Transgene-induced CCWGG methylation does not alter CG methylation patterning in human kidney cells

Taras Shevchuk[1,2], Leo Kretzner[1], Kristofer Munson[1], John Axume[3], Jarrod Clark[1], Olga V. Dyachenko[2], Marie Caudill[3], Yaroslav Buryanov[2] and Steven S. Smith[1,*]

[1]City of Hope National Medical Center and Beckman Research Institute, 1500 E. Duarte Road, Duarte, CA 91010, USA, [2]Laboratory of Plant Biotechnology, Branch of Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia and [3]Department of Human Nutrition and Food Science, College of Agriculture, California State Polytechnic University, 3801 West Temple Avenue, Pomona, CA 91768, USA

ABSTRACT

Several reports suggest that C^{m}CWGG methylation tends not to co-exist with ^{m}CG methylation in human cells. We have asked whether or not methylation at CCWGG sites can influence CG methylation. DNA from cells expressing an M.EcoRII–GFP fusion was actively methylated at CCWGG sites. CG methylation as measured by R.HpaII/R.MspI ratios was unchanged in cells expressing the transgene. Cloned representatives of C^{m}CWGG methylated DNA often contained, or were adjacent to an ALU repeat, suggesting that M.EcoRII-GFP actively methylated gene-rich R-band DNA. The transgenic methyltransferase applied C^{m}CWGG methylation to a representative human promoter that was heavily methylated at CG dinucleotides (the SERPINB5 promoter) and to a representative promoter that was essentially unmethylated at CG dinucleotides (the APC promoter). In each case, the CG methylation pattern remained in its original state, unchanged by the presence of neighboring C^{m}CWGG sites. Q-PCR measurements showed that RNA expression from the APC gene was not significantly altered by the presence of C^{m}CWGG in its promoter. Kinetic studies suggested that an adjacent C^{m}CWGG methylation site influences neither the maintenance nor the *de novo* methylation activities of purified human Dnmt1. We conclude that C^{m}CWGG methylation does not exert a significant effect on CG methylation in human kidney cells.

INTRODUCTION

The signals that set up and maintain somatic methylation patterns are poorly understood. Most of the observed phenomena are consistent with the enzymology of the DNA methyltransferases. Somatic inheritance of CG methylation patterns is consistent with the capacity of a hemimethylated CG dyad to focus Dnmt1 on the unmethylated cytosine residue in the dyad and to stimulate its activity (1,2). The spreading of CG methylation in a region is consistent with the formation and subsequent resolution of hairpin loops (3–6). Moreover, the detection of long-lived hemimethylated sites in DNA (7–10) is consistent with the formation of hairpin loops or recombination intermediates in DNA (2,3,11–13), as are the targeting of Dnmt1 to chromatin during the G_2 phase of the cell cycle (14) and the observation of active DNA methylation in G_2 (15). Activation of *de novo* methyltransferases by residual CG methylation provides a possible explanation for the return of CG methylation patterns after the removal of DNA methyltransferase inhibitors, such as 5-azacytidine (16,17); however, this possibility has not been supported by enzymological evidence. Alternatively, unusual structure formation (3,4,11,12) or methyltransferase targeting to transcription state-specific chromatin proteins (14,18,19) offer possible explanations for this return.

Each of the phenomena above is primarily concerned with the nature of the methylation process at the CG dinucleotide. However, a significant body of evidence suggests that non-CG methylation (cytosine methylation outside the CG dinucleotide pair) is also present in human cells (10,20). The role of this possible epigenetic mark is currently unknown [for reviews see Clark *et al*. (21) and Lorincz and Groudine (22)]. Although its origin is controversial, it is apparently applied by hDnmt3a, hDnmt3b, hDnmt2 or by the non-specific action of

other human DNA methyltransferases (23–28). Methylation at the CCWGG (where W is defined as A or T) sequence is easily detected in certain genes in peripheral blood lymphocytes and its level appears to stand in reciprocal relation to the CG methylation level when neighboring sequences are examined (29,30). Moreover, it is possible that epigenetic marks at CCWGG sites might be functional as appears to be the case in plants (31,32). Of particular interest are reports suggesting that CCWGG methylation tends not to co-exist with CG methylation in human cells in the same region of DNA (29,30,33,34). In this report we have studied the possibility that the presence of CCWGG methylation has a direct effect on CG methylation. This was accomplished by introducing the bacterial DNA (Cytosine-5) methyltransferase M.EcoRII into the HK293 human kidney cell line, and determining CG methylation state in promoter regions that are methylated at CCWGG sites by the bacterial enzyme. Our findings *in vivo* are consistent with direct analyses of the effects of CCWGG methylation containing oligodeoxynucleotides on Dnmt1 methylation kinetics *in vitro*.

MATERIALS AND METHODS

Cloning and expression of the NLS–M.EcoRII–GFP fusion

Plasmid constructs. The EcoRII methyltransferase gene was originally obtained as a 1.8 kb BglII/BamHI insert in the pRT103 vector with the cauliflower mosaic virus 35S promoter upstream of M.EcoRII to drive its expression in plants (31). Synthetic double-strand DNA encoding a translational start ATG and viral (SV40) nuclear localization signal (NLS) was then ligated upstream of and in frame with M.EcoRII. The properties of the expressed M.EcoRII were normal as determined by the methods described previously (35).

For use in human cells, the following series of subclonings were performed. The 35S promoter, M.EcoRII gene (with 5'ATG and NLS), and downstream polyadenylation signal were cut out of pRT103 as an ~3 kb HindIII fragment and inserted into the HindIII site of pEGFP-N3, a mammalian expression vector with the CMV promoter 5' to the gene encoding green fluorescent protein (GFP). This construct was designated p2. It left the plant viral 35S promoter between the CMV promoter and M.EcoRII. An XhoI digest of p2 cut sites immediately 5' and 3' to the 35S promoter, released it and allowed religation of the vector to produce a CMV-driven M.EcoRII gene. This construct was designated p3. It still contained the M.EcoRII stop codon as well as a poly(A) signal between M.EcoRII and GFP. A new HindIII site was created by mutagenesis at the 3' end of M.EcoRII, just before the stop codon. Digestion with R.HindIII removed the M.EcoRII stop codon and poly(A) sequence, and religation of the vector fused M.EcoRII in frame with GFP. This construct was designated p4 or pEcoGFP.

Since HK293 cells are insensitive to selection by G418 and p4 contained the neomycin resistance gene, one more step was taken to link the GFP and M.EcoRII-GFP genes to puromycin resistance. The GFP gene was removed from pGFP-N3 with R.EcoRI and R.NotI and M.EcoRII–GFP was similarly cut out of p4 with R.NheI and R.NotI. These were inserted into the same sites within the polylinker of pIRESpuro3 (BD Biosciences, Palo Alto, CA). This is a CMV-driven bicistronic vector containing an internal ribosome entry site (IRES) downstream of the polylinker, allowing co-transcription of GFP or M.EcoRII–GFP together with the puromycin resistance gene. The same immediate early CMV promoter used in pEGFP-N3 drives expression of the GFP gene and the M.EcoRII–GFP fusion in the pIRESpuro3 subclones. This last pair of constructs was designated pI-GFP and pI-EcoGFP, and used in the experiments described here.

DNA manipulation, gel analysis, cloning and sequencing

DNA isolation. DNA was isolated from HK293 cells essentially as described in Sambrook *et al.* (36) (Chapter 6, Protocol 6), except that an additional chloroform extraction was carried out after the phenol chloroform extraction used in that method. DNA isolated in this fashion was ~12 kb in molecular length as judged by gel electrophoresis.

Digestion with methylation-sensitive isoschizomers. One microgram lots of DNA at 0.1 μg/μl final concentration were digested to completion with either 1 U/μl R.EcoRII or 1 U/μl R.BstNI. The R.EcoRII digestion was carried out in 1× SuRE/Cut Buffer H (Roche Diagnostics GmbH, Mannheim, Germany) at 37°C for 2 h. The R.BstNI digestion was carried out in 1× NEB buffer #2 containing 0.1 μg/ml NEB BSA (New England Biolabs, Beverly, MA) at 65°C for 2 h. For R.HpaII and R.MspI, digestion completeness was judged by co-digestion of ØX174RFI DNA (37). For R.EcoRII and R.BstNI, digestion completeness was assessed by co-digestion of pBR322 DNA isolated from an *Escherichia coli dcm⁻* host. R.EcoRII was inactivated by heating at 65°C for 20 min.

Gel electrophoresis, analysis, cloning and sequencing

Before gel analysis, DNA digests were phenol extracted and ethanol precipitated using the method described in Sambrook *et al.* (36) (Appendix A8.9–A8.15) in order to remove enzymes and BSA present during digestion. DNA was separated by agarose gel electrophoresis using the methods described previously (38).

Number average molecular weights (M_n) were estimated from scans of ethidium bromide stained gels (Figure 1). Using the weight fraction distribution developed in (39) it can be shown that the number average molecular weight occurs at the point in the weight distribution corresponding to 26% of the total area of the scan of the lane measured from the bottom of the gel. The relationship used was obtained as follows:

$$f_N(L) = f(1-f)^L f,$$

where L is the fragment length in bp and f is the cleavage frequency in bp^{-1}.

