5-Halogenated pyrimidine lesions within a CpG sequence context mimic 5-methylcytosine by enhancing the binding of the methyl-CpG-binding domain of methyl-CpG-binding protein 2 (MeCP2)

Victoria Valinluck[1], Pingfang Liu[1,2], Joseph I. Kang Jr[1], Artur Burdzy[1] and Lawrence C. Sowers[1,2,*]

[1]Department of Biochemistry and Microbiology, School of Medicine, Loma Linda University, Loma Linda, CA 92350, USA and [2]Graduate School of Biological Sciences, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, CA 91010, USA

ABSTRACT

Perturbations in cytosine methylation signals are observed in the majority of human tumors; however, it is as yet unknown how methylation patterns become altered. Epigenetic changes can result in the activation of transforming genes as well as in the silencing of tumor suppressor genes. We report that methyl-CpG-binding proteins (MBPs), specific for methyl-CpG dinucleotides, bind with high affinity to halogenated pyrimidine lesions, previously shown to result from peroxidase-mediated inflammatory processes. Emerging data suggest that the initial binding of MBPs to methyl-CpG sequences may be a seeding event that recruits chromatin-modifying enzymes and DNA methyltransferase, initiating a cascade of events that result in gene silencing. MBD4, a protein with both methyl-binding and glycosylase activity demonstrated repair activity against a series of 5-substituted pyrimidines, with the greatest efficiency against 5-chlorouracil, but undetectable activity against 5-chlorocytosine. The data presented here suggest that halogenated pyrimidine damage products can potentially accumulate and mimic endogenous methylation signals.

INTRODUCTION

The DNA of all living organisms is constantly attacked by reactive molecules (1). Many of the lesions formed miscode during DNA replication, leading to genetic mutations that in part underlie the development of cancer (2,3). In addition to coding changes in DNA, perturbations in epigenetic methylation patterns are observed in the majority of human tumors (4). The aberrant loss of epigenetic signals can result in the activation of oncogenes, whereas inappropriate methylation could lead to silencing of tumor suppressor genes. The significance of epigenetic alterations in the development of cancer is becoming more evident; however, the mechanisms by which these altered methylation patterns arise are as yet unknown.

Previous studies demonstrate that the chemical modification of key recognition points of the methylated CpG dinucleotide reverses the increased sequence-specific protein binding affinity afforded by cytosine methylation. The oxidation of the methyl group of 5-methylcytosine or the 8 position of guanine, as well as methylation of the N7 position of guanine, is known to substantially reduce methyl-CpG-binding protein (MBP) affinity (5,6). These findings suggest a possible mechanism by which DNA damage might result in decreased MBP binding and inappropriate gene activation.

Pathways that would lead to increased MBP binding and reduced expression of tumor suppressor genes are less well understood. The chemical alkylation of cytosine in DNA occurs on the exocyclic oxygen atom and N3 ring nitrogen, but not at the C5 position (7). Indeed, most forms of chemical alteration of the CpG dinucleotide would result in decreased MBP binding. One notable exception may be the halogenation of cytosine in DNA. In other systems, halogenated pyrimidines have been shown to mimic a 5-methyl pyrimidine in enhancing sequence-specific DNA–protein interactions (8,9).

Halogenated pyrimidine lesions have recently been shown to result from peroxidase-mediated inflammatory processes

---

*To whom correspondence should be addressed. Tel: +1 909 558 4480; Fax: +1 909 558 4035; Email: lsowers@som.llu.edu
Present address:
Pingfang Liu, Department of Genetics and Complex Disease, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, USA

(10–12). Under inflammatory conditions, phagocytic cells generate peroxidases that can form hypochlorous acid (HOCl) (13,14) and hypobromous acid (HOBr) (15–18), which can attack duplex DNA (19). Eosinophil peroxidases can also form NO_2Cl, which can also chlorinate nucleotides and nucleosides (20). The effect of these halogenated lesions on epigenetic control has not yet been explored.

In this study, we examined the effects of pyrimidine halogenation on the binding of one member of the MBP family, methyl-CpG-binding protein 2 (MeCP2). MeCP2, as well as other members of this family, has a 70–75 amino acid highly conserved domain that is referred to as the methyl-binding domain (MBD) (21–23). Previously, we demonstrated that the MBD of MeCP2 binds a symmetrically methylated duplex oligonucleotide with 100 times greater affinity than the corresponding unmethylated oligonucleotide in electrophoretic mobility shift assays (EMSAs) (5). Here, we extend these studies and demonstrate that the placement of the 5-halopyrimidines, 5-chlorocytosine (ClC), 5-chlorouracil (ClU), 5-bromocytosine (BrC) and 5-bromouracil (BrU) in synthetic oligonucleotides, paired opposite guanine, similarly increases the affinity of the MBD binding. The following MBD affinity to hemimethylated duplexes containing modified pyrimidines on the opposite strand was observed: BrC ≥ ClC ≥ mC > ClU ≫ U ≥ BrU ≫ C. These data suggest that halogenated pyrimidine lesions could mimic 5mC in promoting the binding of proteins containing an MBD, resulting in unintended and potentially heritable gene silencing.

The increased binding of the MBD to the oligonucleotides containing 5-halouracil residues prompted us to investigate an additional member of the MBP family, MBD4. In addition to an MBD domain, MBD4 also has a uracil glycosylase activity (23–26). The MBD4 glycosylase activity is known to excise thymine opposite guanine in a CpG dinucleotide and is thought to function in the repair of lesions arising from the deamination of 5mC (24–26). In the study reported here, MBD4 was found to excise ClU ≥ U > T ≥ BrU > BrC > 5mC within a CpG dinucleotide. The chlorination of a cytosine residue in DNA could generate either ClC or ClU paired with guanine (27). Both lesions would increase the binding affinity of an MBP relative to an unmodified oligonucleotide. The removal of ClU arising from the chlorination of cytosine might be an important function of MBD4. In contrast, no repair activity has yet been demonstrated for ClC. The formation and persistence of ClC in DNA could lead to inappropriate gene silencing by mimicking endogenous methylation signals. We introduce here a previously unexplored mechanism by which inflammation could be involved in cancer development (28–31).

