1038–1045 *Nucleic Acids Research, 1998, Vol. 26, No. 4*

Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci

Yongliang Liu[+], Edward J. Oakeley, Lijie Sun[§] and Jean-Pierre Jost*

Friedrich Miescher-Institut, P.O. Box 2543, Basel, Switzerland

Received October 3, 1997; Revised and Accepted December 19, 1997

ABSTRACT

It has been shown that, during the S-phase of the cell cycle, the mouse DNA methyltransferase (DNA MTase) is targeted to sites of DNA replication by an amino acid sequence (aa 207–455) lying in the N-terminal domain of the enzyme [Leonhardt, H., Page, A. W., Weier, H. U. and Bestor, T. H. (1992) *Cell*, 71, 865–873]. In this paper it is shown, by using enhanced green fluorescent protein (EGFP) fusions, that other peptide sequences of DNA MTase are also involved in this targeting. The work focuses on a sequence, downstream of the reported targeting sequence (TS), which is homologous to the Polybromo-1 protein. This motif (designated as PBHD) is separated from the reported targeting sequence by a zinc-binding motif [Bestor , T. H. (1992) *EMBO J*, 11, 2611–2617]. Primed *in situ* extension using centromeric-specific primers was used to show that both the host DNA MTase and EGFP fusion proteins containing the targeting sequences were localized to centromeric, but not telomeric, regions during late S-phase and mitosis. Also found was that, in ~10% of the S-phase cells, the EGFP fusions did not co-localize with the centromeric regions. Mutants containing either, or both, of these targeting sequences could act as dominant negative mutants against the host DNA MTase. EGFP fusion proteins, containing the reported TS (aa 207–455), were targeted to centromeric regions throughout the mitotic stage which lead to the discovery of a similar behavior of the endogenous DNA MTase although the host MTase showed much less intense staining than in S-phase cells. The biological role of the centromeric localization of DNA MTase during mitosis is currently unknown.

INTRODUCTION

The mouse DNA methyltransferase (MTase) (3–5) contains a long N-terminal regulatory domain of >1000 amino acids (aa) and a shorter C-terminal catalytic domain of 500 aa. An additional 118 aa peptide at the N-terminus of this enzyme has been described by Yoder *et al.* (6) and Tucker *et al.* (7). The start codon of this newly defined N-terminus was the only one present in the murine DNA MTase purified from MEL cells (8). Within the N-terminal domain, there is a DNA replication foci-targeting sequence (aa 207–455) that targets the DNA MTase to the sites where its preferred substrate, hemimethylated DNA, is being synthesized (1). There is a major phosphorylation site (Ser396) lying in this motif which was identified in the DNA MTase from the MEL cells (8). Downstream of this targeting sequence is a short sequence, that has been shown to bind Zn^{2+}, which is homologous to the zinc-binding motif of ALL/TRX proteins (4). This is followed by a stretch of sequence that is homologous to the Polybromo-1 protein (4,9). This sequence designated as Polybromo-1 protein homologous domain (PBHD) has 23% identity in a 270 aa overlap with the Polybromo-1 protein (Liu, unpublished FASTA search). The function of this sequence motif has not yet been defined.

The distribution of DNA MTase changes dynamically in a cell cycle-dependent manner (10,11). By immunostaining of NIH3T3 cells with specific antibodies, it has been shown that the mouse DNA MTase forms toroidal structures in middle and late S-phase cells. In G1 and early S-phase cells the enzyme showed a diffused distribution (1). These S-phase toroidal structures were shown, by *in situ* hybridization to γ-satellite DNA probes, to be at the position of centromeric heterochromatin (1).

The B1 sequence (aa 202–369) (containing DNA binding motifs DB1 and AZn) from the N-terminal domain of human DNA MTase has both zinc and DNA-binding activity (12) and its murine homologue (aa 201–377) is included in its known DNA replication-targeting sequence (12). Constructs over expressing this targeting sequence may therefore have dominant negative effects by competing with the endogenous DNA MTase for the DNA substrate or other factors that recruit the DNA MTase to the DNA replication foci. Measuring changes in the level of genomic DNA methylation could assess this effect. Such a dominant negative mutant could be a useful tool for reducing the level of genomic DNA methylation, particularly as most human tumor

*To whom correspondence should be addressed. Tel: +41 61 697 6688; Fax: +41 61 697 6687; Email: jost@fmi.ch

Present addresses: [+]Department of Cell and Developmental Biology, University of Pennsylvania School of Medicine, 608 Stellar Chance Labs, 422 Curie Boulevard, Philadelphia, PA 19104-6100, USA and [§]Department of Dermatology, University of Pennsylvania School of Medicine, 230 Clinical Research Building, 422 Curie Boulevard, Philadelphia, PA 19104-6100, USA

cells show elevated levels of DNA MTase (13), a dominant negative mutant of DNA MTase could be of therapeutic value.

In this study, the tetracycline-regulated system (14) was used to control the expression of different mouse DNA MTase deletion mutants. These mutants were made as N-terminal fusions with the reporter enhanced green fluorescent protein (EGFP) (Clontech) to provide a fluorescent marker for studying their subcellular localizations. The effects of the expression of these mutants on the DNA methylation level of C3H10T½ were studied by using a modified *Sss*I methyl-accepting assay (15).

MATERIALS AND METHODS

Materials and plasmids

The EGFP N-fusion series of vectors were purchased from Clontech. pMC1NeoPolyA which contained the neomycin resistance gene (*neoR*) was obtained from Stratagene. Dr K. Ballmer-Hoffer provided the plasmid pX343 that contained a hygromycin resistance selectable marker. The autoregulatory tetracycline regulated expression system (14) was a gift from Dr D. Pollays (Life Technologies Inc., USA). Oligonucleotides were synthesized in-house by the FMI oligonucleotide Synthesis Lab. G418 sulfate was purchased from GibcoBRL. Restriction enzymes and the Klenow fragment of DNA polymerase I were obtained from Biofinex. The rapid DNA ligation kit, *Taq* polymerase and Texas Red isothiocyanate (TRITC)-labeled dUTP were purchased from Boehringer Mannheim. Calf intestine alkaline phosphatase was obtained from NEB. Collagen R solutions were purchased from SERVA. The secondary antibodies used for indirect antibody staining were purchased from Sigma and Boehringer Mannheim. BrdU, anti-BrdU monoclonal antibodies and Hygromycin B were obtained from CalBiochem. DAPI was purchased from Fluka. All plasmids were prepared with QIAGEN® plasmid Midi kit (Qiagen GmbH).