Note that

$$(1-f)^L \cong e^{-fL}$$

$$f_N(L) = f^2 e^{-fL} \quad \text{Number average probability}$$

$$f_W(L) = Lf^2 e^{-fL} \quad \text{Weight average probability}$$

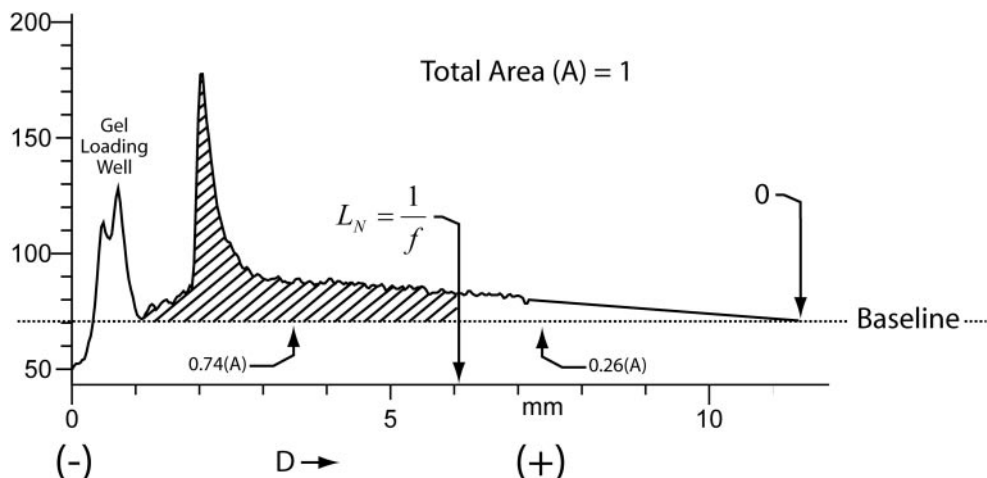$$F_W(L) = \int_{L_1}^{L_2} Lf^2 e^{-fL}\, dL \quad \text{for } L_1 = 0 \text{ and } L_2 = L.$$

Figure 1. Schematic of a restriction digest scan. The scan illustrates the method used in obtaining the number average molecular weight from an ethidium bromide stained gel displaying a weight average distribution of fragment sizes.

$$F_W(L) = \int_0^L Lf^2 e^{-fL} \, dL.$$

$$F_W(L) = 1 - (1 + fL)e^{-fL}.$$

Thus, $F_W(L)$ is the weight fraction $\leqslant L$. If we set $L = L_N$ (the number average molecular weight of the digest) and note $L_N = 1/f$, then the area from $L = 0$ to $L = L_N$:

$$F_W(L_N) = 1 - (1 + fL_N)e^{-fL_N}$$

$$F_W(L_N) = 0.26.$$

Thus, an ethidum bromide stained gel can be photographed and scanned, with an appropriate marker set in an adjacent lane. The molecular weight of the DNA at the point dividing the scan into 0.26 and 0.74 fractions is the number average molecular weight of the digest. A total of six gels were obtained. Scans were collected in duplicate using ImageQuant™ image analysis software (Amersham Biosciences, Piscataway, NJ), and the point in the scan at which 26% of the area occurs was mapped to a molecular weight based on its position in a standard curve relating the position of standards of known molecular length to distance traveled in the gel.

Molecular cloning. For the cloning of M.EcoRII protected DNA, the region of gel containing the high-molecular weight DNA spared by R.EcoRII was excised. The agarose-embedded DNA was then digested with 10 U R.BstNI in NEB Buffer 2 and BSA for 4 h at 60°C. Then, 4 U of Klenow fragment was added for 2 h at 25°C to blunt the DNA. DNA was extracted from the gel using a gel extraction kit (Qiagen, Valencia, CA). Eluted DNA was ligated into pBluescript linearized by R.EcoRV digestion and phosphatase treated. Clones containing an insert were sequenced on an ABI sequencer (Applied Biosystems, Foster City, CA).

CG methylation status in CCWGG methylated promoters by MS-PCR

We used the method of Futscher *et al.* (40) to assess the methylation status of CG sites adjacent to CCWGG

methylated sites in two representative promoter sequences. Genomic DNA from HK293 cells expressing M.EcoRII–GFP or HK293 cells expressing GFP alone was isolated using Qiagen's QIAamp® DNA Blood Kit according to the manufacturer's instructions. Ten micrograms of the M.EcoRII–GFP DNA was digested with 100 U of R.EcoRII in SuRE/Cut Buffer H (Roche GmbH, Mannheim, Germany) for 2 h at 37°C. The digest was then phenol extracted, ethanol precipitated and resuspended in 10 mM Tris and 1 mM EDTA, pH 7.5. The final concentration was determined by spectrophotometry. Five micrograms of uncut GFP and 5 μg of cut EcoRII–GFP 293 genomic DNA was then bisulfite treated using the CpGenome DNA Modification Kit (Chemicon International, Temecula, CA), according to the manufacturer's instructions. An aliquot of 0.5 μg of this bisulfite treated DNA was then amplified as described previously (40) except that for GFP expressors the primer containing an R.EcoRII site matched the unmethylated sequence while that for the M.EcoRII–GFP expressors matched the methylated sequence. The same strategy was used in amplifying and cloning sequences from the promoter of the APC gene, except that the following primers were used: APC EcoRII forward 5′-CAACAAACCTGACCATAATAAC-3′, APC GFP forward 5′-CAACAAACCTAACCATAATAAC-3′, APC reverse 5′-TTAGGGTTAGGTAGGTTGTG-3′. PCR for APC used a single round of amplification utilizing a touchdown procedure of 95°C for 30 s/all cycles, 55°C for 30 s/5 cycles; 53°C for 30 s/5 cycles; 51°C for 30 s/20 cycles, 72°C for 30 s/all cycles. Two microliters of each PCR was used for cloning with a TOPO TA Cloning Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Sequencing of 10 clones for each PCR was carried out using an ABI automated sequencer.

RNA expression from a CCWGG methylated promoter by Q-PCR

RNA expression from the APC gene was determined by use of Quantitative RT–PCR (41). RNA was isolated from 10 week HK293 cells from cells expressing GFP and M.EcoRII–GFP, by use of an RNAqueous™ Kit, (Ambion, Austin, TX). Then

1 µl of isolated RNA was measured using the ND-1000 Nanodrop® spectrophotometer (Ambion), and cross-checked by using RNA microfluidics chips and the 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Q-PCR results were generated using the iScript one-step RT–PCR kit with SYBR green (Bio-Rad, Hercules, CA). A 50 ng to 3.125 ng five-point dilution curve was generated using a total of 19 µl of RNA added to 2.5 µl of 10 µM each FWD and REV primers, 25 µl iScript master mix and 1 µl iScript enzyme. hTERT expression was used as a control, since promoter methylation at CG sites is not obligatorily associated with gene silencing at this locus (42,43). The following primers were used: APC FWD primer: 5′-TTTGTCTTGGCGAGCAGA-3′; the APC REV Primer: 5′-CAAGCTGGACACATTCCG-3′; the hTERT FWD primer: 5′-AGAGTGTCTGGAGCAAGTTGC-3′; and the hTERT REV primer: 5′-CGTAGTCCATGTTCACAATCG-3′. The reaction was then performed in the Rotor-gene 3000® (Corbett Robotics, San Francisco, CA) with the following protocol: 50°C for 10 min, 95°C for 5 min, then 40 cycles of 95°C for 15 s, 58°C for 30 s, 72°C for 60 s, 78°C for 15 s, 85°C for 15 s, and then one hold at 72°C for 1 min, followed by a default melt curve from 72°C to 95°C. All PCR products were verified as the correct size by use of a 2% agarose gel and appropriate DNA molecular weight markers. The slopes of five independent Quantitative RT–PCR runs were used for APC and hTERT RNA expression. The cut point on the fluorescence axis at which the plot of cycle number versus log (fluorescence) achieves the best linear fit of the data (Figure 7) was determined by using the Corbett Research analysis software: Rotor-gene 6® (Corbett Robotics). Analysis of N_0 values [i.e. initial target RNA concentration determined directly from the log-linear region of each amplification curve using the software and the method described previously (44) cross-checked using the second derivative method (45)] gave similar results. Five sets of dilutions ranging from 50 ng to 3.125 ng input total RNA were measured, yielding a total of 25 data points for each cell line.

Cell culture

Human kidney 293 cells were grown at 37°C with humidity and 5% CO_2, in DMEM (Mediatech, Herdon VA) containing 10% fetal calf serum (FCS) and standard antibiotics (25 U/ml penicillin and 25 µg/ml streptomycin) (Invitrogen) in addition to 25 mg/ml erythromycin (Sigma, St Louis, MO). For normal maintenance, cells were passaged once per week at fairly low density (1:10–1:20) and re-fed every 2–3 days. In passaging, cells were trypsinized with 1/2× trypsin in phosphate-buffered saline (PBS)—i.e. a 1:1 mixture of standard trypsin/EDTA (500 U/ml trypsin and 0.5 mM EDTA) and 1× PBS. To minimize typical HK293 cell clumping, 1–2 ml trypsinized and triturated cells were added to ∼5 ml of fresh medium and spun at 1000 *g* for 5 min, whereupon the cell pellet was well resuspended in 0.5 ml of 1/2× trypsin for 40–50 s. An aliquot of 1.5 ml of medium was mixed with this and 1/10 to 1/20 of the volume was added to a new plate with fresh medium. To expand cell numbers, cells were passaged at 1:1–1:3.

Transfections

For transfections, 1×10^6 cells/10 cm dish were plated in 7.5 ml medium. After 20–24 h the cells were re-fed with 5 ml medium, and were transfected 1.5–2.5 h after that. Transfection was by standard $CaPO_4$ method with chloroquine, as described in (36), except for the following: 50 µg plasmid DNA was mixed with 0.5 ml each 2× $CaCl_2$ and 2× HEPES-buffered saline per plate, and transfections were allowed to go overnight (16–20 h). The day transfection began was called 'day 1' of an experiment. Plates were washed once the next day with warm 1× PBS, re-fed with 6 ml medium, and maintained as above until the time of harvesting and analysis. Transfection efficiencies were typically ∼60%. For stable transfections, puromycin (Sigma) was added at 0.5 µg/ml to the medium of the cells on days 3–5, depending on the day of cell sorting, and surviving cells were grown as a pool.

Cell sorting

For cell sorting, cells were processed as above (passaging), with the cell pellets resuspended at $5–10 \times 10^6$/ml in DMEM with 20% FCS and placed on ice. A high speed MoFlo MLS sorter by Dako Cytomation (Fort Collins, CO) was used to sterilely sort for GFP positive cells in the City of Hope Flow Cytometry core facility. Positive cells, fluorescent at 508 nm above background using 488 nm excitation, were sorted into DMEM and 20% FCS, after which they were spun down as above and either re-plated or frozen for future DNA extraction. Plating efficiency and survival of sorted cells was much better on poly-L-lysine treated plates. For this, poly-L-lysine hydrobromide (Sigma) was resuspended in water to 20 µg/ml, filter sterilized, and added to plates at 3 ml/10 cm, and these were incubated at room temperature while gently rocking for 15–20 min. Plates were then washed three times with ∼5 ml sterile water and kept wrapped until use.