MATERIALS AND METHODS

Oligonucleotide synthesis

Oligonucleotide 27mers (Figure 1A) containing a central CpG dinucleotide with cytosine or uracil analogues were prepared by standard solid phase synthesis using either the Gene Assembler Plus (Pharmacia) or Expedite Nucleic Acid Synthesis System (Applied Biosystems) automated DNA synthesizers (32). The sequence used in this study (Figure 1A) was chosen based upon binding experiments previously conducted
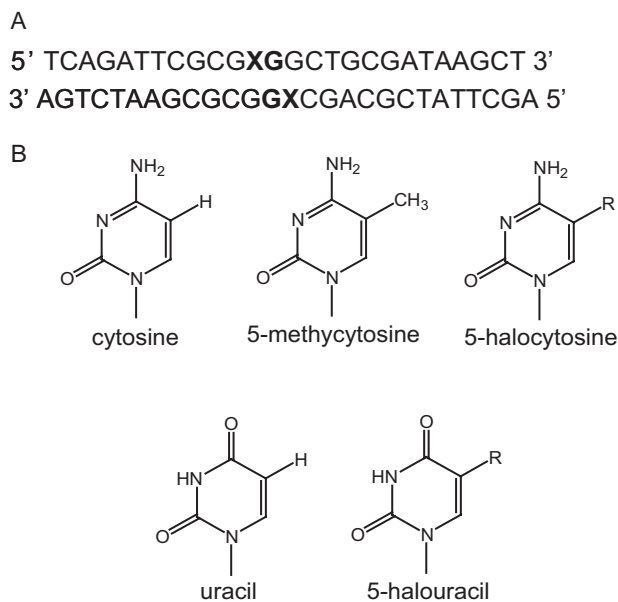
A

5' TCAGATTCGCG**XG**GCTGCGATAAGCT 3'
3' AGTCTAAGCGCG**GX**CGACGCTATTCGA 5'

B

Figure 1. (**A**) Sequence of 27mer oligonucleotide duplex used in EMSA with the MBD of MeCP2. The X position indicates the pyrimidine analogue within the CpG dinucleotide. In representing the duplexes, the pyrimidine modifications at the X position in the upper strand are noted as well as the modification at the X position on the lower strand. For example, a duplex containing 5-methycytosine in the upper strand and 5-chlorocytosine in the lower strand would be represented as 5mC/ClC. (**B**) Structures of cytosine, 5-methylcytosine, 5-halogenated cytosine lesions, uracil and 5-halogenated uracil lesions. R indicates chloro or bromo groups.

with the MBD of MeCP2 (5,22,33). The X symbols indicate the positions where C is substituted with pyrimidine derivatives 5mC, ClC, BrC, U, ClU or BrU. The 5-chlorocytosine (ClC) phosphoramidite was prepared according to the methods developed by this laboratory (27). The 5-chlorouracil (ClU) phosphoramidite was synthesized as previously done by our laboratory according to the method developed by Brandon *et al.* (34,35). All other phosphoramidites used were obtained from Glen Research. Oligonucleotides were removed from the solid support and deprotected in aqueous ammonia (Aldrich) at 60°C overnight. The deprotected oligonucleotides were purified with Poly-Pak II cartridges (Glen Research). The sequence composition of the oligonucleotides was confirmed via high performance liquid chromatography (HPLC) analysis following digest of the oligonucleotides with nuclease P1 (Sigma) at 37°C for 1 h and bacterial alkaline phosphatase (Sigma) at 37°C overnight.

Protein expression and purification

The pAFB105 construct encoding the 6X His-tagged methyl-CpG-binding domain (MBD) of mouse MeCP2, residues 77–165 (22,33), in a pET6H vector, was overexpressed in *Escherichia coli* BL21 (DE3)/pLysS. The expression and purification was done as previously described, according to the protocol of Free *et al.* (5,33).

Electrophoretic mobility shift assay

The 27mer oligonucleotides (Figure 1A) were 5'-^{32}P-end labeled by T4 polynucleotide kinase (New England Biolabs)

with [γ-^{32}P]ATP (MP Biomedicals, LLC) under conditions recommended by the enzyme supplier. After purification using G50 Sephadex columns (Roche), the labeled strand (Figure 1A, lower strand) was incubated with 1.5-fold excess of the complementary unlabeled strand (Figure 1A, upper strand) in 20 mM HEPES pH 7.3, 1 mM EDTA at 95°C for 5 min and then allowed to slowly cool to room temperature for duplex formation. Duplexes were annealed, with either 5mC or C at position X in the lower strand (Figure 1A) and combinations of the halogenated pyrimidines at position X on the upper strand. Previous data indicate that placement of the base lesion in either the upper or lower strand yields the same binding results (5). To confirm duplex formation and the presence of cytosine modifications, the annealed oligonucleotides were digested with MspI and HpaII (New England Biolabs) according to conditions recommended by the enzyme supplier and the products were sized on denaturing 20% (v/v) polyacrylamide gels. MspI was able to cleave duplexes in which the cytosine was modified, both symmetrically and asymmetrically in the CpG site, with bromine or chlorine at the C5 position (data not shown). However, HpaII was able to cleave only the unmodified cytosine containing duplexes (data not shown).

Purified MBD (0, 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 128 and 256 nM) was incubated with 2 nM labeled duplex, 50 ng/μl of poly[dA–dT]·poly[dA–dT] (Sigma) in 20 mM HEPES pH 7.3, 1 mM EDTA, 10 nM (NH$_4$)$_2$SO$_4$, 1 mM dithiothreitol, 0.2% Tween-20, 30 mM KCl for 15 min at room temperature in a 30 μL reaction volume, before the addition of 7.5 μL loading buffer (60% 0.25X TBE, 40% glycerol) (33). The binding reactions were then electrophoresed on 10% non-denaturing polyacrylamide (37.5:1 acrylamide:bis-acrylamide) gels at 250 V for 2.5 h at 4°C after prerunning the gel at 200 V for 1.5 h. Visualization and quantification of the gels were carried out using a phosphorimager and the ImageQuant 5.0 software (Amersham Biosciences).