Construction of EGFP fusion protein expression vectors

Part of the newly discovered N-terminal coding sequence (6–8) was amplified by PCR using two overlapping oligonucleotides and cloned upstream of the *Eco*RI site of the reported cDNA sequence to generate the full length of DNA MTase coding sequence (Fig. 1, WT). The primers used were: 5′-TCTGTCG**CTCGAG**TCGCCACCATGCCAGCGCGAACAGCTCCAGCCCGAGTGCCTGCGCTTGCCTCCCCGGCAGGCTC**GCTCCCGGACCATGTCCGCAGGCGGC**-3′ (forward primer) and 5′-TTTGCAG**GAATTC**ATGCAGTAAGTTTAATTTTCCCTCACACACTCCTTTTCTGTTAAGCCATCTCTTTCCAAGTCTTTGA**GCCGCCTGCGGACATGGTCCGGGAGC**-3′ (reverse primer). The complementary sequence of the above oligonucleotides and the designed *Xho*I and *Eco*RI sites are in bold. The solutions for PCR amplification consisted of 5 µl 10× PCR buffer (Boehringer Mannheim), 5 µl 2 mM dNTPs, 2 µl of each of the above primers (10 µM), 35.5 µl ddH_2O and 0.5 µl Vent polymerase (10 U/µl, New England Biolabs). Amplification was performed in a Stratagene Robocycler for 30 cycles (94°C 1 min, 58°C 1 min and 72°C 40 s). The amplified fragment (183 bp) was digested with *Xho*I and *Eco*RI and ligated to the *Xho*I–*Eco*RI digested construct 2 (Fig. 1).

The deletion of the earlier reported targeting sequence was achieved by replacing the *Nsi*I–*Bgl*II fragment with the following short double stranded oligonucleotides containing *Nsi*I and *Bgl*II overhangs without changing any of the surrounding amino acids: 5′-TAGGAAGCTGGTCTATCA-3′ (upper strand) and 5′-ACGTATCCTTCGACCAGATAGTCTAG-3′ (lower strand).

Different fragments of the cDNA encoding the N-terminal domains of DNA MTase were cloned in the N-fusion EGFP vector (Clontech) (Fig. 1). Those constructs that lacked intrinsic nuclear localization signals (NLS) had the SV40 T antigen NLS added, in-frame, to their N-termini. The expression cassettes were then excised and inserted into the plasmid pTet-Splice down-

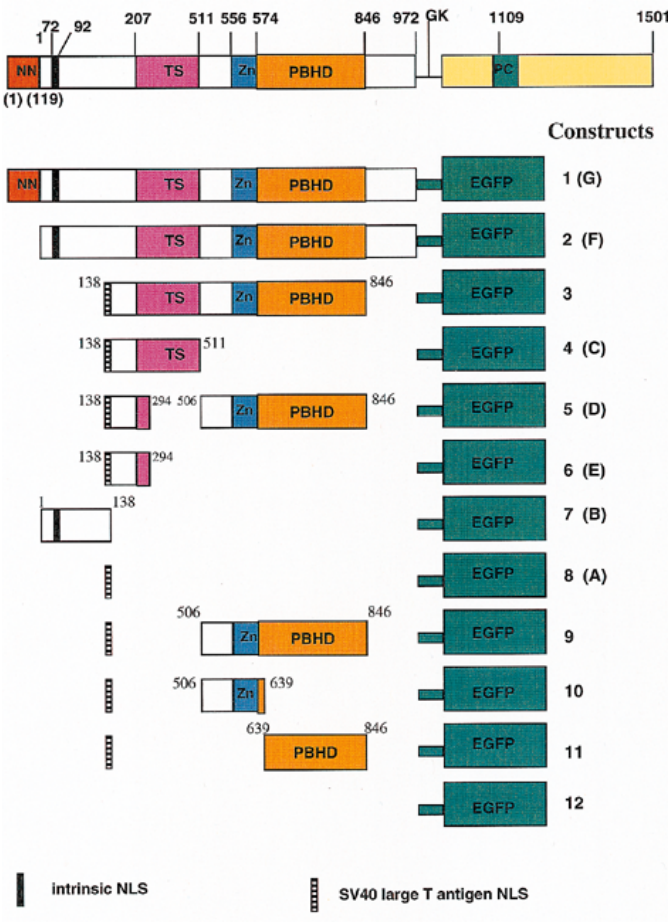


Figure 1. DNA MTase-EGFP fusion construct maps. The map at the top represents the full-length wild-type mouse DNA MTase. 'NN' refers to the new N-terminal sequence that has recently been described for the mouse MTase (6,7) with the numbers underneath in the parenthesis to indicate the start and the end of this sequence; 'TS' is the DNA replication foci-targeting sequence that was described by (1); 'Zn' is the zinc/DNA binding motif (2); PBHD is the Polybromo-1 protein homologous domain; 'PC' is the catalytic site; 'GK' is the linker between the N- and C-terminal domains of DNA MTase and 'EGFP' refers to the enhanced green fluorescent protein sequence used as a marker in the fusion proteins. The black box stands for the intrinsic nuclear localization signal (aa 72–92) and the green box stands for the NLS of SV40 large T antigen. The numbers written against the wild-type protein sequence refer to amino acid positions. The old numbering of the positions of the amino acids is still used here in order to make comparison with the previous report (2). The numbers written on the right-hand side of the figure indicate the construct names for each map and the capital letters in parenthesis correspond to the panels in Figure 2. Constructs 1–8 were based on the tetracycline-regulated expression system in which the promoter (P_{tet}) is controlled by the availability of tetracycline (14). Constructs 9–11 were based on the EGFP fusion vectors (Clontech) using the constitutive, human cytomegalovirus immediate early gene (hCMV) promoter.

stream of the tetracycline-regulated promoter (P_{tet}). The constructs for transient transfection used the constitutive hCMV promoter rather than the tetracycline-regulated promoter.