End counting with the cytosine extension assay

The protocol used was a modification of the method published by Pogribny *et al*. (46,47). Genomic DNA (0.75 µg) was digested for 2 h with a 10-fold excess of methylation-sensitive R.HpaII (New England Biolabs) according to the manufacturer's protocol. A second DNA aliquot (0.75 µg) was digested simultaneously with methylation-insensitive iso-schizomer R.MspI (New England Biolabs). A third DNA aliquot (0.75 µg) was incubated without restriction enzyme and served as a background control. After incubation, the samples were placed on ice for a minimum of 5 min to stop the reaction. The single nucleotide extension reaction was performed in 40 µl of a mixture containing 0.75 µg genomic DNA, 1× NEB Buffer 2 (New England Biolabs), 10 U of Klenow fragment/exo⁻ (New England Biolabs) and 4 µM [³H]dCTP (Perkin Elmer, Boston, MA). After incubating at 15°C for 1 h, the mixture was placed on ice for a minimum of 5 min to stop the reaction. Thirty-five microliters of sample was then applied to DE-81 ion-exchange filter paper (Whatman, Middlesex, UK) and washed three times with sodium phosphate buffer (pH 7.0) at room temperature. The filters were air dried and processed for scintillation counting. Four replicates were generated per sample. The [³H]dCTP incorporation into DNA was expressed as mean disintegrations per minute (d.p.m.) per 0.75 µg of DNA. The absolute percent of double-stranded unmethylated CCGG sites was calculated as follows: (R.HpaII induced d.p.m/R.MspI induced d.p.m) (100%).
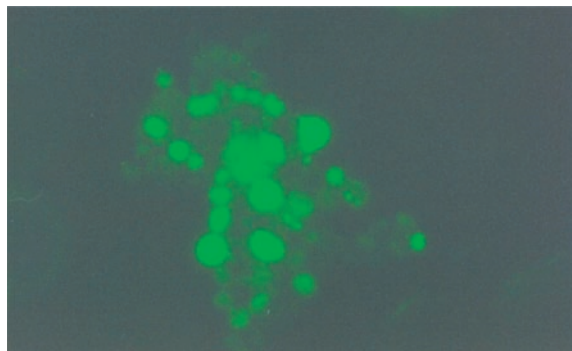
A
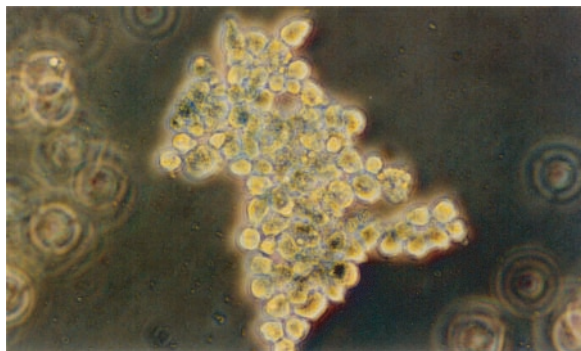
B

Figure 2. Fluorescence detection of GFP and the M.EcoRII–GFP fusion in HK293 cells. Each of the two constructs was transfected into HK293 and subjected to long-term growth under puromycin selection. GFP fluorescence was easily detected with fluorescence microscopy when either construct was used. A set of micrographs is depicted for expression of the M.EcoRII–GFP fusion: (**A**) light microscopy and (**B**) fluorescence microscopy.

In addition to running the undigested control, we included an aliquot of lambda DNA as a positive control suitable for high levels of [³H]dCTP incorporation, and an aliquot of lambda DNA that had been methylated *in vitro* with M.SssI (New England Biolabs) as a negative control suitable for low levels of [³H]dCTP incorporation.

Dnmt1 purification and kinetics

Methods of purification and kinetic analysis for human DNMT1 were those described in Clark *et al.* (48).

RESULTS

Construction of expression vectors

Two constructs have been used. One construct (pI-GFP) carries the GFP gene alone. The other (pI-EcoGFP) carries a bacterial methyltransferase–GFP fusion with a viral (SV40) NLS. In both constructs, expression is driven by a cytomegalovirus (CMV) promoter and selection for expression in human cells is based on the expression of a polycistronic message carrying a 3′ puromycin resistance coding sequence.

Expression of M.EcoRII

Each of the two constructs was transfected into HK293 followed by long-term growth under puromycin selection. GFP fluorescence was easily detected with fluorescence microscopy when either construct was used. A representative light microscope image is shown in Figure 2A and the corresponding fluorescence microscopy image is shown in Figure 2B. Cell sorting data for the GFP expression construct are shown in Figure 3A, and that for the M.EcoRII–GFP fusion is shown in Figure 3B. Cells exhibiting GFP fluorescence after transfection with vectors carrying GFP alone were consistently observed to be present in significantly larger numbers than those exhibiting fluorescence after transfection with the M.EcoRII–GFP fusion (Figure 3). This was true for cells expressing the two constructs for a short time with or without puromycin selection, or for a longer period under puromycin selection.
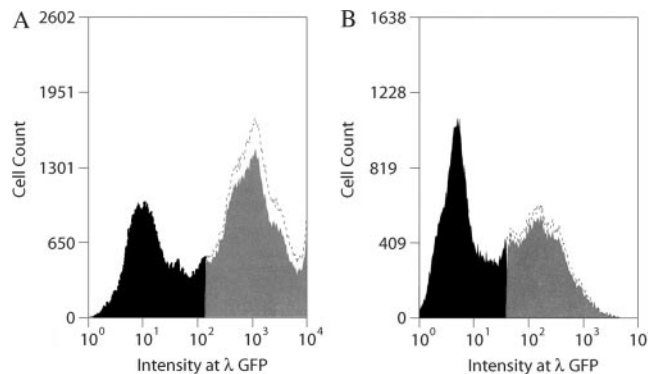

Figure 3. Florescence cell sorting quantification of cells expressing GFP and the M.EcoRII–GFP fusion in human HK293 cells. Each of the two constructs was transfected into HK293 and subjected to long-term growth under puromycin selection. Fluorescence cell sorting at the emission maximum for GFP quantifies the number of cells expressing each protein. (**A**) Quantification of cells expressing the GFP construct. (**B**) Quantification of cells expressing the M.EcoRII–GFP fusion. In each panel background fluorescence is given in black and cellular GFP fluorescence above background is given in grey. Total counts given by the stippled line indicate that not all positive events were collected in the sort.

The cells used in this study were selected as follows: separate lots of cells were transfected with equal amounts of the two vectors and allowed to recover for 3–5 days in non-selective medium. They were then selected for GFP fluorescence by cell sorting. When sorted fluorescent cells were re-plated in equal numbers and subjected to a round of selection by passage in medium containing 0.5 μg/ml puromycin for 16 days, the percentage of fluorescent cells expressing GFP alone was again greater than the percentage of fluorescent cells expressing the M.EcoRII–GFP fusion. Equal numbers of these puromycin-resistant, fluorescent cells were again plated in medium containing 0.5 μg/ml puromycin and passaged in this medium for 21 days. After this final 21-day selection, puromycin-resistant cells expressing fluorescence made up more than 80% of the population from the M.EcoRII–GFP fusion culture, while puromycin-resistant cells expressing fluorescence made up 100% of the GFP culture (Figure 4).
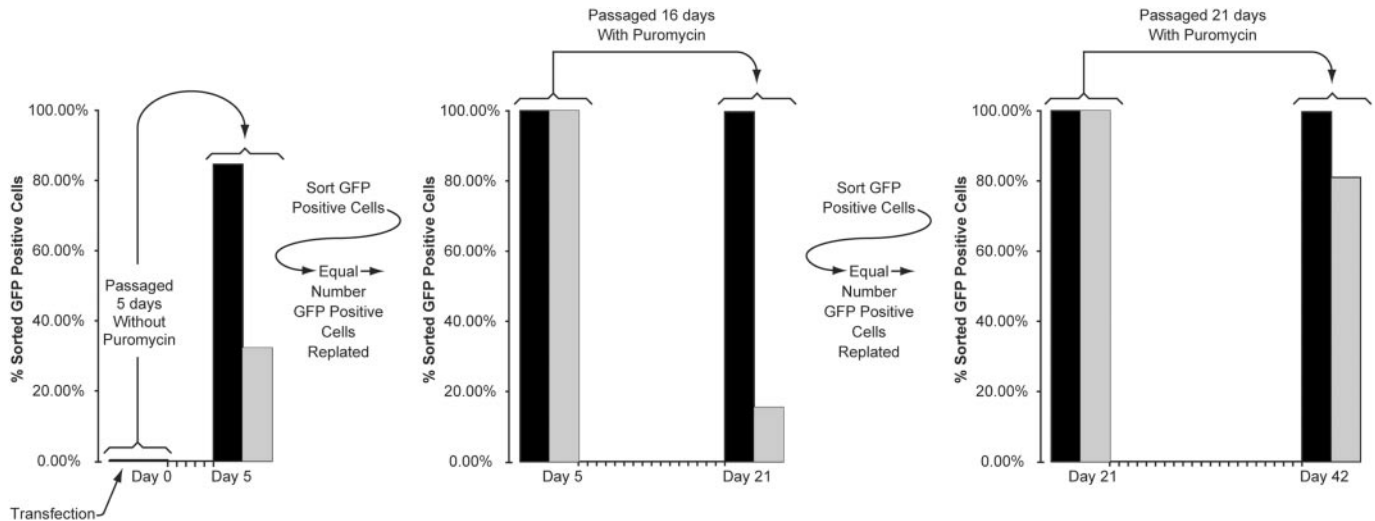
Figure 4. GFP expression during selection with puromycin. Cells plated on puromycin containing medium were sorted for GFP expression at the indicated times. The percentage of cells positive for GFP expression in pI-GFP transfected cells (GFP expressors) is given in black. The percentage of GFP expression in pI-EcoGFP transfected cells (M.EcoRII–GFP fusion expressors) is given in grey. After each round of sorting, equal numbers of GFP positive cells were re-plated for continued growth in the presence of puromycin.