Binding model and data analysis

The non-cooperative, single-site binding scheme previously used to describe the monomeric binding of the MBD of MeCP2 to its symmetrically methylated DNA substrate was also used in this study (5,22). From the EMSA data, the fraction of duplex bound was determined at each protein concentration as follows, where [E] is the concentration of unbound MBD, [O] is the concentration of unbound oligonucleotide and [EO] is the MBD–oligonucleotide complex:

Fraction duplex bound = [EO]/([EO] + [O]).

When [E] ≫ [O], then [E]$_{total}$ − [EO] = [E] ≈ [E]$_{total}$, where [E]$_{total}$ is the total concentration of MBD, both bound and unbound. The following equation can be used to determine the dissociation constant for the oligonucleotide duplexes used in the study (5):

Fraction duplex bound = [E]$_{total}$/([E]$_{total}$ + K$_d$).

The average of a minimum of three data sets obtained for each duplex were fitted to the equation above by non-linear regression using SigmaPlot software (SPSS Science).

MBD4 glycosylase activity assay

The same duplexes used in the EMSAs were used to assay for MBD4 glycosylase excision repair activity. Recombinant human MBD4 (0.6 pmoles with an estimated 10–15% active protein) and 5′-^{32}P-end labeled 27mer DNA duplexes (0.15 pmoles) were incubated at 37°C for 0.5, 1, 2, 5, 10, 30 and 60 min in cleavage buffer (20 mM HEPES-KOH pH 7.5, 1 mM DTT, 1 mM EDTA, 1 mg/ml BSA) in the presence of 2 units of uracil glycosylase inhibitor protein (UGI) to eliminate any potential uracil DNA glycosylase activity present in the MBD4 glycosylase preparation. We define 1 U of UGI (New England Biolabs) as the amount of protein required to inhibit 1 U of *E.coli* UDG in 1 h at 37°C in a total reaction volume of 50 μl; 1 U of UDG is the amount of enzyme that will catalyze the release of 60 pmol of uracil per minute from double-stranded, uracil-containing DNA. Reactions were stopped by heating to 75°C for 5 min, and then cooled to room temperature for 30 min. The abasic sites were cleaved by 1 U of human AP endonuclease (Trevigen) at 37°C for 1 h in reaction buffers provided by the manufacturer. Trevigen defines 1 U of human APE as the amount that cleaves 1 pmol of a ^{32}P-labeled apurinic/apyrimidinic site oligonucleotide in 1 h at 37°C. An equal volume of loading buffer (98% formamide, 0.01 M EDTA, 1 mg/ml xylene cyanol and 1 mg/ml bromophenol blue) was added to stop the reaction. Samples were denatured and electrophoresed on 20% polyacrylamide gels containing 7M urea.

RESULTS

### Electrophoretic mobility shift assays

The binding affinities of the MBD of MeCP2 to oligonucleotide duplexes containing different halogenated pyrimidine lesions were determined. The sequence used in this study was chosen based upon previous binding studies done with the MBD of MeCP2 (5,22,33). The 27mer contains a central CpG dinucleotide (Figure 1A) that was systematically replaced with U, ClU, BrU, 5mC, ClC or BrC (Figure 1B), with either C or 5mC on the opposite strand within that CpG dinucleotide. The duplexes were assayed for binding to the MBD of MeCP2 at different concentrations ranging from 0 to 256 nM. Duplexes containing ClC or BrC in place of both cytosines in the CpG were also tested. The K$_d$ (Table 1) was determined for each duplex using non-linear regression to fit the equation for simple, non-cooperative monomeric binding.

Our results show that 5-halogenated cytosine analogs ClC and BrC in place of a single cytosine, within a methyl-CpG or unmethylated CpG, function as a 5mC residue (Figures 2 and 3, Table 1). The binding curves and K$_d$ determined for the 5mC/ClC, 5mC/BrC and 5mC/5mC duplexes (Figure 3B, Table 1) are very similar, consistent with previous suggestions that the MBD recognizes the methyl group of 5mC via a hydrophobic patch of residues (33,36,37). Also, these data indicate that the MBD tolerates a minimum range of 1.75–2.0 Å at the C5 position of cytosine. Binding with C/ClC, C/BrC and C/5mC (Figure 3C, Table 1) duplexes also showed similar binding affinity, confirming that ClC or BrC within a CpG context is recognized by the MBD. For the 5-halogenated uracil analogues replacing one cytosine in a single CpG

Table 1. Dissociation constants of the MBD of MeCP2 binding to 27mer duplexes containing 5-halogenated pyrimidines within a CpG or methyl-CpG sequence context

Duplex	K_d (nM)	Correlation coefficient (R^2)
5mC/5mC	14.7 ± 1.0	0.99
C/C	1030 ± 20	0.99
5mC/C	127 ± 3	0.99
ClC/ClC	13.4 ± 1.0	0.98
BrC/BrC	9.81 ± 0.78	0.99
5mC/ClC	11.0 ± 0.9	0.99
5mC/BrC	9.32 ± 0.77	0.99
C/ClC	131 ± 4	0.99
C/BrC	102 ± 3	0.99
5mC/U	92.6 ± 4.7	0.99
5mC/ClU	35.6 ± 3.6	0.97
5mC/BrU	110 ± 13	0.93
C/U	421 ± 17	0.99
C/ClU	199 ± 4	0.99
C/BrU	462 ± 18	0.99