Preparation of antiserum against the C-terminal domain of the mouse DNA MTase

The whole C-terminal domain coding-region of the mouse DNA MTase was excised as an *Xho*I–*Spe*I fragment, blunted by Klenow and inserted into the blunted *Bam*HI site of the vector pQE31 (Qiagen) downstream of the 6× His-tag coding sequence. This construct was expressed in *Escherichia coli* and the expressed protein was purified using established protocols (Qiagen). This recombinant protein was used to immunize rabbits and the resultant serum was $(NH_4)_2SO_4$-precipitated as described by Harlow and Lane (16) to raise the polyclonal antibody against the C-terminal domain of the mouse DNA MTase (anti-MTC).

Cell culture and stable transfection

Dr A. Baumeister (FMI, Basel) provided the C3H10T½ cell line. A subclone called C3.3, which was stable after 15 passages, was used for the transfections in this study. Transfections were performed when the cells reached 40–50% confluency. Transient transfections were performed using SuperFect™ (Qiagen) according to the manufacturer's instructions. Stable transfection of C3.3 cells was performed using the calcium phosphate procedure (17). The C3.3 cells were co-transfected with the transactivator expression plasmid pTet-tTAk and the neomycin-resistance plasmid pMC1NeoPolyA (Strategene). The cells were selected using 600 µg/ml of G418 sulfate for 12 days before the clones were picked up. The cloned cells were selected at the same concentration of G418 a week before being tested for transgene expression. Clone TA5, which stably expressed the transactivator, was used for the second round of co-transfection using the mutant expression vectors and the plasmid pX343 carrying hygromycin-resistance gene. Stable transfectants were selected using 150 µg/ml hygromycin B.

Indirect immunofluorescence

The cells were cultured on 18 × 18 mm microscopy coverslips (Menzel-Gläser, Germany). Cells were processed 16 h after induction of expression (stable clones with tetracycline controlled constructs) or 16 h post-transfection (transiently transfected cells with constitutive expression constructs). Prior to fixation they were washed twice in PBS. Fixation was performed by incubating the cells on ice in 4% fresh paraformaldehyde solution for 30 min. The paraformaldehyde was removed by two PBS washes and the cells were then permeablized with 1% Triton-in PBS at room temperature for 10 min. After two washes with PBS, the cells were incubated with blocking solution (10% goat serum in PBS) at room temperature for 30 min followed by incubation with the fractionated anti-C terminal antibody (1:25 dilution) for 1 h at room temperature. After three washes with PBS, the cells were incubated with the secondary antibody using the concentration suggested by the manufacturer. The cells were then washed three times with PBS and mounted on glass slides in Vectorshield (Vector laboratories, Inc., USA) mounting medium.

Conventional and confocal microscopy

Laser confocal microscopic scanning was performed using a Leica confocal microscope and the SCANWare sofware (Leica). Conventional microscopic observations for the co-localization of EGFP fusion proteins and DAPI staining were performed using a conventional Leica fluorescence microscope.

DAPI staining

Fixed cells were permeablized with 1% Triton X100 in PBS at room temperature for 5 min and then incubated with 1 µg/ml DAPI in PBS at room temperature (in the dark) for 10 min and then washed three times with PBS.

Primed *in situ* extension (PRINS)

Cells were cultured on collagen-coated 18 × 18 mm coverslips as described above. Primed *in situ* synthesis was performed according to the protocols described in the Boehringer Mannheim *in situ* hybridization manual (2nd Edition) using a Perkin Elmer PCR machine. The chromosomes were labeled by the incorporation of TRITC-dUTP during the extension of either centromeric- or telomeric-specific primers by *Taq* polymerase. The sequences of the primers were: centromeric-specific (5′-ATTTAGAAATGTC-CACTGTAGGAC-3′) (18) and telomeric-specific (5′-TTAGG-GTTAGGGTTAGGG-3′) (19). Indirect immunofluorescence was performed after the PRINS reaction using an anti-GFP monoclonal antibody (Clontech) and an FITC-conjugated secondary antibody (goat anti-mouse, Sigma) to bring up the intensity of the EGFP as the heating step in the PRINS reaction seriously damaged the EGFP fluorophore.

Genomic DNA extraction and *Sss*I methyl-group accepting assay

Cells that were stably transfected with the tetracycline-regulated constructs (1,4,5,7,8) were passed three times with an overall induction time of 72 h. Cells were scraped off the culture dishes using a rubber policeman and their genomic DNA was extracted according to Sambrook *et al.* (20). To assess the dominant negative effect of the DNA MTase–EGFP fusion proteins the genome wide DNA methylation level was measured using a modified *Sss*I methyl-group accepting assay (15). Briefly each reaction contained 1 µg of genomic DNA digested to completion by *Eco*RI, 6 U of *Sss*I methylase, 2 µM of [^{3}H]S-adenosyl-methionine (SAM) and 18 µM of cold SAM. After 2 h incubation at 37°C, the reactions were stopped by adding 500 µl 20% TCA using 100 µg BSA as carrier. The samples were vortexed for 20 s and incubated on ice for 30 min, the precipitates were spun down, washed three times with 20% TCA and hydrolyzed overnight with 100 µl of formic acid. The samples were then counted for c.p.m. using a scintillation counter. Every genomic DNA sample was tested in triplicate.