De novo methylation of CCWGG sites

In general, the observation of fluorescence from a C-terminal GFP fusion protein indicates that the N-terminal region containing the protein of interest has folded properly because protein folding is initiated at the N-terminus. Thus, we expected that M.EcoRII–GFP fusion would be capable of methylating DNA at CCWGG sites because fluorescence was observed in the nucleus. To determine whether or not CCWGG sites were methylated in transfected cells, high-molecular weight chromosomal DNA was isolated from transfectants expressing the M.EcoRII–GFP fusion or GFP alone. This DNA was then digested with the isoschizomers R.EcoRII and R.BstNI. These enzymes both recognize the CCWGG site, but R.EcoRII cleavage is blocked by C^{m}CWGG methylation (C5-methylation at the internal cytosine) while R.BstNI cleavage is not. As shown in Figure 5, R.EcoRII cleavage is blocked only in cells expressing the M.EcoRII–GFP fusion, indicating that significant methylation at CCWGG sites in human chromosomal DNA had occurred only in those cells expressing the fusion.

The non-digested Eco-GFP DNA (Figure 5) had a number average molecular weight of 2.8 ± 0.1 kb. The R.BstNI digested Eco-GFP DNA had a number average molecular weight of 0.85 ± 0.01 kb and that of the GFP DNA was 0.79 ± 0.01 kb the R.EcoRII digest of the GFP DNA had a number average molecular weight of 1.3 ± 0.1 kb and that of the R.EcoRII digested Eco-GFP DNA was 2.8 ± 0.1 kb. For a genome of 3.0×10^9 bp this implies that about $1.53 \times 10^6/2.51 \times 10^6 = 60\%$ of the genomic CCWGG sites were methylated in HK293 cells expressing GFP, and $2.48 \times 10^6/2.44 \times 10^6$ or ~100% of the genomic CCWGG sites were methyated in the HK293 cells expressing the transgenic methyltransferase.

Representative sequences from the CCWGG methylated DNA

In order to study the protected DNA further, the band of genomic DNA remaining after treatment of Eco-GFP DNA with R.EcoRII was cut from agarose gels, purified and digested with R.BstNI. As a control, the same region of agarose gel from the GFP DNA treated with R.EcoRII, was treated in the same way although no band was visible. The R.BstNI-cleaved DNA was cloned as described in Materials and Methods, and multiple transformants were analyzed by sequencing. Characteristics of representative sequences cloned from the R.EcoRII protected DNA area are depicted in Table 1. Only two clones were obtained from the unprotected DNA from the GFP expressors, in agreement with the low level of R.EcoRII protected DNA shown in Figure 5. However, numerous clones were obtained from the R.EcoRII protected DNA obtained from the cells expressing the M.EcoRII–GFP fusion. Of the 32 clones sequenced 16 contained an ALU sequence or were adjacent to an ALU sequence. Of the total number of bases cloned and sequenced (~3900 bp) ~31% (~1200 bp) were from the various ALU families.

Global CG methylation levels GFP and M.EcoRII–GFP fusion expressors

We wished to determine if ectopic expression of the non-CG methyltransferase could affect the global level of CG methylation. To test this we determined the ratio of cleavage at the CCGG tetranucleotide using end-counting for R.HpaII and R.MspI digests of DNA from the transfectants expressing the M.EcoRII-GFP fusion or GFP alone. The mean dpm values are given in Table 2. The R.HpaII/R.MspI ratios were essentially identical, indicating that no significant change in CG methylation at the CCGG tetranucleotide had occurred in the transfectants expressing the M.EcoRII–GFP fusion.

CG methylation status in CCWGG methylated promoters

Since general levels of CG methylation measured with the end-counting method above do not necessarily reflect the level of CG methylation at promoter sequences, we inspected the CG methylation levels at two representative promoters.
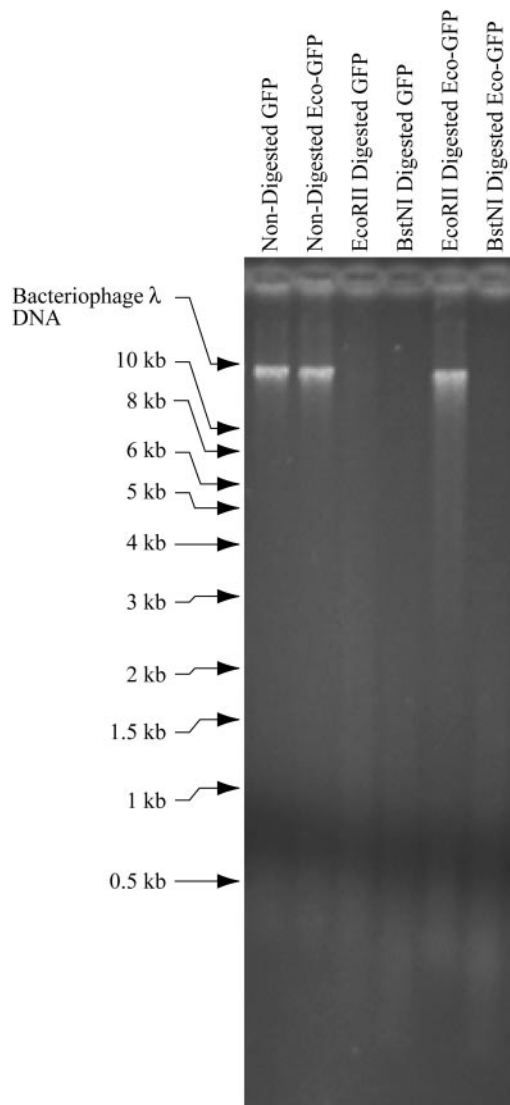
Figure 5. Detection of M.EcoRII activity from the M.EcoRII–GFP fusion expression vector. After 21 days of selection with 0.5 μM puromycin, fluorescent cells were harvested and genomic DNA was prepared. The DNA was digested with the isoschizomers R.EcoRII and R.BstNI. In this pair, R.EcoRII is unable to cleave the C^{m}CWGG site when the internal cytosine is methylated. While R.BstNI is unaffected by cytosine-5 methylation at this site.

One of them, SERPINB5, is methylated at CG sites while the other, the APC promoter, is unmethylated at CG sites in HK293 cells expressing GFP alone. The methylation status of CG sites adjacent to CCWGG methylated sites in the SERPINB5 and APC promoters are shown in Figures 6 and 7, respectively. In each case, the methylation state of CG sequences seen in the cell line expressing GFP alone was not altered by adjacent C^{m}CWGG methylation applied to the promoters by the M.EcoRII–GFP expressors.

RNA expression from a CCWGG methylated promoter

Since C^{m}CWGG has been suggested to be a negative regulator of gene expression in human cells (29,30), it was important to determine whether or not the expression of the APC gene was

Table 1. Representative clones from the EcoRII protected DNA fraction

	Observed in R.EcoRII protected fraction	Genomic expectation
Number ALU family of clones	16/32	–
ALU bp/total cloned bp	1200/3900	1, 179, 211/3, 079, 050, 944
ALU bp frequency	0.31	0.11

Table 2. General DNA level of CG methylation at CCGG sites

Cell line	R.HpaII ends ($n = 4$) (d.p.m.)	R.MspI ends ($n = 4$) (d.p.m.)	HpaII ends/ MspI ends
M.EcoRII–GFP fusion expressor	7829	32 508	0.241
GFP expressor	6114	25 272	0.242

downregulated by C^{m}CWGG methylation in its promoter sequence. The relative abundance of the APC target sequence RNA in total RNA was determined using Q-PCR as described in Materials and Methods. The data are shown in Figure 8. The expression level was not altered significantly in cells expressing M.EcoRII–GFP compared with cells expressing GFP. Expression from the hTERT gene was also unaltered by the presence of the active CCWGG methyltransferase. Thus, C^{m}CWGG methylation does not appear to be a negative regulator of these genes in HK293 cells.

Kinetic analysis of the effect of C^{m}CWGG methylation of Dnmt1

Using an oligodeoxynucleotide sequence from the region of the SERPINB5 promoter, spanning the EcoRII site [labeled: Internal EcoRII, in Figure 6B], we compared the saturation curves for the sequences containing no methylated sites, a single C^{m}CWGG site, a single ^{m}CG site and both a single C^{m}CWGG site, and a single ^{m}CG. As can be seen from the data in Figure 9 human Dnmt1 methylated asymmetrically methylated DNA most rapidly. Moreover, the rate of methylation induced by the presence of the asymmetrically methylated CG site was not significantly affected by the presence of the adjacent asymmetrically methylated CCWGG.

