Human DNA (cytosine-5)methyltransferase selectively methylates duplex DNA containing mispairs

Steven S.Smith*, Thomas A.Hardy and David J.Baker

Division of Surgery, City of Hope National Medical Center, 1500 E.Duarte Rd, Duarte, CA 91010, USA

Received May 8, 1987; Revised and Accepted August 6, 1987

ABSTRACT

The presence of the C·C mispair in a defined duplex oligodeoxynucleotide enhanced its capacity to serve as a substrate for highly purified human DNA methyltransferase. Analysis of tritiated reaction products showed that the C·C mispair acted as a "methylation ₅acceptor" in that it was itself rapidly methylated. The m⁵C·G base pair also enhanced the capacity of the oligodeoxynucleotide to serve as a substrate for the enzyme. However, this complementary base pair was found to act as a "methylation director". That is, the presence of the m⁵C in one strand induced the enzyme to rapidly methylate at the cytosine residue on the opposite strand in an adjacent C·G base pair.

INTRODUCTION

Methylation of cytosine residues at critical CG sites in DNA can prevent the expression of many mammalian genes (1-3). It seems likely that responsive genes methylated at these critical sites will be passed from one somatic cell generation to the next in the transcriptionally inactive state because patterns of cytosine methylation at CG dimers are somatically inherited (4). Clonal gene silencing by DNA methylation may have relevance in several important areas including X-inactivation (5), viral latency (6,7), viral gene expression (8), oncogene expression (9) and carcinogenesis (10).

A role for gene silencing by DNA methylation in the orchestrated process of mammalian development is supported by several lines of experimentation (11). However, this hypothesis must be reconciled with the apparent lack of DNA methylation in Drosophila (12) and Dictyostelium (Smith unpublished). Both of these organisms appear to play out complex developmental programs without appreciable DNA methylation.

Hare and Taylor (13) have obtained evidence suggesting that methylation patterns may provide a signal for strand discrimination during mismatch repair in the vertebrates. We have recently reported (14) that the human DNA (cytosine-5)methyltransferase (E.C.2.1.1.37) recognizes and methylates oligodeoxynucleotides containing a T·G mispair at a moderate rate. Here we report the ability of this same enzyme to selectively methylate duplex DNA structures induced by a mismatched C·C base pair at an enhanced rate.

MATERIALS AND METHODS

Enzymes and Chemicals: S-adenosyl-L-[methyl-³H]methionine ([³H]AdoMet) was obtained from Amersham with a specific activity of 15 Ci/mmol or from Research Products International with a specific activity of 22 Ci/mmol. Dithiothreitol, and 2-mercaptoethanol, were from Sigma Chemical Co. HpaII, MspI, and Sau3AI were from New England Biolabs. Other restriction enzymes and T4 polynucleotide kinase were from Bethesda Research Laboratories.

Purification: DNA(cytosine-5)methyltransferase was purified from human placentas as previously described (15), except that nucleic acid was removed from hydroxylapatite-concentrated material by passage through DEAE-cellulose prior to Dodecyl-Sepharose column chromatography. The purified enzyme was stable for at least 3 months at -70° C.

Enzymatic Methylation: In previous work (15,16) with synthetic M13 DNAs, we studied the effects of salt and AdoMet concentrations on enzyme reaction rate. Both unmethylated and hemimethylated duplex DNAs gave broad optima with sodium chloride under the conditions tested (4 ug/ml DNA and 1.4 uM AdoMet). Hyperbolic kinetics were observed for AdoMet with $S_{0.5}$ (substrate concentration at which the reaction velocity is half maximal) values of 5 uM when synthetic duplex DNA (8 ug/ml) was employed, or 4 uM when hemimethylated DNA (also 8 ug/ml) was employed. The apparent maximal velocities for AdoMet with these DNAs were 0.26 pmole/min and 1.6 pmole/min. For the present work, the most important conclusion drawn from those studies was that the ratio of the activities on the two DNA substrates was

roughly constant as a function of salt or AdoMet concentration. In the studies below 50 mM NaCl was used since this was optimal.

Since it is likely that most enzymes operate in the vicinity of their $S_{0.5}$ values for substrates (17), the AdoMet concentrations used here were chosen so that they were in the vicinity of the observed $S_{0.5}$ values.