RESULTS

Subcellular localization of EGFP fusion proteins containing different parts of the N-terminal domain of the DNA methyltransferase

Stable transfection of C3.3 cells with the full DNA MTase N-terminal domain, under the control of a constitutive promoter, was not successful and may be due to the toxicity of this constitutive expression of the truncated DNA MTase (7). In order

to obtain a high level of stable expression and to reduce any possible toxic side effects, a modified tetracycline-regulated expression system (14) was used. In this system, both the transgene and the transactivator (the tetracycline repressor-VP16 fusion protein) (21) expression were driven by the tetracycline-regulated promoter (P_{tet}) (14). The regulation of expression from this construct was tested by northern blot hybridization (20). The EGFP was used as a fusion partner to monitor the expression of the mutants. The screening of the clones expressing the fusion proteins was performed as described in the method section. Fluorescent clones were then cultured on collagen-coated glass cover slips. Expression was induced for 24 h by changing the culture medium to tetracycline free medium. Figure 2 shows that the mutant DNA MTase–EGFP fusion proteins, that included the DNA replication TS, were all localized to the nucleus and formed toroidal or spotty structures in S-phase cells similar to those reported by Leonhardt *et al*. (1) (Fig. 2C and E–G). Surprisingly, construct 5, in which two thirds of the reported replication foci targeting sequence (aa 294–506) had been deleted, also showed toroidal structures although they were much less abundant (Fig. 2D). This construct contained the reported zinc-binding motif (2) and a stretch of sequence homologous to the chicken Polybromo-1 protein (PBHD, aa574–846) (4,9). Construct 7 (aa 1–138), which did not contain any part of the TS, also showed the toroidal structures as well as a relatively high level of diffused localization (Fig. 2B). On the other hand, the nuclear-localized EGFP alone showed only a diffused nuclear distribution (Fig. 2A) and the wild-type EGFP showed strong fluorescence in both the cytoplasm and the nucleus (Fig. 2H). The above results indicated clearly that, whilst the reported DNA replication foci-targeting sequence was sufficient to target the EGFP fusion proteins to sites of DNA replication, other sequences within the protein might also able to perform this function.

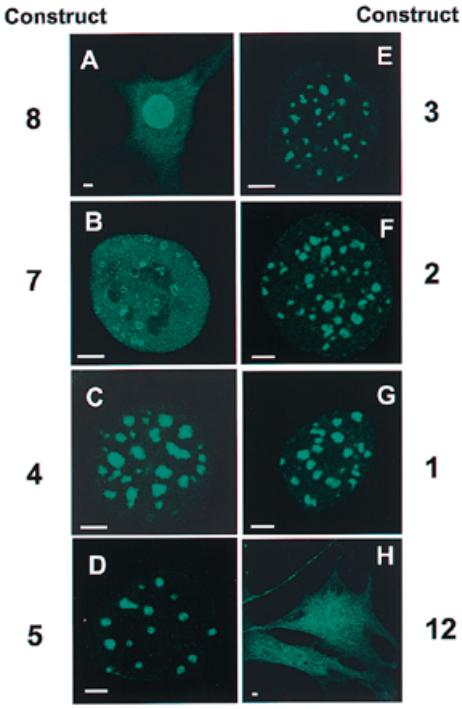


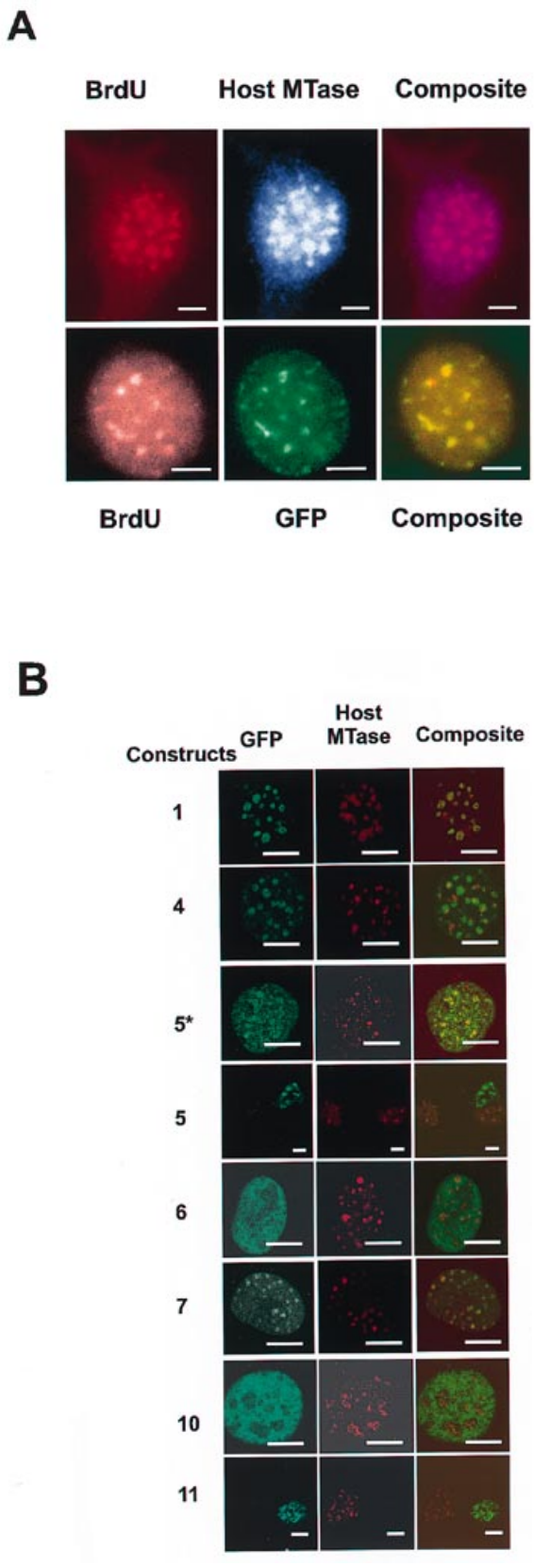
Figure 2. EGFP fluorescence from the MTase fusion constructs. Cells stably transfected with the constructs using the tetracycline-regulated expression vectors were cultured on glass coverslips and induced for expression for 24 h. Confocal scanning microscopy was performed as described in the Materials and Methods. The equatorial section for each construct is shown. (**A**) Construct 8; (**B**) construct 7; (**C**) construct 4; (**D**) construct 5; (**E**) construct 3; (**F**) construct 2; (**G**) construct 1 and (**H**) construct 12 (wild-type EGFP, controlled by P_{tet}). The scale bars are 5 μm in length.