DISCUSSION

Beginning with reports on non-CG methylation in plants, there is now a significant body of evidence suggesting that non-CG methylation is present in most eukaryotes and thus may play a biological role in eukaryotes. For example, methylation levels in *Dictyostelium* (49) can be no more than 0.2% of cytosine, although CG deficiency in this A+T-rich genome suggests that some methylation is present (50). Adult *Drosophila* have no CG methylation (51); however, there are data supporting the presence of non-CG methylation in this organism (52–55). In human cells non-CG methylation may be the predominant form of methylation (20). The existence of an active alternative methyltransferase with non-CG specificity can be inferred from reports showing that the majority of 5mC in humans is on repeated sequences at non-CG sites (10,20), coupled with *in vitro* evidence suggesting that hDnmt1 is not capable of
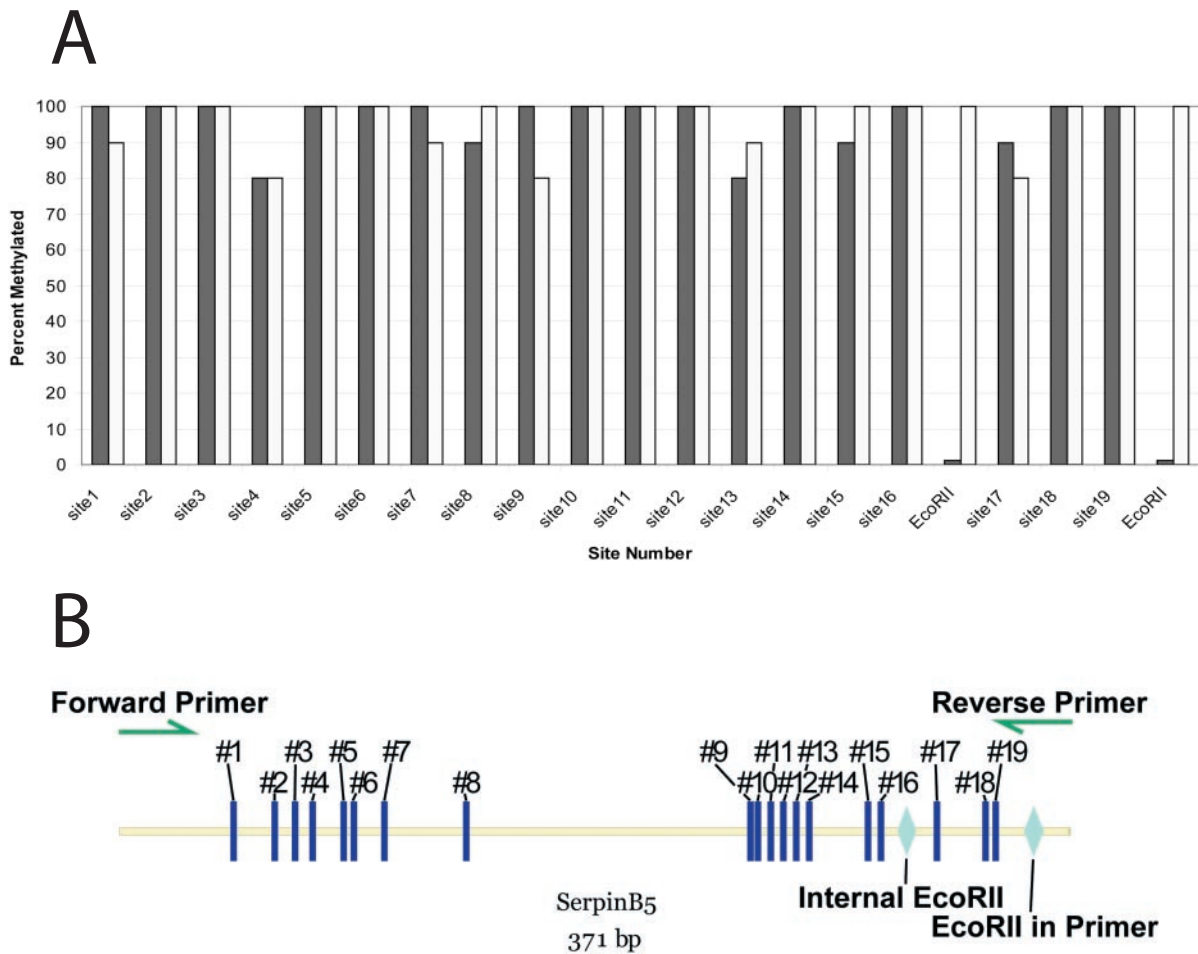
Figure 6. Effect of C^{m}CWGG methylation at a promoter that is heavily methylated at CG sites. Bisulfite treated DNA from the SERPINB5 promoter was amplified, cloned and sequenced as described in Materials and Methods. The 372 bp amplicon contains 19 CG dinucleotide pairs. It also contains two CCWGG sites, one in the reverse primer and one downstream from the reverse primer. Ten clones were sequenced in order to determine the methylation state at each of the 19 CG sites and the two CCWGG sites in each clone. (**A**) Bar graph depicting the observed fraction of methylated sequences at each site in the two types of transformed cells. (**B**) Those expressing GFP alone are shown in grey (left), and those expressing the M.EcoRII-GFP fusion are shown in white (right). Map of the amplicon depicting the relative position of each of the sites in the bar graph.

mediating the somatic inheritance of CWG methylation patterns (2). The specificity of target recognition of methyltransferases is generally preserved in the fusion proteins that have been studied (56–58); thus, it does not seem unreasonable to attempt to study sequence recognition in cloned mammalian methyltransferase fusion proteins of different types. The best evidence using these fusions suggests that CCWGG methylation observed in mammalian cells may be generated by the action of hDnmt3a, hDnmt3b, hDnmt2 or by the non-specific action of other human DNA methyltransferases (23–28).

Adenine-N6 methylation does not occur in higher eukaryotes. Gingeras and co-workers (59) have shown that the expression of an adenine-N6 methyltransferase in mammalian cells does not appear to alter the phenotype of mammalian cells. However, Baylin and co-workers (60) found that expressing the bacterial cytosine-5 methyltransferase M.HhaI in mouse 3T3 cells results in the introduction of CG methylation at the GCGC tetramer and the selection of significant phenotypic alterations.

Our experiments did not identify phenotypic changes associated with M.EcoRII expression in human embryonic kidney cells. Since the cells were sorted based on the level of fluorescence detected from GFP or from the M.EcoRII–GFP fusion, the difference we observe in the number of fluorescent cells could result from either a difference in transcription or translation efficiency that decreases the level of expression of the M.EcoRII–GFP fusion protein relative to the level of expression of GFP alone, or from some form of toxicity caused by the expression of M.EcoRII–GFP. We favor the possibility that translation from a subset of the M.EcoRII–GFP integrants is diminished (possibly by alterations near the two ATG's near its translation start site). This would explain the requirement for extensive cell sorting to homogenize the population for high levels of M.EcoRII–GFP expression.

Even so, the initial drop in the number of GFP positive cells expressing the M.EcoRII–GFP fusion can also be explained if expression levels are high in all integrants and cell sorting is required in order to select for cells that can tolerate the expression of the fusion. The apparent toxic effects of the expression might result from the capacity of the M.EcoRII–GFP fusion protein to methylate DNA at CCWGG sites. Alternatively, these effects might result from the capacity of the
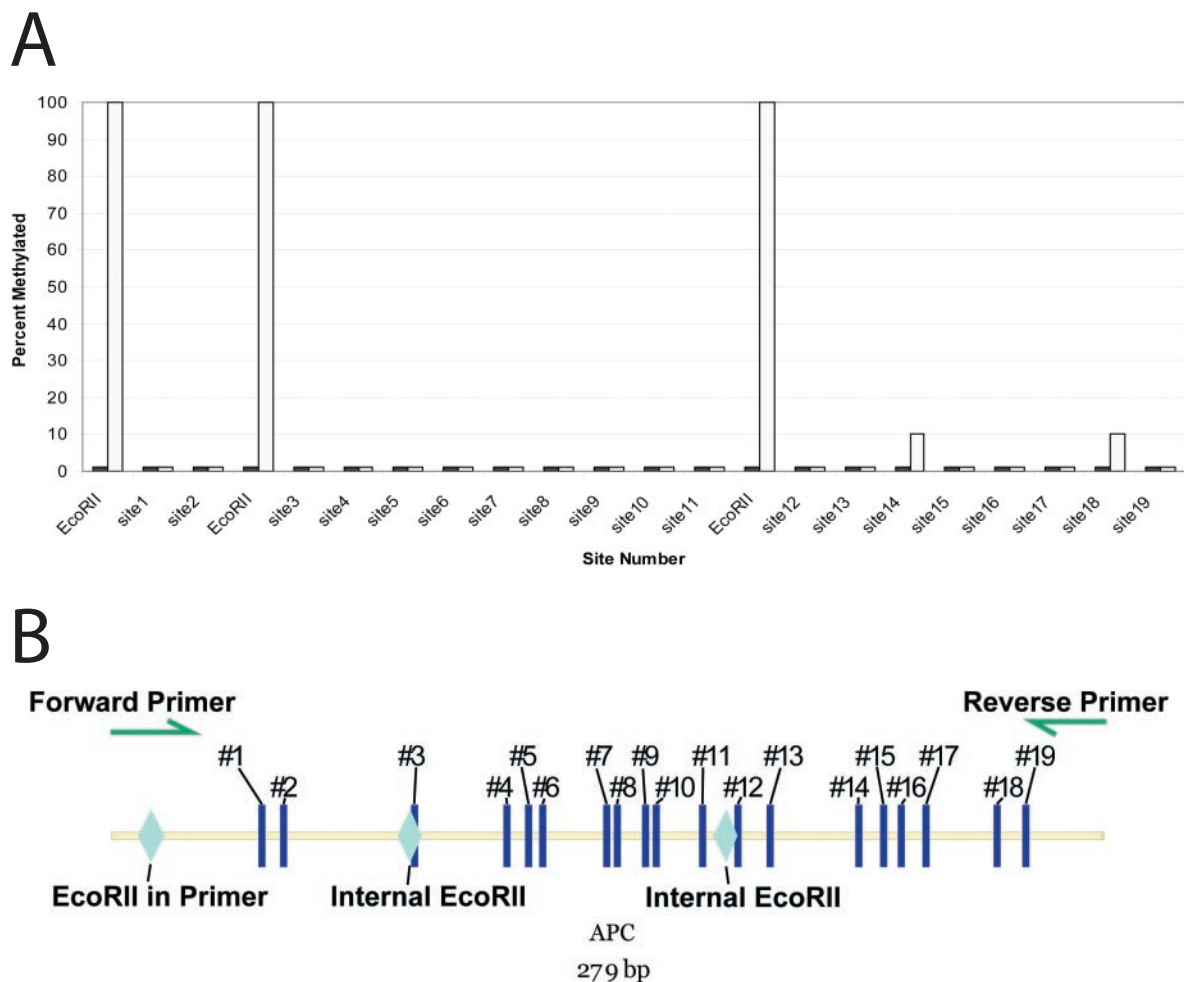
A

B

Figure 7. Effect of C^mCWGG methylation at a promoter that is unmethylated at CG sites. Bisulfite treated DNA from the APC promoter was amplified cloned and sequenced as described in Materials and Methods. The 279 bp amplicon contains 19 CG dinucleotide pairs. It contains three CCWGG sites, one in the forward primer, and two downstream from the forward primer. Ten clones were sequenced in order to determine the methylation state at each of the 19 CG sites and the three CCWGG sites in each clone. (**A**) Bar graph depicting the observed fraction of methylated sequences at each site in the two types of transformed cells. (**B**) Those expressing GFP alone are shown in grey (left), and those expressing the M.EcoRII-GFP fusion are shown in white (right). Map of the amplicon depicting the relative position of each of the sites in the bar graph.

