, CpG islands loose their reluctance to be good substrates for the eukaryotic enzyme, and thereby methylation of these sequences undergoes a greater stimulation than that seen with neighboring sequences. On the contrary, an A+T-rich DNA such as the plasmid pHbβ1S, which contains a

3.5 kb-long insert of genomic DNA of approximately 60% (A+T) content, becomes a very poor substrate for the enzyme when about one-third of the CpGs are pre-methylated. All these effects appear to be specific to the eukaryotic enzyme, M.*Sss*I having a consistently lower level of activity toward pre-methylated substrates.

Quantitatively, the extent of stimulation due to partial pre-methylation of double-stranded DNA is lower than that seen in presence of a hemimethylated site. This is not surprising, since maintenance methylation is the main activity of DNA MTase, while de novo methylation occurs mainly during early embryogenesis. Recent evidence indeed suggests the existence of a distinct de novo activity during gametogenesis (Tucker et al., 1996). We have studied instead the de novo activity of the enzyme species present in adult tissues, whose incorrect control might give origin to the altered methylation patterns found in neoplastic cells. In recent experiments using partially methylated synthetic oligodeoxyribonucleotides, Tollefsbol and Hutchinson (1995, 1997) have shown a stimulating effect of the same order of magnitude as those reported here, or smaller.

Stimulation of in vitro de novo methylase activity by pre-existing methyl groups has previously been reported with single-stranded oligodeoxyribonucleotides (Carotti et al., 1986; Christman et al., 1995; Lindsay & Adams, 1996). The present data show that stimulation by pre-existing methyl groups also takes place on double-stranded DNA, particularly of high CpG content. In retrospect, the present findings might provide a molecular basis for the observation that the activity of human placenta DNA MTase using a plasmid containing the *Ha-ras*-1 gene promoter, which is a CpG island, was stimulated after previous methylation by M.*Hpa*II and M.*Hha*I (Rachal et al., 1989).

The studies with single-stranded DNA previously suggested that the enzyme recognizes not only the cytosine to be methylated but also another site that contains another cytosine or 5-methylcytosine moiety. The two sites need not be on the same DNA segment (Lindsay & Adams, 1996). A second putative DNA-binding site has been identified recently in the N-terminal domain of cloned human DNA MTase (Chuang et al., 1996), and has been suggested to be involved in sensing the appropriate length of newly synthesized DNA, before methylation. Since some single-stranded oligodeoxyribonucleotides, used in the previous studies, contained short stretches of complementary sequences; the differences observed upon introduction of 5-methylcytosine moieties have been related to the possibility of formation of mispaired 5-methylcytosines in a stem-loop structure (Christman et al., 1995). We believe that in our experiments that possibility can be ruled out, given that very similar results were obtained with both supercoiled CpG-rich plasmids and with linear synthetic poly(dG-dC)·poly(dG-dC), i.e., a polymer in which the formation of such structures is highly unlikely, in our experimental conditions. Moreover, similar structures, if forming at all, should be present in poly(dI-dC)·poly(dI-dC), where instead pre-methylation has an inhibitory rather than a stimulatory effect.

The present data do not pinpoint the molecular mechanisms by which the eukaryotic DNA MTase responds to the presence of pre-existing 5-methylcytosines. However, it is tempting to speculate that a second DNA binding site with regulatory functions may exist which may sense the presence of "regulatory" 5-methylcytosines, even in double-stranded DNA.

The ability of such paired methylated CpGs to stimulate DNA MTase suggests that the so-called methylation spreading (i.e., the appearance of de novo methylated cytosine moieties in genomic regions which already contain methylated cytosines), which has been observed to occur *in vivo* upon viral integration in genomic DNA (Toth et al., 1990) or in repetitive sequence regions such as the CCG repeat present in the fragile X syndrome (Oberlé et al., 1991), does not necessarily require the existence of single-stranded DNA regions capable of activating methylation at distant sites. It is possible that methylation centers, from which de novo methylation is thought to spread at a fixed distance (Hasse & Schulz, 1994; Mummaneni et al., 1995), adopt in vivo unusual structures that have been shown to be preferentially methylated by the DNA MTase (Smith et al., 1992), and thus they acquire the ability to stimulate the subsequent methylation of adjacent sequences.