Co-localization of the GFP fusion proteins with the host DNA MTase showed that the PBHD sequence could also target the enzyme to DNA replication foci

To test whether or not construct 5 was indeed targeted to sites of DNA replication, cells were pulse-labeled with BrdU after 16 h of induction of expression, and stained with an anti-BrdU antibody as described by Leonhardt *et al*. (1). Figure 3A (top row) shows that, in the wild-type cells, host DNA MTase always co-localized with the incorporated BrdU. Construct 5 co-localized with the incorporated BrdU (Fig. 3A, bottom row) but was observed in the same cell as the host DNA MTase at an extremely low frequency (Fig. 3B, construct 5*) and often showed mutual localization exclusivity with the host MTase (Fig. 3B, construct 5). While construct 6 showed a diffused nuclear localization, the host DNA MTase in the same cell showed toroidal or granular structures (Fig. 3B). This indicated that the sequence responsible for the DNA replication foci targeting in construct 5 was lying between aa 506 and 846, which contains the zinc binding motif and the PBHD. To further restrict the sequence required for this localization, three other constructs 9, 10 and 11 (Fig. 1) were made under the control of the constitutive hCMV promoter. Cells were transiently transfected and stained with anti-MTC. The construct 10 showed a diffused nuclear distribution (Fig. 3B) whereas construct 11 which contained the PBHD sequence sometimes showed co-localization with the host DNA MTase similar to construct 5 (Fig. 3B, construct 5*) and sometimes showed mutual exclusivity (Fig. 3B, construct 11), that is to say cells either had strong intensities of EGFP or host MTase staining but not both (Fig. 3B, constructs 5, 7, 11). The more EGFP expression observed the less host MTase was found. Construct 9 showed a similar behavior to construct 11 (data not shown). This indicated that the PBHD was responsible for the DNA replication foci targeting. Construct 1, which contained the full length of N-terminal domain of the DNA MTase containing both the TS and PBHD, showed in general a good co-localization with the host MTase but in some regions it showed more red (Host MTase) than green (EGFP) or *vice versa* (Fig. 3B, construct 1). Construct 4 in some cases showed co-localization with the host MTase and some times showed mutual exclusivity (Fig. 3B, construct 4). All of the above results indicate that these MTase–EGFP fusions could be, in principle, acting as competitors for the host protein on the DNA binding sites.

DNA MTase was found to be targeted to chromosomal regions other than just the centromeric regions

We determined the localization of the host MTase and the mutant MTase–GFP fusions (construct 1, 4, 5 and 9) by immuno-fluorescence and compared their distributions to those of the centromeric and telomeric sequences in interphase and S-phase cells. The localization of the centromeres and telomeres was made by PRINS reactions using specific primer sets. The specificity of the PRINS was tested in the following ways:



(i) single-strand DNA breaks could be blocked by a pre-treatment with ddNTPs without preventing the PRINS from working; (ii) the pattern of spots obtained was primer dependent [i.e. no spots were obtained from cells incubated without primers (data not shown)]; (iii) PRINS experiments, performed on condensed chromosomes, showed that the spots obtained with the primers were localized to the telomeric or centromeric chromosomal regions (Fig. 4A). The chromosomal localization of the DNA MTase appeared to be cell cycle dependent, in that ~10% of the cells failed to show co-localization with the centromeric regions (Fig. 4B, constructs 1*, 4*and 5*) whereas the other 90% did localize (Fig. 4B, constructs 1, 4, 5 and 9). We never observed telomeric localization of either the host MTase or the EGFP constructs (data not shown).

To determine the distribution of the EGFP-fusion protein along the condensed chromosomes, DAPI staining of these cells was performed. The DAPI-stained cells were observed by conventional fluorescence microscopy. Figure 4C shows composite images of green (EGFP) and blue (DAPI) mitotic cells (constructs 1 and 4). It was clear that the EGFP fusions containing the TS localized to the structures reminiscent of centromeres during mitosis. Given the apparent EGFP localization, the question about the location of the host MTase was addressed by staining untransfected cells with anti-MTC. DAPI staining was used to visualize the mitotic chromosomes. Mitotic cells showed a lower staining level of host MTase than their S-phase counterparts, however, it was clear that the host MTase did indeed localize to centromeric regions preferentially (data not shown).

The truncated DNA MTase–GFP fusion proteins can affect the action of the host DNA MTase in a dominant negative manner

The mutual exclusivity between the GFP fusion proteins and the host DNA MTase indicated that the mutants might exert a dominant negative effect based on a competition mechanism. The potential effect of the mutants was assessed by the changes in the genome-wide DNA methylation of clones stably expressing the mutants using the *Sss*I methyl-group accepting assay (15). The assay was designed to carry the reactions to completion. The more [^{3}H]SAM the DNA samples incorporate, the less methylated they are. The differences between the incorporation into the control DNA (isolated from the cells that only express the nuclear-localized EGFP; Fig. 1, construct 8) and the other constructs tested are summarized in Table 1. A negative value indicates that the DNA methylation level is higher than the control and a positive value

Figure 3. Co-localization of host DNA MTase with replication foci and the MTase–EGFP fusion protein. (**A**) Shows Br-dUTP incorporation into the replication foci, in red, of either untransfected (top row) or construct 5-expressing cells (bottom row). Host MTase staining is shown in blue (top row) and EGFP fluorescence, in green, from construct 5 (bottom row). The-right-hand column shows composite images of the left and middle columns. (**B**) Shows confocal images of the co-localization between the host MTase and the EGFP fusion constructs. Numbers on left-hand side of the panels stand for the names of the constructs. 5* shows that, under rare circumstances, construct 5 which contains mainly the polybromo domain PBHD and a small part of the targeting sequence TS, indeed co-localized with the host DNA MTase whereas in most cases it showed mutual exclusivity with the host MTase (Fig. 3B, construct 5). For constructs 5 and 11 there are two nuclei per panel whereas in all the other panels there is only one nucleus and as a consequence the magnification of these two panels is reduced. The scale bars are 5 μm in length.