M.EcoRII–GFP fusion protein to bind DNA and thereby interfere with normal replication and transcription functions in human cells. We initially sought to distinguish between these two forms of toxicity by constructing a catalytically inactive form of the enzyme for transformation under the conditions used here. However, we have noted (J. Clark, E. Singer, and S. S. Smith, unpublished data) that catalytically inactive DNA methyltransferases bind tightly to DNA *in vitro* based on gel retardation experiments. Thus such an experiment would be inconclusive.

Although we cannot determine the mechanism from the data at hand, it is clear that cell cultures in which a high percentage of cells express both the *N*-acetyl-transferase protein from the puromycin resistance gene and M.EcoRII–GFP fusion protein from the methyltransferase gene are more difficult to select than cell cultures expressing both the puromycin resistance enzyme and the GFP protein alone.

As expected, only the cell line expressing the M.EcoRII–GFP fusion protein acquired methylation at CCWGG sites. Methylation at this site was shown to span relatively large

regions of DNA (Figure 5) where CCWGG methylated segments as large as 10 and 12 kb could be detected. Although the sample of 32 clones is small, a significant over-representation of ALU sequences is present in the R.EcoRII-protected DNA. Since ALU sequences are preferentially found in the R-band, gene-rich fractions of the genome, it would appear that the M.EcoRII–GFP fusion gains access to the gene-rich fractions of the HK293 genome.

In spite of the significant level of C^mCWGG methylation generated by the expression of the M.EcoRII–GFP fusion, it had no detectable effect on global methylation at CG sites as measured by end-counting with the cytosine extension assay (Table 2). A similar lack of effect on bulk methylation was observed in a previous study when the CG methyltransferase M.HhaI was expressed in mouse cells (60). In that case it was suggested that the absolute level of CG methylation was maintained constant whereas the patterning of CG methylation was shifted. Given reports suggesting that CCWGG methylation is not found adjacent to CG methylation in human cells (29,30,33,34), it was possible that the pattern of CG
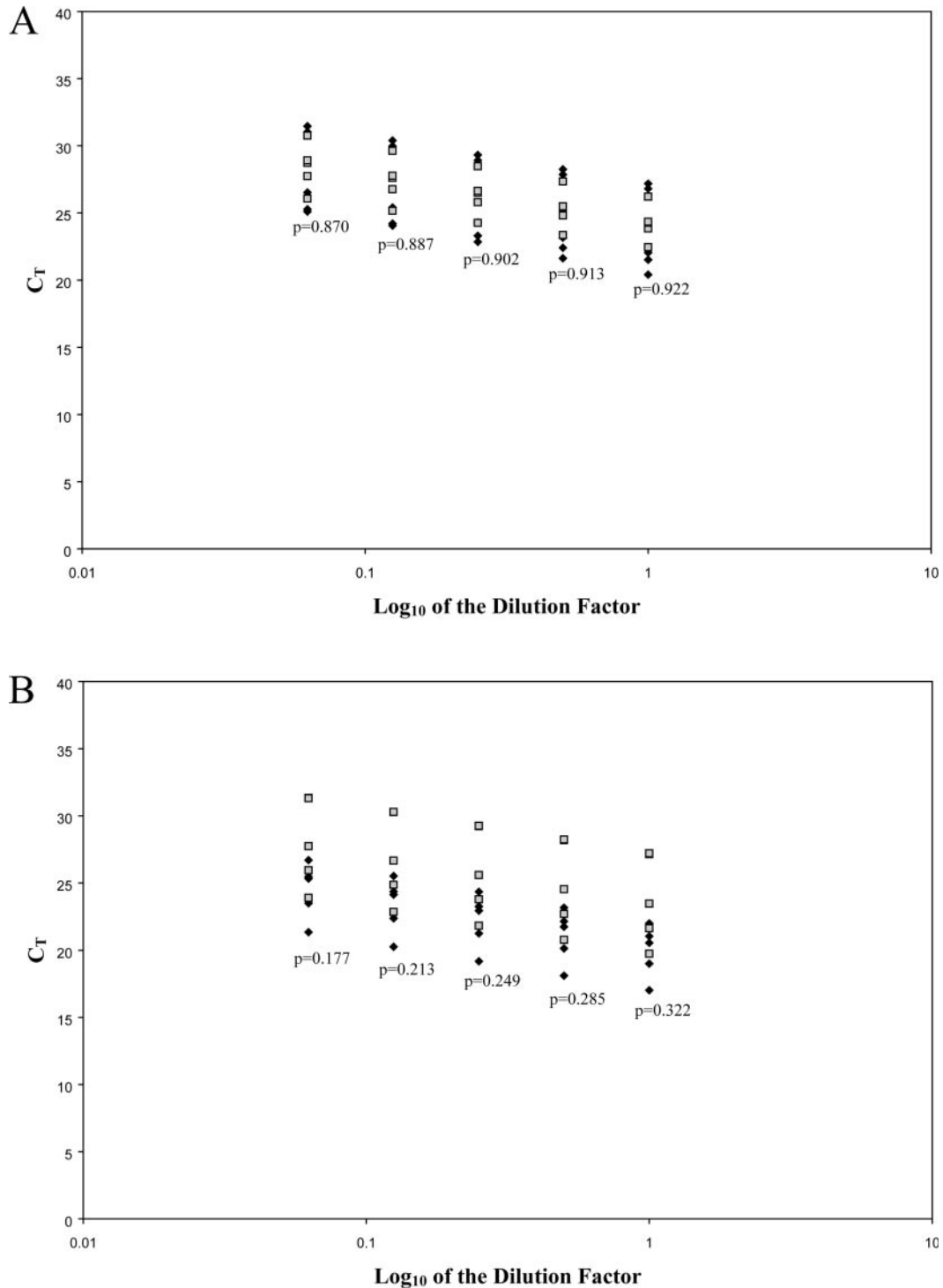
Figure 8. Comparison of RNA expression levels from the APC gene and the hTERT gene in cells expressing GFP alone, or the M.EcoRII–GFP fusion. C_t values obtained at each fold dilution of input RNA are plotted. Two independent RNA preparations were used, one preparation was tested in duplicate and one preparation was tested in triplicate. *P*-values from χ^2 analyses are listed below each set of data points. Each value is >>0.05, indicating that there is no significant difference in the level of RNA expression when comparing the two cell lines, for either of the two genes tested. (**A**) Data for the hTERT gene. closed squares, cells expressing the M.EcoRII–GFP fusion. Closed diamonds, cells expressing the GFP alone. (**B**) Data for the APC gene. Closed squares, cells expressing the M.EcoRII–GFP fusion. Closed diamonds, cells expressing the GFP alone.

methylation was shifted in the cells we used. However, this is consistent neither with the enzymological evidence for Dnmt1 (Figure 9) nor with the data on the individual promoters we investigated (Figures 6 and 7).

Since the patterning of CG methylation appears to be carefully controlled at promoter sequences, it was of interest to determine whether or not C^mCWGG methylation applied by the M.EcoRII–GFP fusion altered CG methylation patterning
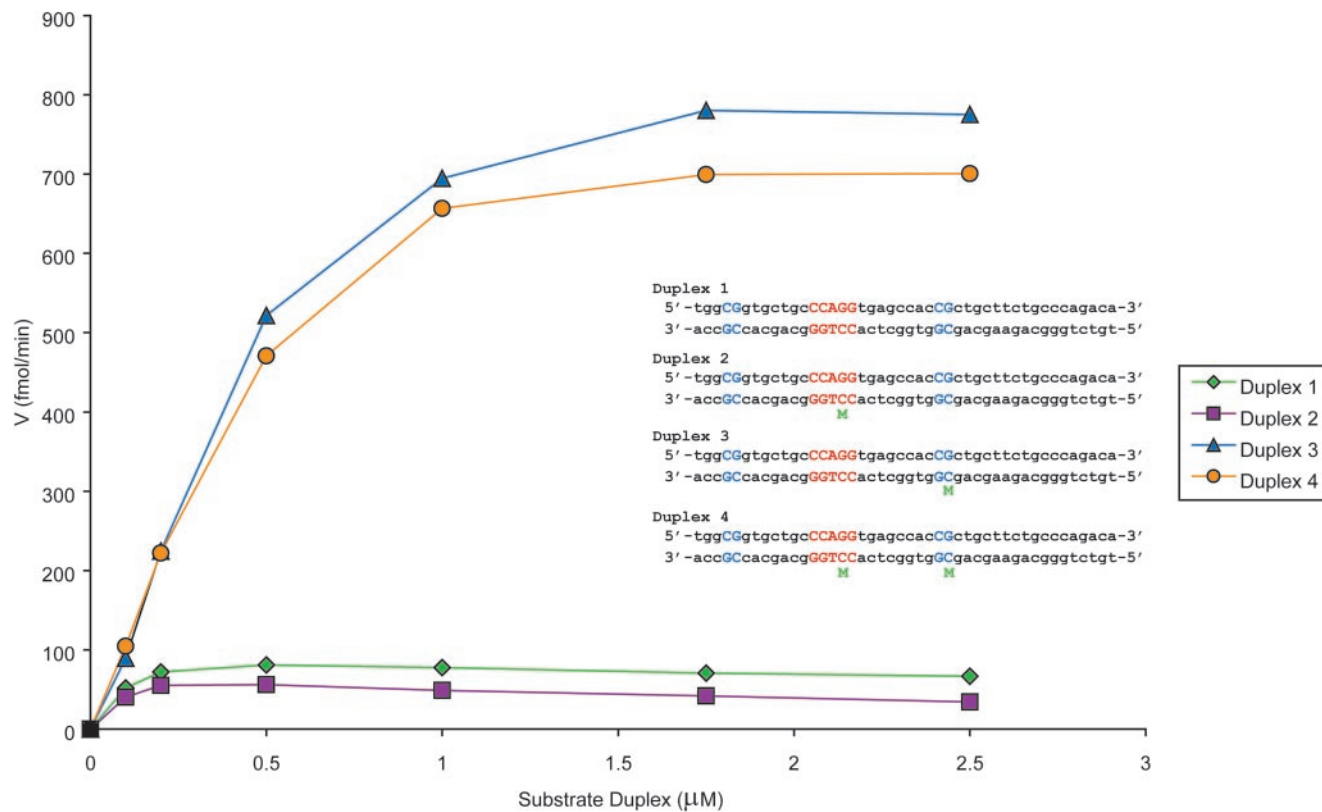
Figure 9. Kinetics of methylation by human DNMT1 at C^mCWGG methylated oligodeoxynucleotide substrates from the SERPINB5 promoter. Saturation curves for SERPINB5 oligodeoxynucleotide substrates with varying states of methylation at CCWGG and adjacent CG site are depicted. Duplex 1: saturation curve for the unmethylated duplex. Duplex 2: saturation curve for the duplex carrying a hemimethylated CCWGG site. Duplex 3: saturation curve for the duplex carrying a hemimethylated CG site. Duplex 4: saturation curve for the duplex with both the CCWGG and the CG sites in the hemimethylated state.