Using SigmaPlot, the K_d for MBD binding to each of the duplexes was determined from non-linear regression of the three or more sets of data obtained from EMSA. The plot of the average percentage binding of each duplex against the concentration of MBD was fitted to the equation for simple, non-cooperative, monomeric binding (see Materials and methods). The ranges seen in the K_d values are the standard error for the regression analysis. The K_d and R^2 for MBD binding to the 5mC/5mC duplex were obtained from previous studies (5).
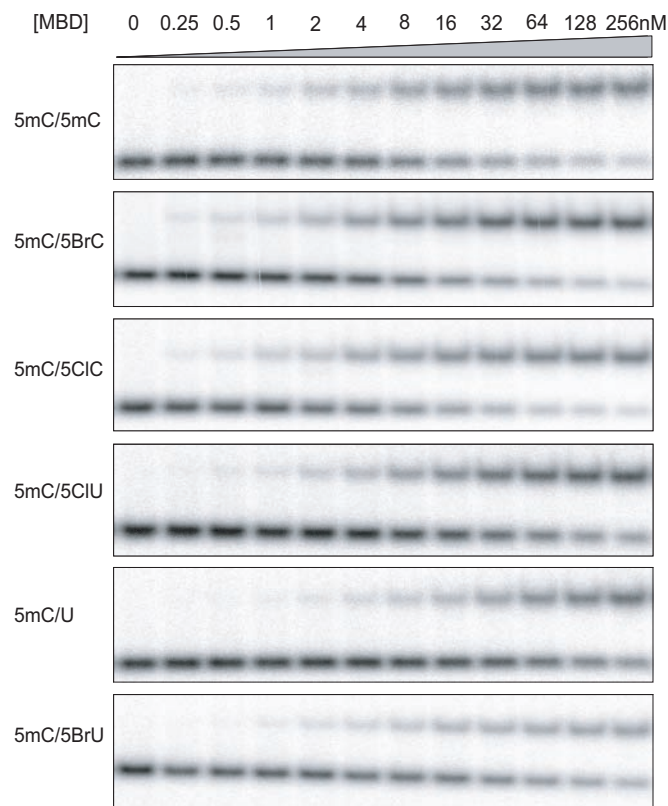
Figure 2. Binding of 5mC/5mC, 5mC/BrC, 5mC/ClC, 5mC/ClU, 5mC/U and 5mC/BrU duplexes to varying concentrations of MBD from 0 to 256 nM. The following affinities for MBD to 5-halogenated pyrimidine lesions within a methyl-CpG sequence are observed: BrC ≥ ClC ≥ 5mC ≥ T > ClU >> U ≥ BrU >> C (5).
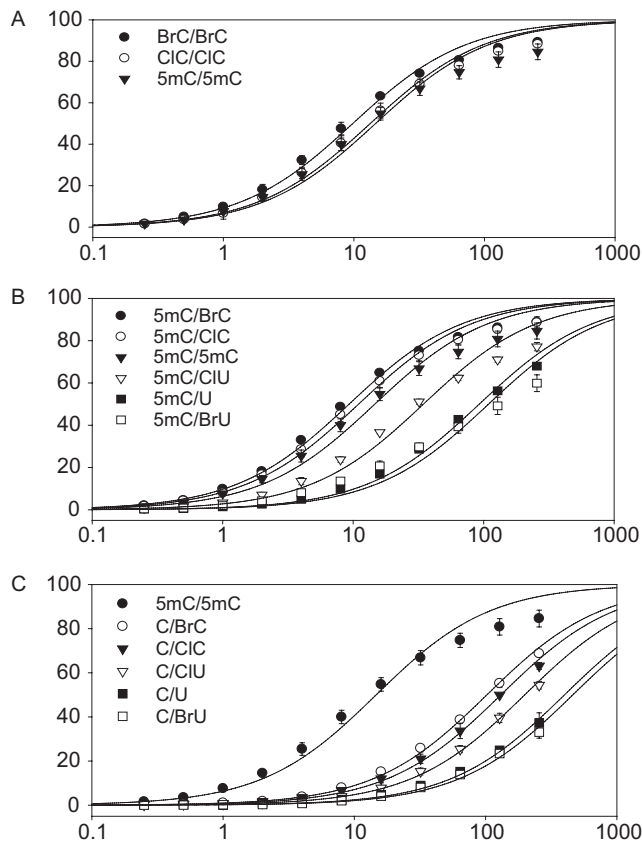
Figure 3. Non-linear regression of the plot of average percentage binding (determined from three or more sets of titrations per duplex) and concentration of MBD for selected modified duplexes tested in EMSA assay. Halogenated pyrimidine lesions in a CpG sequence context mimic 5mC by binding to the MBD with high affinity. The x-axis represents MBD concentration (nM). The y-axis represents present binding.

dinucleotide sequence, the observed binding affinities are as follows: ClU > U > BrU (Table 1). This trend holds true whether the uracil analog is replacing a cytosine in a methyl-CpG or an unmethylated CpG sequence.