We would like to suggest a possible mechanism leading to CpG island methylation upon prolonged cell culture or tumor transformation. DNA MTase levels tend to be high in actively replicating (Szyf et al., 1985) and neoplastic (Kautiainen & Jones, 1986; Issa et al., 1993; Belinsky et al., 1996; Lee et al., 1996) cells. In these conditions of abnormally high enzyme activity, sporadic de novo methylation might occur on DNA sequences otherwise not subjected to methylation (i.e., on CpG islands). CpG island DNA has in fact been shown to be susceptible to methylation, though with a somewhat decreased efficiency compared to conventional DNA sequences (Carotti et al., 1989). Partially methylated sequences would become a very suitable substrate for the enzyme which will thereafter methylate them as efficiently as any other sequence (Figure 1). In this framework, it is interesting that recent data have shown that certain CpG islands become methylated in cells overexpressing DNA MTase; a non-random pattern of methylation has been observed in both human fibroblasts (Vertino et al., 1996) and in human fibroblast × fibrosarcoma hybrids (Kuerbitz & Baylin, 1996) that express high levels of DNA MTase. In both cases, locus-specific features and/or clonal variations were suggested to account for the apparent specificity of methylation, beside the high levels of DNA MTase. The present experiments support the relevance of the sequence context in which the extra methyl groups need to be included in order to exert their stimulating effect. In a CpG island, CpGs are separated by approximately 11 nucleotides, on a statistical basis, a number roughly corresponding to the helix pitch. The activity of eukaryotic DNA MTase could be impaired (as shown by the reduced methyl-accepting capability of CpG islands) when the enzyme, in its processive scanning of the DNA double helix, meets, with its two hypothetical DNA binding sites, two unmethylated CpGs located at a suitable distance on the same side of the double helix. If, instead, one of the two sites is occupied by a methylated CpG, the methylation of the other one is favored. The results obtained with the pHbβ1S plasmid suggest that when the two CpG sites are very far apart the stimulatory effect is no more effective, and the resulting methylated DNA is less efficiently methylated because of its lower affinity for the enzyme. In support of this hypothesis, in vitro proteolytic cleavage of the N-terminal

regulatory domain of murine DNA MTase stimulates the de novo activity of the C-terminal domain, but interestingly such stimulation is not seen with poly(dI-dC)·poly(dI-dC) (Bestor, 1992). M.*Sss*I, that lacks the N-terminal domain (Renbaum et al., 1990), is capable of methylating CpG island DNA as efficiently as any other DNA. Accordingly, the inhibitory effect of poly(dG-m^{5}dC)·poly(dG-m^{5}dC) is much more evident on eukaryotic than on prokaryotic methylases. There is also an apparent inverse correlation between the extent of stimulation by pre-methylation and extent of inhibition by poly(dG-m^{5}dC)·poly(dG-m^{5}dC), as shown by the comparison of the similar sets of data obtained with human and murine DNA MTase (Figures 3 and 4). Prokaryotic enzymes perform both de novo and maintenance methylation with the same efficiency and do not show any stimulation by pre-existing methylation, thus rather exhibit a lower affinity for methylated DNA, which becomes a worse substrate than its unmethylated counterpart.

How extra methyl groups might specifically stimulate methylation of a G+C-rich and CpG-rich stretch of DNA remains to be clarified. CpG methylation has been shown to have various effects on the DNA structure. Methylation reportedly stabilizes the Z form of poly(dG-dC)·poly(dG-dC) (Behe & Felsenfeld, 1981), increases the T_m of G·C base pairs by several degrees (Szer & Shugar, 1966), and can induce a modification of the local structure of DNA (Hodges-Garcia & Hagerman, 1992) or of DNA flexibility (Hodges-Garcia & Hagerman, 1995), in particular, by increasing the helicoidal torsion (Kerwood et al., 1991). Studies with synthetic polymers of defined sequence have shown that the curvature of a short (dA)*n*·(dT)*n* DNA segment is altered by the presence of 5-methylcytosine(s) within the segment itself or at a certain distance from it. This effect depends on both the DNA sequence and on the position of the methylated cytosine(s) (Hodges-Garcia & Hagerman, 1992). ^{1}H-NMR and ^{31}P-NMR studies on the structural and thermal stability behavior of CpG-containing octanucleotides, in both unmethylated (El-Antri et al., 1993; Lefebvre et al., 1995a) and methylated forms (Lefebvre et al., 1995b), have shown effects strongly dependent on the nature of the flanking steps. Since most studies have been performed on A+T-rich oligomers, the results cannot, unfortunately, be extrapolated to our G+C-rich substrates. However, they clearly indicate that CpG methylation is able to cause a local alteration of the DNA structure in a way dependent on the DNA sequence that contains the methylated cytosine(s).

Eukaryotic DNA MTase reportedly prefers, with a reaction mechanism similar to that hypothesized for bacterial M.*Hha*I (Klimasauskas et al., 1994), mispaired cytosines (Smith, 1994), possibly because their rotation out of the DNA helix requires less energy than rotation of those which are normally hydrogen bonded. In double-stranded DNA, CpG methylation is likely to induce a local structural alteration capable of increasing the accessibility of adjacent cytosine, included in a CpG dinucleotide, to the further action of the eukaryotic DNA MTase.