at representative promoters. Although the ALU cloning evidence suggested that the transgenically expressed enzyme gained access to gene-rich sequences, it was not clear that it would be able to methylate a heavily CG methylated promoter that presumably resides in condensed chromatin. In fact studies with yeast expressing transgenic M.BspRI (61), the *E.coli Dam* methyltransferase (62) and M.SssI (63) suggested that condensed chromatin sites might not be accessible to prokaryotic methyltransferases. Thus, we inspected the SERPINB5 promoter region in HK293 cells expressing the M.EcoRII–GFP fusion because it is known to be heavily methylated at CG sites in this cell line (40). Consistent with the enzymological evidence (Figure 9), no difference in patterning was observed for CG sites in this promoter when C^mCWGG was introduced (Figure 6), nor was the bacterial enzyme activity in the M.EcoRII–GFP fusion excluded from the promoter sequence. In addition, no difference in methylation patterning at CG sites was observed (Figure 7) when C^mCWGG was introduced at the APC promoter (a promoter that is not methylated at CG sites in this cell line). Both findings are consistent with the enzymology of wild-type Dnmt1 isolated from human placenta as reported previously (1,48,64–66). This makes it unlikely that the selection process we used has selected for an altered form of Dnmt1.

Thus, strong conclusions can be made that are independent of the nature of the cells that result from selecting for the expression of M.EcoRII–GFP. First, we conclude that the

transgenically expressed M.EcoRII–GFP fusion can methylate gene-rich genomic DNA sequences in both condensed and more open chromatin conformations in viable human cells. Second, we conclude that neither the transient presence of M.EcoRII–GFP nor the C^mCWGG that it applies have a direct effect on either bulk levels of CG methylation or the patterning on CG methylation at CG sites in HK293 cells.

The data on RNA expression from the APC gene are also consistent with this conclusion. Continued expression from the C^mCWGG methylated promoter suggests that methylation at CCWGG sites may not be tightly coupled to gene expression. However, if we assume that non-CG methylation does not naturally occur at this promoter, then the lack of response might be expected. Further, Malone *et al.* (30) have suggested that the effects of CCWGG methylation may be mediated by the CCWGG binding proteins they observe in lymphoma cells. Thus, the lack of effect we observe could also indicate that these proteins are absent in HK293 cells.

Finally, it is important to point out that the Q-PCR showing no significant change in expression from APC is supported by the bisulfite sequencing data on APC. This is because the promoter region inspected has been studied by many groups interested in the expression of the APC tumor-suppressor gene (67). The results of those studies show that this region is methylated at CG sites in tumor cells that do not express the gene. Thus, CG methylation in this region, although it may or may not play an active role in downregulating the

gene, is clearly a marker of the downregulated state. For this reason, continued expression of RNA from the APC gene provides an independent confirmation of the lack of influence of CCWGG methylation on CG methylation in the APC promoter. Based on the preponeance of evidence in the field, it is reasonable to expect that if the CCWGG methylation had downregulated APC expression, CG methylation would now be expected to mark the downregulated state. APC modulation of the hDnmt1 promoter is also consistent with this result, since downregulation of APC would be expected to enhance CG methylation levels through an elevation of hDnmt1 expression as reported in (68).

In summary then, all of the data presented are consistent with the conclusion that CCWGG methylation does not directly force changes in CG methylation.

ACKNOWLEDGEMENTS

REFERENCES

1. Smith,S.S., Kaplan,B.E., Sowers,L.C. and Newman,E.M. (1992) Mechanism of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation. *Proc. Natl Acad. Sci. USA*, **89**, 4744–4748.
2. Smith,S.S., Hardy,T.A. and Baker,D.J. (1987) Human DNA (cytosine-5)methyltransferase selectively methylates duplex DNA containing mispairs. *Nucleic Acids Res.*, **15**, 6899–6916.
3. Smith,S.S., Lingeman,R.G. and Kaplan,B.E. (1992) Recognition of foldback DNA by the human DNA (cytosine-5-)-methyltransferase. *Biochemistry*, **31**, 850–854.
4. Christman,J.K., Sheikhnejad,G., Marasco,C.J. and Sufrin,J.R. (1995) 5-Methyl-2′-deoxycytidine in single-stranded DNA can act in *cis* to signal *de novo* DNA methylation. *Proc. Natl Acad. Sci. USA*, **92**, 7347–7351.
5. Chen,X., Mariappan,S.V., Catasti,P., Ratliff,R., Moyzis,R.K., Laayoun,A., Smith,S.S., Bradbury,E.M. and Gupta,G. (1995) Hairpins are formed by the single DNA strands of the fragile X triplet repeats: structure and biological implications. *Proc. Natl Acad. Sci. USA*, **92**, 5199–5203.
6. Smith,S.S., Laayoun,A., Lingeman,R.G., Baker,D.J. and Riley,J. (1994) Hypermethylation of telomere-like foldbacks at codon 12 of the human c-Ha-ras gene and the trinucleotide repeat of the FMR-1 gene of fragile X. *J. Mol. Biol.*, **243**, 143–151.
7. Burden,A.F., Manley,N.C., Clark,A.D., Gartler,S.M., Laird,C.D. and Hansen,R.S. (2005) Hemimethylation and non-CpG methylation levels in a promoter region of human LINE-1 (L1) repeated elements. *J. Biol. Chem.*, **280**, 14413–14419.
8. Laird,C.D., Pleasant,N.D., Clark,A.D., Sneeden,J.L., Hassan,K.M., Manley,N.C., Vary,J.C.,Jr, Morgan,T., Hansen,R.S. and Stoger,R. (2004) Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc. Natl Acad. Sci. USA*, **101**, 204–209.
9. Genereux,D.P., Miner,B.E., Bergstrom,C.T. and Laird,C.D. (2005) A population-epigenetic model to infer site-specific methylation rates from double-stranded DNA methylation patterns. *Proc. Natl Acad. Sci. USA*, **102**, 5802–5807.
10. Woodcock,D.M., Lawler,C.B., Linsenmeyer,M.E., Doherty,J.P. and Warren,W.D. (1997) Asymmetric methylation in the hypermethylated CpG promoter region of the human L1 retrotransposon. *J. Biol. Chem.*, **272**, 7810–7816.
11. Smith,S.S. (1991) DNA methylation in eukaryotic chromosome stability. *Mol. Carcinog.*, **4**, 91–92.
12. Smith,S.S. and Crocitto,L.E. (1999) DNA methylation in eukaryotic chromosome stability revisited: DNA methyltransferase in the management of DNA conformation space. *Mol. Carcinog.*, **26**, 1–9.
13. Bestor,T.H. and Tycko,B. (1996) Creation of genomic methylation patterns. *Nature Genet.*, **12**, 363–367.
14. Easwaran,H.P., Schermelleh,L., Leonhardt,H. and Cardoso,M.C. (2004) Replication-independent chromatin loading of Dnmt1 during G_2 and M phases. *EMBO Rep.*, **5**, 1181–1186.
15. Evans,H.H., Evans,T.E. and Littman,S. (1973) Methylation of parental and progeny DNA strands in *Physarum polycephalum*. *J. Mol. Biol.*, **74**, 563–572.
16. Issa,J.P., Gharibyan,V., Cortes,J., Jelinek,J., Morris,G., Verstovsek,S., Talpaz,M., Garcia-Manero,G. and Kantarjian,H.M. (2005) Phase II study of low-dose decitabine in patients with chronic myelogenous leukemia resistant to imatinib mesylate. *J. Clin. Oncol.*, **23**, 3948–3956.
17. Flatau,E., Gonzales,F.A., Michalowsky,L.A. and Jones,P.A. (1984) DNA methylation in 5-aza-2′-deoxycytidine-resistant variants of C3H 10T1/2 C18 cells. *Mol. Cell. Biol.*, **4**, 2098–2102.
18. Rountree,M.R., Bachman,K.E. and Baylin,S.B. (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nature Genet.*, **25**, 269–277.
19. Liu,Y., Oakeley,E.J., Sun,L. and Jost,J.P. (1998) Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. *Nucleic Acids Res.*, **26**, 1038–1045.
20. Woodcock,D.M., Crowther,P.J. and Diver,W.P. (1987) The majority of methylated deoxycytidines in human DNA are not in the CpG dinucleotide. *Biochem. Biophys. Res. Commun.*, **145**, 888–894.
21. Clark,S.J., Harrison,J. and Frommer,M. (1995) CpNpG methylation in mammalian cells. *Nature Genet.*, **10**, 20–27.
22. Lorincz,M.C. and Groudine,M. (2001) C(m)C(a/t)GG methylation: a new epigenetic mark in mammalian DNA? *Proc. Natl Acad. Sci. USA*, **98**, 10034–10036.
23. Pinarbasi,E., Elliott,J. and Hornby,D.P. (1996) Activation of a yeast pseudo DNA methyltransferase by deletion of a single amino acid. *J. Mol. Biol.*, **257**, 804–813.
24. Hermann,A., Schmitt,S. and Jeltsch,A. (2003) The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. *J. Biol. Chem.*, **278**, 31717–31721.
25. Gowher,H. and Jeltsch,A. (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *J. Mol. Biol.*, **309**, 1201–1208.
26. Fatemi,M., Hermann,A., Gowher,H. and Jeltsch,A. (2002) Dnmt3a and Dnmt1 functionally cooperate during *de novo* methylation of DNA. *Eur. J. Biochem.*, **269**, 4981–4984.
27. Baylin,S. and Bestor,T.H. (2002) Altered methylation patterns in cancer cell genomes: cause or consequence? *Cancer Cell*, **1**, 299–305.
28. Ramsahoye,B.H., Biniszkiewicz,D., Lyko,F., Clark,V., Bird,A.P. and Jaenisch,R. (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl Acad. Sci. USA*, **97**, 5237–5242.
29. Franchina,M. and Kay,P.H. (2000) Evidence that cytosine residues within 5′-CCTGG-3′ pentanucleotides can be methylated in human DNA independently of the methylating system that modifies 5′-CG-3′ dinucleotides. *DNA Cell Biol.*, **19**, 521–526.
30. Malone,C.S., Miner,M.D., Doerr,J.R., Jackson,J.P., Jacobsen,S.E., Wall,R. and Teitell,M. (2001) CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing. *Proc. Natl Acad. Sci. USA*, **98**, 10404–10409.
31. Shevchuk,T.V., Zacharchenco,N.S., Dyachencko,O.V. and Buryanov,Y.I. (2001) The influence of EcoRII DNA methyltransferase gene on a phenotype of transgenic tumor lines of *Nicotiana tabacum* and methylation level of CpNpG sequences in their genomes. *Russ. J. Plant Physiol.*, **48**, 478–482.
32. Tariq,M. and Paszkowski,J. (2004) DNA and histone methylation in plants. *Trends Genet.*, **20**, 244–251.