Initial velocity determinations were made using 250 units/ml of fraction VI enzyme with a minimum specific activity of 12,000 units/mg, 50mM NaCl, and 2.0 μ M [3 H] AdoMet. A unit of activity is that amount of enzyme that will catalyze the incorporation of 1 pmole of methyl groups into trichloroacetic acid insoluble material in one hour at 37°C in a 100 μ l reaction mixture containing: 150 μ g/ml single-stranded M. lysodeikticus DNA, 50 mM Tris-Cl, pH 7.8; 50mM NaCl; 10% vol/vol glycerol; 1.4 μ M [3 H]AdoMet; 10 mM EDTA; 10 μ g/ml RNase; and 2 mM dithiothreitol. Omission of the RNase and EDTA did not alter the observed rates with the highly purified enzyme used here. Reaction rates with oligonucleotides were found to be linear for about 30 minutes. Initial velocities were estimated from a single time point taken at 20 minutes.

When restriction analysis was used to localize label on the reaction products formed by incubation of duplex oligodeoxynucleotides with human DNA methyltransferase and [3 H]AdoMet, reactions were allowed to continue for 45 minutes at 37°C. Best results were obtained with the standard reaction mixture (15) containing 4 to 8 μ g/ml DNA, 6 μ M [3 H]AdoMet and 500 U/ml methyltransferase in a final reaction volume of 1.0 ml. The increased AdoMet and enzyme concentrations permitted higher levels of incorporation without altering the relative kinetic ordering for substrate observed in initial velocity studies. Reactions were initiated with enzyme after a 15 minute preincubation at 37°C. They were terminated by chilling on ice, and the DNA was precipitated with ethanol using 1.0 μ g/ml glycogen as carrier.

Oligonucleotide Design: Our purpose in designing the oligonucleotides used in these studies was to construct GC-rich duplexes that could be used as vehicles for the determination of primary sites of methylation in the vicinity of complementary

and mispaired bases. Several factors were considered in the design of oligonucleotides. First they are 30-mers containing two CG sites spaced about 15 bp apart so that they meet criteria established by Bolden et al. (18) for DNA methyltransferase substrates. In addition, one of the CG sites was placed in a CCGG tetramer in order to permit conclusions regarding cleavage by the isoschizomers HpaII and MspI. And last, the overall molecule was made GC-rich for two reasons. First, the observation of short GC-rich sequences in the control regions of developmentally regulated genes (1) makes such sequences inherently interesting in studies of DNA methylation. Second, the inherent stability of GC-rich DNA increases the likelihood that the presence of mispaired bases will permit the formation of stable helical structures.

Our aim was to use short reaction times for labelling studies, and initial velocities for kinetic determinations in order to determine the primary specificity of the enzyme. The advantage of this approach is that secondary reactions that depend on weaker affinities or on the accumulation of significant amounts of reaction products are avoided. The disadvantage of the approach is that direct sequencing (19-20) can not be used to determine the position of the methyl groups applied. since even primary sites are only partially methylated. The analysis of the labelling specificity depends on restriction enzyme cleavage sites introduced into the oligonucleotides that permit the assignment of labelled sites in the DNA. Thus the HaeIII cleavage site allows the inspection of the right end of the molecule, while the MboI and MspI sites allow the inspection of the center and left end of the molecule.

Oligodeoxynucleotide synthesis and characterization:

Oligodeoxynucleotides were prepared by the Synthesis Facility of the City of Hope Cancer Center using the phosphoramidite method (21) and were purified by HPLC and polyacrylamide gel electrophoresis (22). The sequence of each strand used was verified in this laboratory using T4 polynucleotide kinase end-labelling and Maxam-Gilbert (20) sequencing.

To form duplex molecules, molar equivalencies were estimated from absorbance at 260 nm. Strands were annealed by heating

a solution containing 0.1 ug/ul of each complementary strand to 95° C for 5 minutes in: 10 mM Tris, 1 mM EDTA, 100mM NaCl. Subsequent overnight incubation at 50° C was followed by cooling to room temperature and then to 0°C prior to use. To confirm the structures in solution, single strands were end-labelled with T4 polynucleotide kinase. Duplexes were formed by annealing as above. To confirm the structures, ³²P end-labelled duplexes were digested to completion with HpaII, MspI, HaeIII, MboI, Sau3AI, and EcoRII, and the products were sized on non-denaturing 12% acrylamide gels (23).