A final point concerns the peculiar behavior of poly(dI-dC)·poly(dI-dC). Why is this polymer the preferred substrate for the de novo methylation reaction, and is instead extremely unfavored in its partially methylated form? As compared to its G-containing counterpart, poly(dI-dC)·poly(dI-dC) is characterized by the occurrence of two (instead of three) hydrogen bonds between paired bases, by the absence of the guanine NH_2 group, that lies on the bottom of the minor groove, and by a somewhat narrower minor groove. Poly-(dI-dC)·poly(dI-dC) has therefore some structural similarity to poly(dA-dT)·poly(dA-dT). Might these features account for the preference shown by the eukaryotic DNA MTase? Some motifs present in the N-terminal domain of the eukaryotic enzyme (Bestor et al., 1992) are thought to preferentially bind the minor groove of A+T-rich DNAs (Churchill & Suzuki, 1989). In addition, we have shown (Carotti et al., 1996) that only the eukaryotic enzyme is inhibited by histone H1, which also contains similar peptide sequences. At variance with poly(dA-dT)·poly(dA-dT), however, poly(dI-dC)·poly(dI-dC) is susceptible of being methylated. It will be worth investigating which structural changes are brought about by the introduction of methyl groups on cytosine moieties and why they make it a bad substrate for eukaryotic DNA MTase.

ACKNOWLEDGMENT

We thank Patrizia Lavia for critical reading of the manuscript. We are grateful to Vincenzo Peresempio for technical assistance.

REFERENCES

Antequera, F., Boyes, J., & Bird, A. (1990) *Cell 62*, 503−514.

Behe, M., & Felsenfeld, G. (1981) *Proc. Natl. Acad. Sci U.S.A. 78*, 1619−1623.

Belinsky, S. A., Nikula, K. J., Baylin, S. B., & Issa, J. P. J. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 4045−4050.

Bestor, T. H. (1992) *EMBO J. 7*, 2611−2617.

Bestor, T. H., Gundersen, G., Kolsto, A. B., & Prydz, H. (1992) *Genet. Anal.: Tech. Appl. 9*, 48−53.

Bird, A. P. (1986) *Nature 321*, 209−213.

Bolden, A. H., Ward, C. A., Nalin, C. M., & Weissbach, A. (1986) *Prog. Nucleic Acid Res. Mol. Biol. 33*, 231−250.

Brandeis, M., Dale, F., Keshet, I., Zahava, S., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., & Cedar, H. (1994) *Nature 371*, 435−438.

Bressan, A., Somma, M. P., Lewis, J., Santolamazza, C., Copeland, N., Gilbert, D., Jenkins, N., & Lavia, P. (1991) *Gene 103*, 201−209.

Carotti, D., Palitti, F., Mastrantonio, S., Rispoli, M., Strom, R., Amato, A., Campagnari, F., & Whitehead, E. P. (1986) *Biochim. Biophys. Acta 866*, 135−143.

Carotti, D., Palitti, F., Lavia, P., & Strom, R. (1989) *Nucleic Acids Res. 17*, 9219−9229.

Carotti, D., Funiciello, S., Lavia, P., Caiafa, P., & Strom, R. (1996) *Biochemistry 35*, 11660−11667.

Christman, J. K., Sheikhnejad, G., Marasco, C. J., & Sufrin, J. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7347−7351.

Chuang, L. S. H., Ng, H. H., Chia, J. N., & Li, B. F. L. (1996) *J. Mol. Biol. 257*, 935−948.

Churchill, M. E. A., & Suzuki, M. (1989) *EMBO J. 8*, 4189−4195.

Collins, F. S., & Weissman, S. (1984) *Prog. Nucleic Acid Res. Mol. Biol. 31*, 315−462.

El Antri, S., Mauffret, O., Monnot, M., Lescot, E., Convert, O., & Fermandjian S. (1993) *J. Mol. Biol. 230*, 373−378.

Gonzales-Zulueta, M., Bender, C. M., Yang, A. S., Nguyen, T., Beart, R. W., Van Tornout, J. M., & Jones, P. A. (1995) *Cancer Res. 55*, 4531−4535.

Greger, V., Passarge, E., Hopping, W., Messmer, E., & Horsthemke, B. (1989) *Hum. Genet. 83*, 155−158.

Hasse, A., & Schulz, W. A. (1994) *J. Biol. Chem. 269*, 1821−1826.