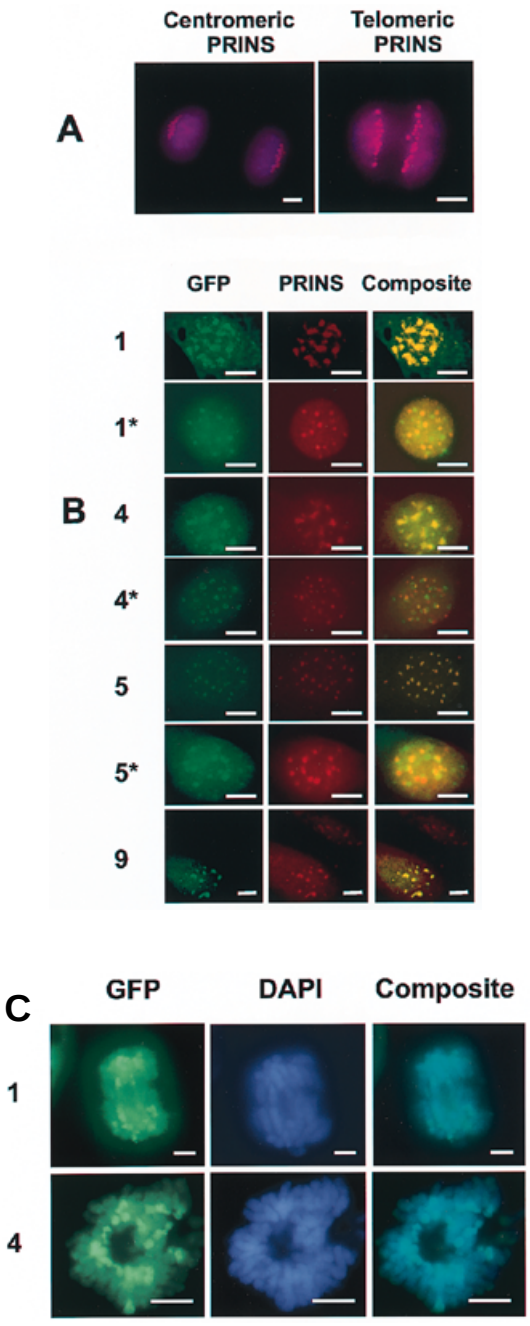


Figure 4. Primed *in situ* extension (PRINS) detection of centromeric and telomeric sequences. (**A**) Shows two control PRINS experiments using either the centromeric or telomeric primers to demonstrate their specificity. (**B**) Shows EGFP expression and centromere-specific PRINS. The numbers on the left are the construct names. Those marked with a '*' failed to co-localize with the centromeric PRINS (~10% of S-phase cells). In (B)9 there are two nuclei whereas all other panels have only one. (**C**) Shows EGFP (constructs 1 and 4, green) localized to the chromosomal structures reminiscent of centromeres. The right hand panels shows the composite images of the EGFP (green) and DAPI staining (blue). The scale bars are 5 µm in length.

indicates lower methylation. Construct 7 caused a non-significant ($P > 0.05$) decrease in the methyl group (CH_3) accepting capacity, whereas the other three constructs (constructs 1, 4 and 5) significantly increased the CH_3 accepting capacity, indicating a decrease in the level of genomic DNA methylation. The most extreme case was construct 4, which resulted in the largest degree of demethylation and also produced the highest and most stable levels of EGFP fusion protein expression (data not shown). The delta values were converted into percentage changes of genomic DNA methylation according to the formulae described in the legend for Table 1. These changes were 3.8 ± 0.3%, 5.7 ± 0.6%, 2.9 ± 1.7% and 1.0 ± 0.9% for the constructs 1, 2, 4 and 7, respectively (Table 1, right hand column).

DISCUSSION

Multiple domains are involved in the targeting of the mouse DNA MTase to the DNA replication foci

DNA MTase–EGFP fusion proteins, which contained the DNA replication foci TS of the mouse DNA MTase, behaved as suggested by Leonhardt *et al.* (1). However, construct 5 (aa 138–294; 507–846), which contained only the first 87 aa of the TS, and construct 7 (aa 1–138), which did not contain any part of the TS, also formed the toroidal structures typical for MTase staining of S-phase cells (Fig. 2B and D). Further analysis showed that construct 5 also co-localized with the incorporated BrdU (Fig. 3A, bottom row) and construct 7 showed co-localization with the host DNA MTase (Fig 3B, construct 7) while the host DNA MTase always co-localized with the BrdU incorporation (Fig. 3A, top row). This indicated that, while the TS may be sufficient for targeting, other sequences in the N-terminal domain are also involved in this targeting. However, this targeting was not always equal in efficiency. For example, the fusion proteins expressed from construct 7 showed both diffused distributions, reminiscent of the nuclear-localized EGFP control and toroidal structures similar to those observed from the constructs containing the published TS. Neither the nuclear-localized EGFP (construct 8) nor the wild-type EGFP (construct 12) controls ever showed granular structures (Fig. 2A and H). Construct 10 (aa 138–294) showed a similar diffused distribution as construct 8 while the host DNA MTase showed granular structures in the same cell. These data indicate that EGFP targeting to the replication foci may be performed by the PBHD and the sequence between aa 1–138, albeit less efficiently than the TS. The precise localization of the EGFP-constructs to replication foci was confirmed by their co-localization with BrdU (Fig. 3A and B). While these data show that there is more than one sequence involved in the DNA replication foci targeting they also show that the TS sequence is sufficient to fulfil this targeting on its own.

Given the lack of an obvious DNA-binding motif in construct 7, it is possible that it might be targeted to DNA replication sites via an indirect process such as an interaction with other proteins or nucleic acids that are themselves targeted to the replication foci. While as yet we are unable to support or reject this hypothesis, it should be noted that this protein region is highly charged and part of the peptide sequence shows high similarity to the 70 kDa U1 snRNP and to the Su (w^a) protein (1). The construct 5 contained a sequence from the N-terminal part of the known targeting sequence (aa 207–294, the full length of the TS is from aa 207–455) which includes the sequences corresponding to the reported DNA binding motifs DB1 and AZn from the human DNA MTases (12). However, construct 6 (aa 138–294), containing the homologous sequences of these motifs, did not

show any targeting (Fig. 3B) indicating that the DNA binding activity alone was not sufficient for the targeting. In this paper, we have restricted the sequence responsible for DNA replication targeting in construct 5 to the PBHD sequence.