MBD4 glycosylase activity assays

As MBD4 glycosylase is believed to play a role in the maintenance of CpG integrity (24–26), MBD4 was tested for repair activity against the halogenated pyrimidines paired with guanine in a CpG and mCpG context. MBD4-mediated removal of the target pyrimidines from the 27mer oligonucleotides generates an abasic site that is subsequently cleaved by human AP endonuclease or by alkaline conditions resulting in the formation of a shorter 13mer fragment. Our results show that MBD4 excises ClU ≥ U > T ≥ BrU > BrC > 5mC in both a CpG and mCpG site, but excision of ClC was not detected under our assay conditions (Figure 4A). In contrast to previous studies (25), our study shows that MBD4 glycosylase rates are modestly slower in a methylated CpG sequence than in a non-methylated CpG sequence (Figure 4B and C). These results are perhaps explained by the dual methyl-binding and glycosylase function of MBD4. The results suggest that the MBD4 may indeed be turning over faster within an unmethylated CpG oligonucleotide, whereas it plateaus with the methylated sequence.
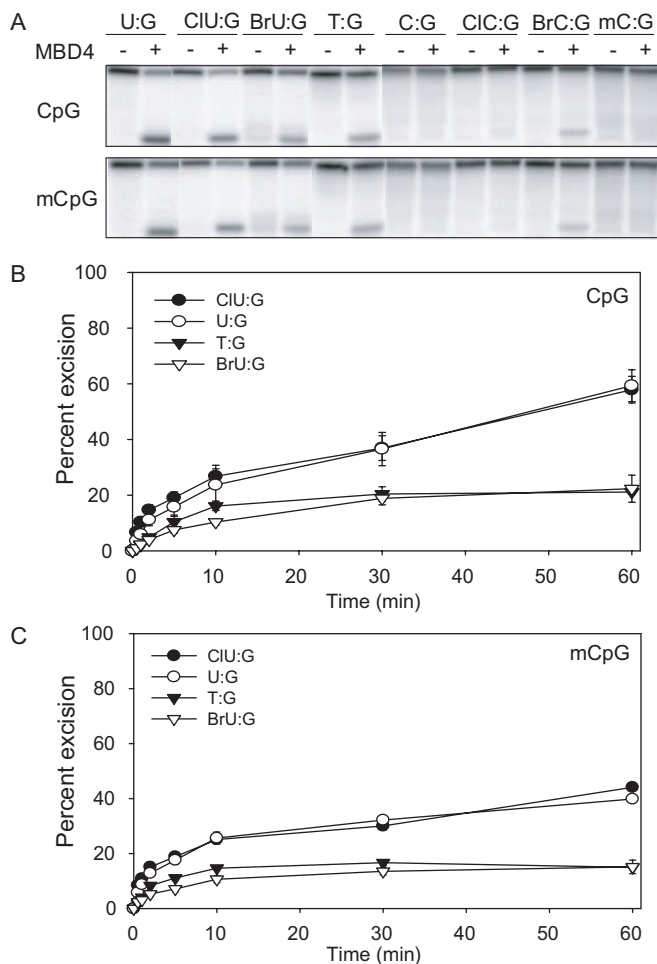
A

Figure 4. (**A**) MBD4 glycosylase excision activity at 60 min against a panel of 5-substituted pyrimidines paired with guanine. (**B**) Time course (0.5, 1, 2, 5, 10, 30 and 60 min) of MBD4 glycosylase excision repair of 5-halogenated uracil lesions within a CpG. (**C**) Time course (0.5, 1, 2, 5, 10, 30 and 60 min) of MBD4 glycosylase excision repair of 5-halogenated uracil lesions within a mCpG. MBD4 excises ClU ≥ U > T ≥ BrU > BrC > 5mC in both a CpG and an mCpG site, but no excision of ClC was detected under our assay conditions. To date, no repair activity of ClC has been detected in human cells.

DISCUSSION

Human tumors are observed to have a multitude of genetic changes, including chromosomal rearrangements, loss of entire chromosomes, point mutations and changes in cytosine methylation patterns. It is currently believed that these events develop over a period of years, and that a characteristic sequence of changes is often observed in the development of specific tumor types (38,39).

Changes in methylation patterns have been observed as both early and late events in tumor development. In prostate carcinogenesis, for example, methylation of the GSTpi gene is observed frequently in very early stages, including premalignant lesions (31,40,41). Inappropriate methylation of the promoter region of the GSTpi gene has been shown to result in transcriptional silencing in prostate cancer cells (4,40,41). The list of human tumors in which inappropriate promoter region methylation results in the silencing of tumor suppressor genes is rapidly growing (4).

Although aberrant promoter region methylation is frequently observed in human cancer, the mechanisms by which the promoter region becomes methylated have not yet been identified. Overexpression or inappropriate activity of the DNA methyltransferases has been proposed to explain hypermethylation, however, no consensus has yet been reached (42). Our laboratory has been investigating potential mechanisms by which DNA damage could be translated into altered methylation signals. However, most forms of DNA damage would interfere with both methyltransferase-mediated methylation as well as the binding of methyl-binding proteins involved in the recruitment of chromatin-modifying enzymes. One exception to this paradigm is pyrimidine halogenation, which is known to occur as a result of reactive molecules generated in the process of inflammation. Halogenation of cytosine in DNA can generate the corresponding 5-halocytosine and 5-halouracil derivatives (27).

Previous studies have demonstrated that the halogen, bromine, can substitute for the thymine methyl group in sequence-specific DNA protein interactions (8,9). Similar studies with chlorine-substituted pyrimidines were not possible until suitable synthetic methods were developed. Using recently developed methods (27), we constructed a series of oligonucleotide templates containing cytosine as well as 5-methyl-, bromo- and chloro-cytosine analogues to investigate the potential role of these analogues in MBP binding.

Our results (Table 1) demonstrate that both 5-chloro and 5-bromo substituents on cytosine substantially increase the affinity of the MBP for target oligonucleotides. Decreases in K_d of ~10- to 100-fold are observed with the ClC- and BrC-containing oligonucleotides in both unmethylated and hemimethylated sequences as compared with oligonucleotides containing unmodified cytosine. These results suggest that inflammation-mediated halogenation of cytosine residues in the DNA of previously unmethylated chromatin regions could increase the binding affinity of MBPs, which in turn would recruit additional chromatin-modifying enzymes, initiating a cascade of chromatin condensation and silencing. Previously, it was demonstrated that BrC could substitute for 5mC in directing methylation of the complementary strand by the human maintenance methylase (43). That methyltransferase study, in conjunction with the results reported here, suggests that cytosine halogenation could have dual roles in driving the aberrant methylation of previously unmethylated chromatin regions.