Herman, J. G., Latif, F., Weng, Y., Lerman, M. J., Zbar, B., Liu, S., Samid, D., Duan, D-S. R., Gnarra, J. R., Linehan, W. M., & Baylin, S. B. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 9700−9704.
Hodges-Garcia, Y., & Hagerman, P. J. (1992) *Biochemistry 31*, 7595−7599.
Hodges-Garcia, Y., & Hagerman, P. J. (1995) *J. Biol. Chem. 270*, 197−201.
Issa, J. J., Vertino, P. M., Wu, J., Sazawal, S., Celano, P., Nelkin, B. D., Hamilton, S. R., & Baylin, S. (1993) *J. Natl. Cancer Inst. 85*, 1235−1240.
Kautiainen, T. L., & Jones, P. (1986) *J. Biol. Chem. 261*, 1594−1598.
Kerwood, D. J., Zon, G., & James, T. L. (1991) *Eur. J. Biochem. 197*, 583−595.
Klimasauskas, S., Kumar, S., Roberts, R. J., & Cheng, X. (1994) *Cell 76*, 357−369.
Knight, S. J. L., Flannery, A. V., Hirst, M. C., Campbell, L., Christodoulou, Z., Phelps, S. R., Pointon, J., Middleton-Price, H. R., Barnicoat, A., Pembrey, M. E., Holland, J., Oostra, B. A., Bobrow, M., & Davies, K. E. (1993) *Cell 74*, 127−134.
Kuerbitz, S. J., & Baylin, S. B. (1996) *Cell Growth Differ. 7*, 847−853.
Lavia, P., Macleod, D., & Bird, A. (1987) *EMBO J. 6*, 2773−2779.
Lee, P. J., Washer, L. L., Law, D. J., Boland, C. R., Horon, I., & Feinberg, A. P. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 10366−10370.
Lefebvre, A., Mauffret, O., Hartmann, B., Lescot, E., & Fermandjian, S. (1995a) *Biochemistry 34*, 12019−12028.
Lefebvre, A., Mauffret, O., El Antri, S., Monnot, M., Lescot, E., & Fermandjian, S. (1995b) *Eur. J. Biochem. 229*, 445−454.
Li, E., Beard, C., & Jaenisch, R. (1993) *Nature 366*, 451−456.
Lindsay, H., & Adams, R. (1996) *Biochem. J. 320*, 473−478.
Macleod, D., Charlton, J., Mullins, J., & Bird, A. P. (1994) *Genes Dev. 8*, 2282−2292.
Monk, M. (1995) *Dev. Genet. 17*, 188−197.
Mummaneni, P., Walker, K. A., Bishop, P. L., & Turker, M. L. (1995) *J. Biol. Chem. 270*, 788−792.
Oberlé, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boué, J., Bertheas, M. F., & Mandel, J. L. (1991) *Science 252*, 1097−1102.
O'Gara, M., Roberts, R. J., & Cheng, X. (1996) *J. Mol. Biol. 263*, 597−606.
Pedrali-Noy, G., & Weissbach, A. (1986) *J. Biol. Chem. 261*, 7600−7602.
Rachal, M. J., Yoo, H., Becker, F. F., & Lapeyre, J. N. (1989) *Nucleic Acids Res. 17*, 5135−5147.
Renbaum, P., & Razin, A. (1992) *FEBS 313*, 243−247.
Renbaum, P., Abrahamove, D., Fainsod, A., Wilson, G. G., Rotten, S., & Razin, A. (1990) *Nucleic Acids Res. 18*, 1145−1152.
Singer-Sam, J., & Riggs, A. (1993) in *DNA Methylation: Molecular Biology and Biological Significance* (Jost, J. P., & Saluz, H. P., Eds.) pp 358−384, Birkhaeuser Verlag, Basel, Switzerland.
Smith, S. S. (1994) *Prog. Nucleic Acid Res. Mol. Biol. 49*, 65−111.
Smith, S. S., Kaplan, B. E., Sowers, L. C., & Newman, E. M. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4744−4748.
Szer, W., & Shugar, D. (1966) *J. Mol. Biol. 17*, 174−187.
Szyf, M. (1996) *Pharmacol. Ther. 70*, 1−37.
Szyf, M., Kaplan, F., Mann, V., Giloh, H., Kedar, E., & Razin, A. (1985) *J. Biol. Chem. 260*, 8653−8656.
Tollefsbol, T. O., & Hutchinson III, C. A. (1995) *J. Biol. Chem. 270*, 18543−18550.
Tollefsbol, T. O., & Hutchinson III, C. A. (1997) *J. Mol. Biol. 269*, 494−504.
Toniolo, D., Persico, M., & Alcalay, M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 851−855.
Toth, M., Muller, U., & Doerfler, W. (1990) *J. Mol. Biol. 214*, 673−683.
Tucker, K. L., Bears, C., Dausman, J., Jackson-Grusby, L., Laird, P. W., Lei, H., Li, E., & Jaenisch R. (1996) *Genes Dev. 10*, 1008−1020.
Vertino, P. M., Yen, R. W. C., Gao, J., & Baylin S. B. (1996) *Mol. Cell. Biol. 16*, 4555−4565.

BI971031I