33. Marinitch,D.V., Vorobyev,I.A., Holmes,J.A., Zakharchenko,N.S., Dyachenko,O.V., Buryanov,Y.I. and Shevchuk,T.V. (2004) Hypermethylation of 5′-region of the human calcitonin gene in leukemias: structural features and diagnostic significance. *Biochemistry (Moscow)*, **69**, 340–349.

34. Buryanov,Y.I., Shevchuk,T.V., Zakharchenko,N.S., D'yachenko,O.V., Marinich,D.V. and Vorob'ev,I.A. (2000) The absence of the CpNpG methylation at the 5′-terminal region of the human calcitonin cene in norm and leukemias. *Russ. J. Bioorg. Chem.*, **26**, 358–360.

35. Buryanov,Y.I., Bogdarina,I.G. and Bayev,A.A. (1978) Site specificity and chromatographic properties of *E. coli* K12 and EcoRII DNA-cytosine methylases. *FEBS Lett.*, **88**, 251–254.

36. Sambrook,J. and Rusell,D.W. (2001) *Molecular Cloning: A Laboratory Manual, 3rd edn*. Cold Spring Harbor Laboratory Press, NY.

37. Tolberg,M.E. and Smith,S.S. (1984) Methylation of a middle repetitive DNA sequence class during differentiation in Friend erythroleukemia cells. *FEBS Lett.*, **176**, 250–254.

38. Smith,S.S., Gilroy,T.E. and Ferrari,F.A. (1983) The influence of agarose—DNA affinity on the electrophoretic separation of DNA fragments in agarose gels. *Anal. Biochem.*, **128**, 138–151.

39. Botchan,M., McKenna,G. and Sharp,P.A. (1974) Cleavage of mouse DNA by a restriction enzyme as a clue to the arrangement of genes. *Cold Spring Harb. Symp. Quant. Biol.*, **38**, 383–395.

40. Futscher,B.W., Oshiro,M.M., Wozniak,R.J., Holtan,N., Hanigan,C.L., Duan,H. and Domann,F.E. (2002) Role for DNA methylation in the control of cell type specific maspin expression. *Nature Genet.*, **31**, 175–179.

41. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.

42. Devereux,T.R., Horikawa,I., Anna,C.H., Annab,L.A., Afshari,C.A. and Barrett,J.C. (1999) DNA methylation analysis of the promoter region of the human telomerase reverse transcriptase (hTERT) gene. *Cancer Res.*, **59**, 6087–6090.

43. Dessain,S.K., Yu,H., Reddel,R.R., Beijersbergen,R.L. and Weinberg,R.A. (2000) Methylation of the human telomerase gene CpG island. *Cancer Res.*, **60**, 537–541.

44. Ramakers,C., Ruijter,J.M., Deprez,R.H. and Moorman,A.F. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci. Lett.*, **339**, 62–66.

45. Luu-The,V., Paquet,N., Calvo,E. and Cumps,J. (2005) Improved real-time RT–PCR method for high-throughput measurements using second derivative calculation and double correction. *Biotechniques*, **38**, 287–293.

46. Pogribny,I., Yi,P. and James,S.J. (1999) A sensitive new method for rapid detection of abnormal methylation patterns in global DNA and within CpG islands. *Biochem. Biophys. Res. Commun.*, **262**, 624–628.

47. Pogribny,I., Raiche,J., Slovack,M. and Kovalchuk,O. (2004) Dose-dependence, sex- and tissue-specificity, and persistence of radiation-induced genomic DNA methylation changes. *Biochem. Biophys. Res. Commun.*, **320**, 1253–1261.

48. Clark,J., Shevchuk,T., Kho,M.R. and Smith,S.S. (2003) Methods for the design and analysis of oligodeoxynucleotide-based DNA (cytosine-5) methyltransferase inhibitors. *Anal. Biochem.*, **321**, 50–64.

49. Smith,S.S. and Ratner,D.I. (1991) Lack of 5-methylcytosine in *Dictyostelium discoideum* DNA. *Biochem. J.*, **277**, 273–275.

50. Eichinger,L., Pachebat,J.A., Glockner,G., Rajandream,M.A., Sucgang,R., Berriman,M., Song,J., Olsen,R., Szafranski,K., Xu,Q. *et al.* (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature*, **435**, 43–57.

51. Smith,S.S. and Thomas,C.A.,Jr (1981) The two-dimensional restriction analysis of Drosophila DNAs: males and females. *Gene*, **13**, 395–408.

52. Tweedie,S., Ng,H.H., Barlow,A.L., Turner,B.M., Hendrich,B. and Bird,A. (1999) Vestiges of a DNA methylation system in *Drosophila melanogaster*? *Nature Genet.*, **23**, 389–390.

53. Tang,L.Y., Reddy,M.N., Rasheva,V., Lee,T.L., Lin,M.J., Hung,M.S. and Shen,C.K. (2003) The eukaryotic DNMT2 genes encode a new class of cytosine-5 DNA methyltransferases. *J. Biol. Chem.*, **278**, 33613–33616.

54. Lyko,F., Ramsahoye,B.H., Kashevsky,H., Tudor,M., Mastrangelo,M.A., Orr-Weaver,T.L. and Jaenisch,R. (1999) Mammalian (cytosine-5) methyltransferases cause genomic DNA methylation and lethality in *Drosophila*. *Nature Genet.*, **23**, 363–366.

55. Gowher,H., Leismann,O. and Jeltsch,A. (2000) DNA of *Drosophila melanogaster* contains 5-methylcytosine. *EMBO J.*, **19**, 6918–6923.

56. Klimasauskas,S., Kumar,S., Roberts,R.J. and Cheng,X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357–369.

57. Reinisch,K.M., Chen,L., Verdine,G.L. and Lipscomb,W.N. (1995) The crystal structure of HaeIII methyltransferase convalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. *Cell*, **82**, 143–153.

58. Smith,S.S. and Baker,D.J. (1997) Stalling of human methyltransferase at single-strand conformers from the Huntington's locus. *Biochem. Biophys. Res. Commun.*, **234**, 73–78.

59. Kwoh,T.J., Kwoh,D.Y., McCue,A.W., Davis,G.R., Patrick,D. and Gingeras,T.R. (1986) Introduction and expression of the bacterial PaeR7 methylase gene in mammalian cells. *Proc. Natl Acad. Sci. USA*, **83**, 7713–7717.

60. Wu,J., Herman,J.G., Wilson,G., Lee,R.Y., Yen,R.W., Mabry,M., de Bustros,A., Nelkin,B.D. and Baylin,S.B. (1996) Expression of prokaryotic HhaI DNA methyltransferase is transforming and lethal to NIH 3T3 cells. *Cancer Res.*, **56**, 616–622.

61. Feher,Z., Kiss,A. and Venetianer,P. (1983) Expression of a bacterial modification methylase gene in yeast. *Nature*, **302**, 266–268.

62. Singh,J. and Klar,A.J. (1992) Active genes in budding yeast display enhanced *in vivo* accessibility to foreign DNA methylases: a novel *in vivo* probe for chromatin structure of yeast. *Genes Dev.*, **6**, 186–196.

63. Kladde,M.P., Xu,M. and Simpson,R.T. (1999) DNA methyltransferases as probes of chromatin structure *in vivo*. *Methods Enzymol.*, **304**, 431–447.

64. Smith,S.S., Kan,J.L., Baker,D.J., Kaplan,B.E. and Dembek,P. (1991) Recognition of unusual DNA structures by human DNA (cytosine-5)methyltransferase. *J. Mol. Biol.*, **217**, 39–51.

65. Smith,S.S. (1994) Biological implications of the mechanism of action of human DNA (cytosine-5)methyltransferase. *Prog. Nucleic Acid Res. Mol. Biol.*, **49**, 65–111.

66. Kho,M.R., Baker,D.J., Laayoun,A. and Smith,S.S. (1998) Stalling of human DNA (cytosine-5) methyltransferase at single-strand conformers from a site of dynamic mutation. *J. Mol. Biol.*, **275**, 67–79.

67. Esteller,M., Sparks,A., Toyota,M., Sanchez-Cespedes,M., Capella,G., Peinado,M.A., Gonzalez,S., Tarafa,G., Sidransky,D., Meltzer,S.J. *et al.* (2000) Analysis of adenomatous polyposis coli promoter hypermethylation in human cancer. *Cancer Res.*, **60**, 4366–4371.

68. Campbell,P.M. and Szyf,M. (2003) Human DNA methyltransferase gene DNMT1 is regulated by the APC pathway. *Carcinogenesis*, **24**, 17–24.