Table 1. Dominant negative effect of the mutants tested using *Sss*I methyl-group accepting assay

Constructs	Δpmol CH_3/μg DNA
8 (reference)	0
1	1.7 ± 0.14
4	2.6 ± 0.28
5	1.3 ± 0.77
7	–0.45 ± 0.39

One microgram of genomic DNA of each clone was used for the *Sss*I methyl-group (CH_3) accepting assay. The incorporated c.p.m. were converted into pmol CH_3 accepted per μg DNA. The value of the control sample (Construct 8, the nuclear localized EGFP) was taken as the basal level (18.1 pmol/μg DNA) and the differences in the incorporation between other clones to the control were calculated. The left column shows the constructs tested. The middle column shows the differences in the incorporation of [^{3}H]CH_3 compared with the control (construct 8) ± S.E. The right hand column shows an estimate of the changes in genomic DNA methylation assuming that the control cells have 60% of their total CG dinucleotides methylated at C5 (34). The formulae used for this calculation were as follows: X / Z = 1 – 0.6 and C = (Y – X)/ Z. Where 'X' is the amount of [^{3}H]CH_3 incorporated into 1 μg of the control genomic DNA (18.1 pmol in this case). 'Z' is the total number of CG dinucleotides in 1 μg of genomic DNA (45.25 pmol in this case). 'Y' is the amount of [^{3}H]CH_3 incorporated into the listed constructs. 'C' is the percentage change in genomic DNA methylation.

The apparent exclusivity between the host MTase and the 'targeting sequence' containing EGFP fusion proteins indicated a competition between these two. Because of this competition, the relative intensity of the red (host MTase) and green (EGFP-fusion protein) was dependent on the time at which the transgene was induced. If it was induced late in S-phase then some of the sites were presumably already occupied by the host MTase prior to transgene expression. Thus, when the transgene was expressed it had much less opportunity to access these sites. However, if the transgene was expressed at a high level in G1 phase, where the host MTase is at its lowest level (11), the EGFP fusion protein would occupy more of the DNA replication foci as the cell cycle progressed towards S-phase. While this may explain the observed exclusivity between the host MTase and the EGFP fusion proteins, potential qualitative differences between these two could not be excluded. The EGFP fusion proteins that contained the PBHD and the zinc finger motif, in the absence of the TS, showed similar localization patterns to construct 5 which contained only a small part of the TS and the zinc finger motif, in addition to the PBHD (Fig. 3B, construct 5). They do not, however, always co-localize with the endogenous DNA MTase (Fig. 3B, constructs 5 and 11). In many cases the exclusivity between these PBHD-containing constructs and the host MTase was stronger than for those containing either the TS alone (Fig 3B, construct 4) or the TS and PBHD together (Fig 3B, construct 1). These differences could possibly be explained by the intrinsic property of PBHD to favor stronger protein–protein interactions and thus show a stronger sequestration of the replication machinery.

Chromatin association of the DNA MTase–GFP fusion proteins

It has been shown that DNA MTase can be associated with chromatin (22–24). An extensive study of the association of the MTase–EGFP fusion proteins to the chromatin was performed using the PRINS techniques with either centromeric or telomeric specific primers. Consistent with the previous report (1), it was found that ~90% of the centromeric PRINS co-localized with the EGFP fusions containing any of either the TS or PBHD (Fig 4B). However, it was also found that, in ~10% of the cells examined, the EGFP fusions formed granular structures that failed to co-localize with the centromeric PRINS (Fig. 4B, constructs 1*, 4* and 5*). The satellite DNA sequences located near the centromeric regions of the mouse chromosomes are heavily methylated at CpG sites (24) and are replicated at the very end of S-phase (25–29). After DNA replication, there would be a large number of hemimethylated sites, therefore, it is not surprising that a mechanism existing to recruit the DNA MTase to a high concentration to ensure the methylation pattern faithfully passed from mother cells to daughter cells. The other 10% of cases may have reflected genomic regions other than centromeric regions that also have high concentration of methylated CpG sites. Although subtelomeric satellite DNA has also been shown to be heavily methylated in mouse somatic cells (30), we never observed the EGFP fusion proteins co-localize with the telomeric PRINS (data not shown).

EGFP-fusion proteins were also found to localize to the centromeric regions in mitotic cells (Fig. 4C, constructs 4 and 8) while the nuclear-localized EGFP showed a diffused distribution at this stage (data not shown). The host DNA MTase showed lower staining in mitotic cells than in S-phase cells, but it was also preferentially localized to the centromeres. Although it has been reported that DNA methylation occurs several hours after replication (31), it was striking that both the EGFP fusion proteins and the host MTase were localized to heterochromatic regions during mitosis. The biological role of this heterochromatic localization of the DNA needs to be further investigated.

The toxicity of the mutants may be due to the sequestration of the DNA replication machinery rather than a change in genomic DNA methylation

Initial attempts to express mutants, containing the entire N-terminal domain of MTase, constitutively resulted in no stable-expressing clones (data not shown). This could be due to the toxic effects of these DNA MTase mutants as has been discussed by Tucker *et al.* (7). Subsequent experiments made use of the autoregulatory tetracycline-controlled expression system (14) to permit us to study these potentially toxic proteins.

Because the mutants were targeted to the DNA replication foci, it is possible that the high levels of their expression might titrate out several critical components of the DNA replication machinery or else upset the regulation of DNA replication through the direct binding of the mutants to genomic DNA. The work of Chuang *et al.* (12) revealed the presence of multiple DNA binding motifs in the N-terminal domain of the human DNA MTase. Tucker *et al.* (7) proposed that the toxicity of MTase overexpression might be attributed to the reported TS. Our construct 4, which contained the TS and a very short sequence (aa 138–206) upstream of it, was able to stably overexpress this sequence at a very high level and could be passed for prolonged periods (2 weeks) without obvious

toxic effects whereas any clone containing the PBHD would be lost in less than a week of passing with a significant loss of fluorescent cells 24–48 h post-induction (Liu, unpublished observations). Thus we hypothesize that this apparent toxic effect is most probably due to PBHD sequence and that its over-expression sequesters critical components of the DNA replication machinery through protein-protein interaction resulting in either the loss of the construct or cell death. This idea of protein–protein interactions is consistent with the observations that DNA MTase can be tightly bound to the nuclear matrix (32), possibly for enzyme storage, that it can be associated with chromatin (22–24) and that it can also interact with histone H1 strongly (Suetake *et al*., Poster, *FASEB*, Biological methylation, 1997).