Analysis of Enzymatically Tritiated DNAs: After ethanol precipitation as described above, between 10,000 and 25,000 cpm of labelled product was resuspended in restrictionbuffer. Good results were obtained with final DNA concentrations between 0.02 and 0.1 ug/ul. Restriction was for two hours at 37°C using 30 U enzyme/ug DNA. After restriction, 10 ul of the digestion product was mixed with denaturation buffer, heat denatured and separated by electrophoresis through 20% acrylamide DNA sequencing gels (20). Labelled DNA fragments were visualized by low temperature autoradiography after treating the gels with "En³Hance" (NEN Research Products) using the method described by the manufacturer. Autoradiographs were quantified using gel scans produced and integrated with a Gilford Response densitometer.

RESULTS

Characterization of component single strands: The central 27 bases of each sequence used in the study (including the positions of m⁵C) residues were confirmed by Maxam-Gilbert sequencing (19,20).

Characterization of duplex substrates: To test for the formation of the appropriate duplex structures, component single strands were end-labelled and annealed under conditions expected to give stoichiometric formation of duplex 30-mers. The products were separated by gel electrophoresis under non-denaturing conditions (not shown). Unannealed 30-mer runs well ahead of the duplex 30-mer in these gels. In each case more than 95% of the label ran at the position of the duplex with only a trace of

Table I. Sequences and Properties of the Duplex Oligonucleotides

SEQUENCE AND NAME OF DUPLEX	V _i	DUPLEX CUT BY				
		Hp	Mp	He	Sa	RII
<u>unmethylated complementary duplex</u>						
5'GTCCACCAGATCCGGGCTACCTGGCCTCGA3'	2.9	+	+	+	+	+
3'CAGGTGGTCTAGGCCCGATGGACCGGAGCT5'						
<u>methylated complementary duplex</u>						
M 5'GTCCACCAGATCCGGGCTACCTGGCCTCGA3'	0.7	-	+	+	+	+
3'CAGGTGGTCTAGGCCCGATGGACCGGAGCT5' M						
<u>hemimethylated complementary duplex</u>						
M 5'GTCCACCAGATCCGGGCTACCTGGCCTCGA3'	132.0	-	+	+	+	+
3'CAGGTGGTCTAGGCCCGATGGACCGGAGCT5' *						
<u>unmethylated mispaired (C·C) duplex</u>						
5'GTCCACCAGATCCCGGCTACCTGGCCTCGA3'	20.5	-	+/-	+	+	+
3'CAGGTGGTCTAGGCCCGATGGACCGGAGCT5' *						

Sites of methylation that can be unequivocally assigned from the autoradiographic and kinetic data are marked with an asterisk. Probable methylation sites deduced from the data are underlined. Average initial velocities (V_i) from five independent determinations for each substrate are given in fmole/min incorporated at 4 ug/ml oligodeoxynucleotide DNA. Complete cleavage of a given duplex by a given restriction enzyme is shown with a (+) lack of cleavage is indicated by a (-) and partial cleavage by a (+/-). HpaII (Hp); MspI (Mp); HaeIII (He); Sau3AI (Sa); EcoRII (RII).

unannealed 30-mer evident in autoradiographs.

Each duplex was cleaved with a battery of restriction enzymes, and the products were separated using non-denaturing gels. The results are summarized in Table I. HaeIII, Sau3AI and EcoRII were found to cleave all of the oligonucleotides in the expected fashion. HpaII, which is sensitive to methylation in DNA at the central cytosine in the CCGG tetramer, was able to cleave the unmethylated complementary duplex but was unable to

Fig. 1. Localization of Methyl Groups on Hemimethylated and Mismatched Duplexes

Oligodeoxyribonucleotides labelled by the DNA methyltransferase were cleaved with the restriction enzymes shown, then denatured and separated by electrophoresis through 20% acrylamide DNA sequencing gels. Representative low temperature autoradiographs of the labelled restriction fragments are shown.

cleave the fully methylated or the hemimethylated complementary duplexes. MspI, which is insensitive to methylation at that same site, cleaved all of these same duplexes. Although the duplex containing the C·C mispair was not cleaved by HpaII, partial cleavage was obtained with MspI.