The most frequent point mutation observed in human tumors is the C to T transition at a CpG dinucleotide (44,45). It is widely presumed that the deamination of 5mC to T drives this mutation. Previous studies have demonstrated that the mispair formed by T and G at a CpG dinucleotide still serves as a high-affinity binding site for an MBP (5,24). Those studies showed that the presence of the methyl group, in either the 5mC:G or wobble T:G mispair, was sufficient to promote binding. It is known that the halogenation of cytosine in DNA can generate the cytosine derivatives discussed above, as well as the corresponding uracil derivatives by deamination. We therefore investigated whether ClU and BrU could similarly facilitate MBP binding. As shown in Table 1, we observe that the conversion of cytosine to ClU increases binding of MBP by a factor of five, only 2-fold lower than the affinity of the oligonucleotide containing the G:T mispair. Surprisingly,
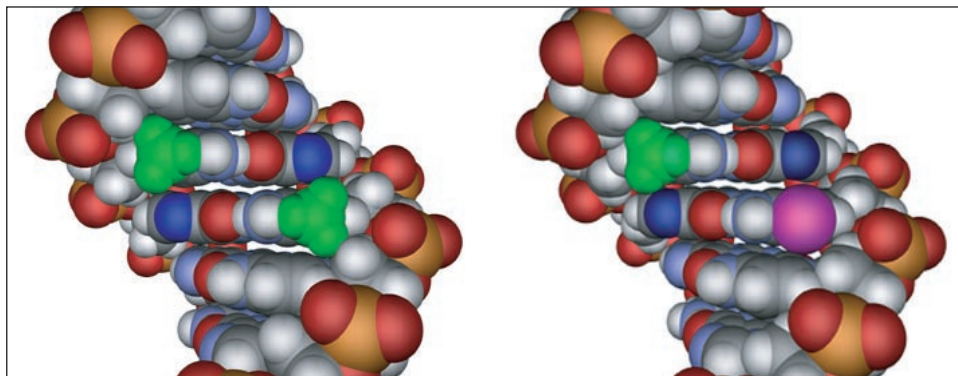
Figure 5. Molecular model of the sequence 5′-GC5mCGGC-3′/3′-CGGClCCG-5′. The methyl group of 5mC is depicted in green, the N7 position of the guanines within the methyl-CpG dinucleotide is depicted in dark blue, the chlorine atom is depicted in magenta. Multiple sites within the methylated CpG dinucleotide are needed for strong binding by the MBD. All four major sites of contact, two methyl groups and two hydrogen bond accepting nitrogens, are in the major groove of the DNA within close proximity of one another. Replacement of a methyl group with a chlorine results in a point of contact with nearly identical dimensions and positioning of the original methyl group, explaining the high affinity of MBD for ClC-containing CpG dinucleotides.

the corresponding conversion of cytosine to BrU does not increase MBP affinity. Although bromine is similar in size to the methyl group, it has a slightly larger van der Waals radius, and the carbon–bromine bond length is longer. When in the wobble geometry, the slightly larger footprint of the bromine substituent might result in a steric clash with the MBP.

The high frequency of transition mutations at methylated CpG dinucleotides is attributed, in part, to the relative inefficiency of repair of the deaminated thymine of the T:G mispair. Whereas uracil, the deamination product of cytosine, is rapidly repaired by an array of uracil DNA glycosylases, the range and activity of glycosylases that target and remove mispaired thymine residues are substantially less (46). Among the activities that could potentially repair a mispaired thymine residue derived from 5mC deamination is MBD4. MBD4 appears to have dual roles as both a member of the MBP family and a uracil/thymine glycosylase. MBD4 might have a significant role in protecting against transition mutations at CpG dinucleotides by residing in methylated areas of chromatin and initiating thymine repair (23–26). The ability to bind to both 5mC:G and T:G could facilitate this activity.

Upon the basis of the potential role of MBD4 in repairing T:G mispairs, we investigated the capacity of MBD4 to repair the halogenated uracil analogues that could arise by halogenation and deamination of cytosine in DNA (Figure 4). We observe that MBD4 is highly active against ClU mispaired with guanine, and that ClU is as good a substrate as uracil. These results are in accord with previous studies, in which it was shown that both 5-fluorouracil and uracil are good substrates for MBD4 (25). We observed that MBD4 activities against mispaired BrU and mispaired thymine are similar; however, the repair activity against these two lesions is less than the activity against other pyrimidines. Previously, it was shown that the glycosylase activity against thymine is less than uracil (24). Our data suggest that both the size and the inductive properties of the 5-substituent influence the repair efficiency of the lesion (47). Surprisingly, MBD4 also removes BrC, but not ClC. The formation of ClU at a CpG dinucleotide could facilitate MBP binding, as well as serve as a potentially

miscoding DNA lesion. The corresponding formation of BrU would result primarily in the introduction of a lesion that would code predominantly like thymine. The formation of either ClU or BrU from cytosine would give rise to a C to T transition mutation indistinguishable from a point mutation resulting from hydrolytic 5mC deamination.

In conclusion, the results reported here demonstrate that the halogenation of cytosine in a CpG dinucleotide does mimic the effect of enzymatic cytosine methylation with respect to the increased binding affinity of MBPs (Figure 5). To date, no repair activity has been demonstrated against ClC paired with guanine (27), although MBD4 has some activity against BrC. The ClC lesion may therefore be a persistent lesion in DNA.

The initial binding of the MBPs is thought to recruit further chromatin-modifying activities that together result in a chromatin structure with reduced transcriptional activity (48–50). Recent studies have demonstrated that MBPs can bind with and recruit methyltransferases (51). The ClC lesion may then serve a dual role in the recruitment of methyltransferase, as well as potentially directing the methylase-mediated methylation of the complementary strand of a hemihalogenated CpG dinucleotide. The initial cytosine halogenation could then serve as a signal for subsequent methylation seeding in a previously unmethylated chromatin region. Through this mechanism, it is proposed that the halogenation of cytosine by inflammation-mediated reactive species could result in the inappropriate inactivation of tumor suppressor genes.