The constructs containing the TS and PBHD could serve as dominant negative competitors for the host DNA MTase and cause hypomethylation of the genome (Table 1). Construct 4 which expressed the highest level of the fusion protein, showed the highest effect on the DNA methylation level, but was less toxic than the others (data not shown) indicating the toxic effect of the mutants might not be due to the changes in the DNA methylation level. This was supported by the fact that embryonic stem cells that had only 1/3 of the genomic DNA methylation level of the wild-type cells could grow normally (33).

ACKNOWLEDGEMENTS

We are grateful to Dr D. Pollays of Life Technologies Inc., USA for providing the tetracycline-regulated expression system and Prof. T. Bestor for his original DNA methyltransferase construct. We would like to thank Dr J. Hagmann for assisting confocal scanning microscopy, Drs B. Ludin and S. Kaech for helping with the conventional fluorescence microscopy. We would also like to thank Dr M. Sinnreich for providing the GFP antibody. We greatly appreciate Drs J. Paskowsky and S. Schwarz for critical reading of this manuscript.

REFERENCES

1 Leonhardt, H., Page, A. W., Weier, H. U. and Bestor, T. H. (1992) *Cell*, **71**, 865–873.
2 Bestor, T. H. (1992) *EMBO J.*, **11**, 2611–2617.
3 Adams, R. L. P. (1995) *BioEssays*, **17**, 139–145.
4 Bestor, T. H. (1996) In Riggs, A. D. Matienssen, R. A. and Russo, V. E. A. (ed.), *Epigenetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 61–76.
5 Jost, J. P. and Saluz, H. P. (1993) *DNA Methylation: Molecular Biology and Biological Significance*. Birkhäuser Verlag, Basel, Boston, Berlin, pp.572.
6 Yoder, J. A., Yen, R. W. C., Vertino, P. M., Bestor, T. H. and Baylin, S. B. (1996) *J. Biol. Chem.*, **271**, 31092–31097.
7 Tucker, K. L., Talbot, D., Lee, M. A., Leonhardt, H. and Jaenisch, R. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 12920–12925.
8 Glickman, J. F., Pavlovich, J. G. and Reich, N. O. (1997) *J. Biol. Chem.*, **272**, 17851–17857.
9 Nicolas, R. H. and Goodwin, G. H. (1996) *Gene*, **175**, 233–240.
10 Adams, R. L. P. and Burdon, R. H. (eds) (1985) *Molecular Biology of DNA Methylation*. Springer Verlag, New York.
11 Vogel, M. C., Papadopoulos, T., Muller-Hermelink, H. K., Drahovsky, D. and Pfeifer, G. P. (1988) *FEBS Lett.*, **236**, 9–13.
12 Chuang, L. S., Ng, H. H., Chia, J. N. and Li, B. F. (1996) *J. Mol. Biol.*, **257**, 935–948.
13 Szyf, M. (1996) *Pharmacol. Ther.*, 70, 1–37.
14 Shockett, P., Difilippantonio, M., Hellman, N. and Schatz, D. G. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 6522–6526.
15 Schmitt, F., Oakeley, E. J. and Jost, J. P. (1997) *J. Biol. Chem.*, **272**, 1534–1540.
16 Harlow, E. and Lane, D. (1988) *Antibodies, A Laboratory Guide*. Cold Spring Harbor Laboratory University Press, Cold Spring Harbor, NY.
17 Chen, C. and Okayama, H. (1987) *Mol. Cell. Biol.*, **7**, 2745–2752.
18 Vissel, B. and Choo, K. H. (1989) *Genomics*, **5**, 407–414.
19 Kipling, D., Wilson, H. E., Thomson, E. J. and Cooke, H. J. (1995) *Hum. Mol. Genet.*, **4**, 1007–1014.
20 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
21 Gossen, M. and Bujard, H. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 5547–5551.
22 Creusot, F., Christman, J. K. (1981) *Nucleic Acids Res* **9**, 5359–5381.
23 Caiafa, P., Mastrantonio, S., Cacace, F., Attina, M., Rispoli, M. and Strom, R. (1988) *Biochim. Biophys. Acta*, **951**, 191–200.
24 Caiafa, P., Mastrantonio, S., Attina, M., Rispoli, M., Reale, A. and Strom, R. (1988) *Biochem. Int.*, **17**, 863–875.
25 Selig, S., Ariel, M., Goitein, R., Marcus, M. and Cedar, H. (1988) *EMBO J.*, **7**, 419–426.
26 Church, K. (1965) *Genetics*, **52**, 843–849.
27 Dev, V. G., Grewal, M. S., Miller, D. A., Kouri, R. E., Hutton, J. J. and Miller, O. J. (1971) *Cytogenetics*, **10**, 436–451.
28 Hsu, T. C. and Markvong, A. (1975) *Chromosoma*, **51**, 311–322.
29 Miller, O. J. (1976) *Chromosoma*, **55**, 165–170.
30 de Lange, T., Shiue, L., Myers, R. M., Cox, D. R., Naylor, S. L., Killery, A. M. and Varmus, H. E. (1990) *Mol. Cell. Biol.*, **10**, 518–527.
31 Woodcock, D. M., Adams, J. K. and Cooper, I. A. (1982) *Biochim. Biophys. Acta*, **696**, 15–22.
32 Hubscher, U., Pedrali-Noy, G., Knust-Kron, B., Doerfler, W., Spadari, S. (1985) *Anal. Biochem.*, **150**, 442–448.
33 Li, E., Bestor, T. H. and Jaenisch, R. (1992) *Cell*, **69**, 915–926.
34 Leonhardt, H. and Bestor, T.H. (1993) In Jost and Saluz (ed.) *DNA Methylation: Molecular Biology and Biological Significance*. Birkhäuser Verlag, Basel, Boston, Berlin, pp. 109–119.