Methylation of Complementary Oligodeoxynucleotide Duplexes: Initial velocities determined for the complementary duplexes are given in Table I. Measured incorporations in the initial velocity studies ranged between 120 and 22,000 cpm above background. The data give the following ordering for the substrates studied here: methylated complementary duplex < unmethylated

Sequence of Duplex Fragments		Enzyme Used
<u>Hemimethylated complementary duplex</u>		
"12" M "18"		
5'GTCCACCAGATC	CGGGCTACCTGGCCTCGA3'	
"14" "16"		
3'CAGGTGGTCTAGGC	CCGATGGACCGGAGCT5'	<u>MspI</u>
	*	
"8" M "22"		
5'GTCCACCA	GATCCGGGCTACCTGGCCTCGA3'	
"12" "18"		
3'CAGGTGGTCTAG	GCCCGATGGACCGGAGCT5'	<u>MboI</u>
	*	
"24" M "6"		
5'GTCCACCAGATCCGGGCTACCTGG	CCTCGA3'	
"24" "6"		
3'CAGGTGGTCTAGGCCCGGATGGACC	GGAGCT5'	<u>HaeIII</u>
	*	
<u>Unmethylated C·C mispaired duplex</u>		
"8" "22"		
5'GTCCACCA	GATCCGGGCTACCTGGCCTCGA3'	
"12" "18"		
3'CAGGTGGTCTAG	GCCCGATGGACCGGAGCT5'	<u>Mbo I</u>
"13" "17"		
5'GTCCACCAGATCC	CGGCTACCTGGCCTCGA3'	
"14" "16"		
3'CAGGTGGTCTAGGC	CCGATGGACCGGAGCT5'	<u>Msp I</u>

Fig. 2. Schematic of Restriction Analysis of Labelled Oligonucleotides
The sequences of the hemimethylated complementary duplex and the unmethylated duplex containing a C·C mispair are shown along with exploded views of the fragments expected from the duplex after cleavage by Msp I and Mbo I followed by denaturation. Sites of methylation that can be unequivocally assigned are marked with an asterisk, while other probable methylation sites deduced from the data are underlined.

complementary duplex< unmethylated mispaired duplex< hemi-methylated complementary duplex.

A direct assignment of the methylated cytosine(s) in a given oligonucleotide can be made using autoradiographs of restriction fragments run on 20% sequencing gels. Autoradiography of MspI digests of the hemimethylated duplex after in vitro labelling showed that of the four possible products, only the 14-mer was labelled. MboI digests of the same product revealed that only the 18-mer was labelled (see Fig. 1). HaeIII

Fig. 3. Localization of Methyl groups on Unmethylated and Mismatched Duplexes.

Enzymatically labelled duplexes were cleaved with the indicated restriction enzymes, denatured, and separated on 20% sequencing gels. Molecular length markers are shown on the left of each array. The labelled fragments have the indicated single stranded molecular lengths.

cleavage produced only a labelled 24-mer (not shown).

The expected restriction fragments for the hemimethylated duplex are listed in Fig. 2. Since only the cytosine underlined in Fig. 2 is common to both the 14-mer and the 18-mer the experiment allows the unequivocal assignment of the label to that unique cytosine in this example.