Occasionally, some of the reaction intermediates might also deaminate, generating the corresponding halouracil analogues. We observe that ClU, but not BrU, enhances MBD binding. Fortunately, the glycosylase activity of MBD4, a putative repair enzyme to remove thymine derived from the deamination of 5mC, has the greatest efficiency against ClU of all pyrimidines tested to date. These results suggest that MBD4 could serve to remove uracil, ClU, BrU and T, all of which would give rise to C to T transition mutations if unrepaired. The inflammation-mediated halogenation of cytosine residues in DNA may account for several of the DNA alterations observed in human tumors, including both point mutations and alterations in methylation patterns.

ACKNOWLEDGEMENTS

REFERENCES

1. Lindahl,T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709–715.

2. Loeb,L.A. (1989) Endogenous carcinogenesis: molecular oncology into the twenty-first century—Presidential address. *Cancer Res.*, **49**, 5489–5496.

3. Mullaart,E., Lohman,P.H., Berends,F. and Vijg,J. (1990) DNA damage metabolism and aging. *Mutat. Res.*, **237**, 189–210.

4. Esteller,M. (2002) CpG island hypermethylation and tumor suppressor genes: a booming present, a bright future. *Oncogene*, **21**, 5427–5440.

5. Valinluck,V., Tsai,H-H., Rogstad,D.K., Burdzy,A., Bird,A. and Sowers,L.C. (2004) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of MeCP2. *Nucleic Acids Res.*, **32**, 4100–4108.

6. Watanabe,S., Ichimura,T., Fujita,N., Tsuruzoe,S., Ohki,I., Shirakawa,M., Kawasuji,M. and Nakao,M. (2003) Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin. *Proc. Natl Acad. Sci. USA*, **100**, 12859–12864.

7. Singer,B. (1975) The chemical effects of nucleic acid alkylation and their relation to mutagenesis and carcinogenesis. *Prog. Nucleic Acid Res. Mol. Biol.*, **15**, 219–284.

8. Brennan,C.A., Van Cleve,M.D. and Gumport,R.I. (1986) The effects of base analogue substitutions on the methylation by the EcoRI modification methylase of octadeoxyribonucleotides containing modified EcoRI recognition sequences. *J. Biol. Chem.*, **261**, 7279–7286.

9. Brennan,C.A., Van Cleve,M.D. and Gumport,R.I. (1986) The effects of base analogue substitutions on the cleavage by the EcoRI restriction endonuclease of octadeoxyribonucleotides containing modified EcoRI recognition sequences. *J. Biol. Chem.*, **261**, 7270–7278.

10. Henderson,J.P., Byun,J., Takeshita,J. and Heinecke,J.W. (2003) Phagocytes produce 5-chlorouracil and 5-bromouracil, two mutagenic products of myeloperoxidase, in human inflammatory tissue. *J. Biol. Chem.*, **278**, 23522–23528.

11. Jiang,Q., Blount,B.C. and Ames,B.N. (2003) 5-Chlorouracil, a marker of DNA damage from hypochlorous acid during inflammation. *J. Biol. Chem.*, **278**, 32834–32840.

12. Henderson,J.P., Byun,J., Mueller,D.M. and Heinecke,J.W. (2001) The eosinophil peroxidase-hydrogen peroxide-bromide system of human eosinophils generates 5-bromouracil, a mutagenic thymine analogue. *Biochemistry*, **40**, 2052–2059.

13. Harrison,J.E. and Schultz,J. (1976) Studies on the chlorinating activity of myeloperoxidase. *J. Biol. Chem.*, **251**, 1371–1374.

14. Foote,C.S., Goyne,T.E. and Lehrer,R.I. (1983) Assessment of chlorination by human neutrophils. *Nature*, **301**, 715–716.

15. Henderson,J.P., Byun,J., Williams,M.V., Mueller,D.M., McCormick,M.L. and Heinecke,J.W. (2001) Production of brominating intermediates by myeloperoxidase. A transhalogenation pathway for generating mutagenic nucleobases during inflammation. *J. Biol. Chem.*, **276**, 7867–7875.

16. Weiss,S.J., Test,S.T., Eckmann,C.M., Roos,D. and Regiani,S. (1986) Brominating oxidants generated by human eosinophils. *Science*, **234**, 200–203.

17. Mayeno,A.N., Curran,A.J., Roberts,R.L. and Foote,C.S. (1989) Eosinophils preferentially use bromide to generate halogenating agents. *J. Biol. Chem.*, **264**, 5660–5668.

18. Thomas,E.L., Bozeman,P.M., Jefferson,M.M. and King,C.C. (1995) Oxidation of bromide by the human leukocyte enzymes myeloperoxidase and eosinophil peroxidase. Formation of bromamines. *J. Biol. Chem.*, **270**, 2906–2913.

19. Whiteman,M., Jenner,A. and Halliwell,B. (1997) Hypochlorous acid-induced base modifications in isolated calf thymus DNA. *Chem. Res. Toxicol.*, **10**, 1240–1246.

20. Whiteman,M., Spencer,J.P., Jenner,A. and Halliwell,B. (1999) Hypochlorous acid-induced base modification: potentiation by nitrite: biomarkers of DNA damage by reactive oxygen species. *Biochem. Biophys. Res. Commun.*, **257**, 572–576.

21. Nan,X., Ng,H.H., Johnson,C.A., Laherty,C.D., Turner,B.M., Eisenmann,R.N. and Bird,A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involved a histone deacetylase complex. *Nature*, **393**, 386–389.

22. Nan,X., Meehan,R.R. and Bird,A.P. (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res.*, **21**, 4886–4892.

23. Hendrich,B. and Bird,A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding protiens. *Mol. Cell Biol.*, **18**, 6538–6547.

24. Hendrich,B., Hardeland,U., Ng,H.H., Jiricny,J. and Bird,A. (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature*, **401**, 301–304.

25. Petronzelli,F., Riccio,A., Markham,G.D., Seeholzer,S.H., Stoerker,J., Genuardi,M., Yeung,A.T., Matsumoto,Y. and Bellacosa,A. (2000) Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase. *J. Biol. Chem.*, **275**, 32422–32429.

26. Parsons,B.L. (2003) MED1: a central molecule for maintenance of genome integrity and response to DNA damage. *Proc. Natl Acad. Sci. USA*, **100**, 14601–14602.

27. Kang,J.I. Jr, Burdzy,A., Liu,P. and Sowers,L.C. (2004) Synthesis and characterization of oligonucleotides containing 5-chlorocytosine. *Chem. Res. Toxicol.*, **17**, 1236–1244.

28. Rutter,M., Saunders,B., Wilkinson,K., Rumbles,S., Schofield,G., Kamm,M., Williams,C., Price,A., Talbot,I. and Forbes,A. (2004) Severity of inflammation is a risk factor for colorectal neoplasia in ulcerative colitis. *Gastroenterology*, **126**, 451–459.

29. Ballaz,S. and Mulshine,J.L. (2003) The potential contributions of chronic inflammation to lung carcinogenesis. *Clin. Lung Cancer*, **5**, 46–62.

30. Ohshima,H., Tatemichi,M. and Sawa,T. (2003) Chemical basis of inflammation-induced carcinogenesis. *Arch. Biochem. Biophys.*, **417**, 3–11.

31. Nelson,W.G., De Marzo,A.M. and Isaacs,W.B. (2003) Prostate cancer. *N. Engl. J. Med.*, **349**, 366–381.

32. Nielsen,J., Taagaard,M., Marugg,J.E., van Boom,J.H. and Dahl,O. (1986) Application of 2-cyanoethyl-N,N,N′,N′-tetraisopropylphosphoramidite for *in situ* preparation of deoxyribonucleoside phosphoramidites and their use in polymer supported synthesis of oligodeoxyribonucleosides. *Nucleic Acids Res.*, **14**, 7391–7403.

33. Free,A., Wakefield,R.I., Smith,B.O., Dryden,D.T., Barlow,P.N. and Bird,A.P. (2001) DNA recognition by the methyl-CpG binding domain of MeCP2. *J. Biol. Chem.*, **276**, 3353–3360.

34. Baker,D., Liu,P., Burdzy,A. and Sowers,L.C. (2002) Characterization of the substrate specificity of a human 5-hydroxymethyluracil glycosylase activity. *Chem. Res. Toxicol.*, **15**, 33–39.

35. Brandon,M.L., Mi,L., Chaung,W., Teebor,G. and Boorstein,R.J. (2000) 5-Chloro-2′-deoxyuridine cytotoxicity from base excision repair of uracil subsequent to thymidylate synthase inhibition. *Mutat. Res.*, **459**, 161–169.

36. Wakefield,R.I., Smith,B.O., Nan,X., Free,A., Soteriou,A., Uhrin,D., Bird,A.P. and Barlow,P.N. (1999) The solution structure of the domain from MeCP2 that binds to methylated DNA. *J. Mol. Biol.*, **291**, 1055–1065.

37. Ohki,I., Shimotake,N., Fujita,N., Jee,J., Ikegami,T., Nakao,M. and Shirakawa,M. (2001) Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell*, **105**, 487–497.

38. Jackson,A.L. and Loeb,L.A. (1998) The mutation rate and cancer. *Genetics*, **148**, 1483–1490.
39. Kinzler,K.W. and Vogelstein,B. (1996) Lessons from hereditary colorectal cancer. *Cell*, **87**, 159–170.
40. Lin,X., Tascilar,M., Lee,W.H., Vles,W.J., Lee,B.H., Veeraswamy,R., Asgari,K., Freije,D., van Rees,B., Gage,W.R. *et al.* (2001) GSTP1 CpG island hypermethylation is responsible for the absence of GSTP1 expression in human prostate cancer cells. *Am. J. Pathol.*, **159**, 1815–1826.
41. Lee,W.H., Morton,R.A., Epstein,J.I., Brooks,J.D., Campbell,P.A., Bova,G.S., Hsieh,W.S., Isaacs,W.B. and Nelson,W.G. (1994) Cytidine methylation of regulatory sequences near the pi-class glutathione S-transferase gene accompanies human prostatic carcinogenesis. *Proc. Natl Acad. Sci. USA*, **91**, 11733–11737.
42. Baylin,S. and Bestor,T.H. (2002) Altered methylation patterns in cancer cell genomes: cause or consequence? *Cancer Cell*, **1**, 299–305.
43. Hardy,T.A., Baker,D.J., Newman,E.M., Sowers,L.C., Goodman,M.F. and Smith,S.S. (1987) Size of the directing moiety at carbon 5 of cytosine and the activity of human DNA(cytosine-5) methyltransferase. *Biochem. Biophys. Res. Commun.*, **145**, 146–152.
44. Duncan,B.K. and Miller,J.H. (1980) Mutagenic deamination of cytosine residues in DNA. *Nature*, **287**, 560–561.
45. Sved,J. and Bird,A. (1990) The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc. Natl Acad. Sci. USA*, **87**, 4692–4696.
46. Liu,P., Burdzy,A. and Sowers,L.C. (2003) Repair of the mutagenic DNA oxidation product, 5-formyluracil. *DNA Repair*, **2**, 199–210.
47. Liu,P., Burdzy,A. and Sowers,L.C. (2002) Substrate recognition by a family of uracil-DNA glycosylases: UNG, MUG, and TDG. *Chem. Res. Toxicol.*, **15**, 1000–1009.
48. Jones,P.L. and Wolffe,A.P. (1999) Relationships between chromatin organization and DNA methylation in determining gene expression. *Sem. Cancer Biol.*, **9**, 339–347.
49. Boyes,J. and Bird,A. (1991) DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell*, **64**, 1123–1134.
50. Antequera,F., Boyes,J. and Bird,A. (1990) High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. *Cell*, **62**, 503–514.
51. Kimura,H. and Shiota,K. (2003) Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. *J. Biol. Chem.*, **278**, 4806–4812.