Labelling patterns were not determined for the methylated duplex nor were they determined for any of the single-stranded oligodeoxynucleotides, since reaction rates with these substrates were extremely slow. Nevertheless, ^3H incorporation into the DNA of the methylated duplex was detected as autoradiographic signal running with the denatured 30-mers in sequencing

Table II. Restriction Fragments Generated From Labelled Duplex Oligonucleotides

Duplex	Relative Intensity of Methyltransferase Labelling									
	<u>Msp</u> I						<u>Mbo</u> I			
	Fragment Length						Fragment Length			
	12	13	14	16	17	18	8	12	18	22
<u>unmethylated</u> <u>complementary</u> <u>duplex</u>	13	0	37	14	0	36	<2	0	40	58
<u>hemimethylated</u> <u>matched duplex</u>	0	0	100	0	0	0	0	0	100	0
<u>unmethylated</u> <u>mispaird (C·C)</u> <u>duplex</u>	0	0	73	0	27	0	0	0	62	38

The length and relative intensity for Msp I and Mbo I fragments observed in autoradiographs like those shown in Figs. 1 and 3 are listed for each duplex structure tested. In each autoradiograph approximately the same amount of total radioactivity was applied. The data given show how this label was distributed among the fragments produced from a single labelling experiment with the total label found in fragments taken as 100%.

gels (not shown) indicating bona fide labelling of the DNA. An autoradiograph showing the labelling pattern obtained with the unmethylated complementary duplex is shown in Fig. 3.

Only three of the four possible fragments produced by MboI digestion of the unmethylated complementary duplex were found to be tritium labelled: a 22-mer, an 18-mer, and a an 8-mer. The 22-mer, and 18-mer were labelled much more intensely than the 8-mer (Fig. 3, Table II). All four possible fragments produced by MspI digestion were found to be tritium labelled. However, the 18-mer and the 14-mer were more intensely labelled than the 16-mer and the 12-mer (Fig. 3 and Table II).

Methylation at the C·C mispair: The unmethylated duplex containing a C·C mispair was methylated more than 7 fold faster than the unmethylated complementary duplex. When the enzymatically methylated product was cleaved by MspI, tritium was detected in a 14-mer and in a 17-mer. MboI cleavage produced an 18-mer, and a 22-mer (see Figs.1-3 and Table II). HaeIII

Fig. 4. Mismatch induced alteration in MspI cleavage. Duplex DNA that had been 32 P labelled at the 5' ends of each strand was cleaved with MspI, denatured and separated on sequencing gels. The molecular lengths of the digestion products observed with the mismatched duplex (left) and the hemimethylated duplex (right) indicate cleavage by MspI at the sites indicated by the arrows.

cleavage produced labelled 24-mer only (Fig.3). No label was detected in the 6-mers that contain the CG dimer near the end of the duplex structure.

A comparison of the autoradiographic signal in fragments produced with MboI and HaeIII from the mispaired duplex is given in Fig. 3. In that experiment approximately equal amounts of radiolabel were loaded in each lane. The scanned area under the peaks corresponding to the MboI 18-mer and the MboI 22-mer were found to be equal to one another indicating symmetrical labelling on both strands of the duplex. Further, the area under the peak corresponding to the HaeIII 24-mer was almost exactly equal to the sum of these two areas, indicating labelling of each of the two 24-mers and negligible labelling of the 6-mers produced by this enzyme.

The observation of a 17-mer in MspI digests suggests cleavage between the two cytosines in the mispaired 30-mer at positions 13 and 14 (from the 5' end on the "top" strand shown

in Fig. 2) to produce a 13-mer and a 17-mer from the parent 30-mer. To check this assignment oligonucleotides used to form the hemimethylated complementary duplex, and the unmethylated duplex containing the C·C mispair were end-labelled prior to restriction with MspI. After denaturation and separation on a sequencing gel, the fragments were visualized by autoradiography. A 16-mer and a 12-mer were produced by MspI cleavage of the complementary duplex, while a 16-mer and 13-mer were produced from the mispaired duplex (see Fig. 4).

DISCUSSION

Methylation of Duplex DNA: In evaluating the enzymatic methylation of DNA it is useful to distinguish two types of base pair in substrate DNA molecules. If a base pair enhances the methylation of an adjacent base pair but is not itself methylated during the reaction catalyzed by the DNA methyltransferase, then it can be termed a director or directing base pair. If the base pair is methylated, then it can be termed an acceptor or accepting base pair. In the most frequently studied reaction (5,16,18, and 24), a hemimethylated CG dinucleotide in duplex DNA is converted to a fully methylated CG dinucleotide as shown in (I). In this simple example, the base pair on the left is identified as the directing base pair, and that on the right is identified as the accepting base pair.

Specificity of the Purified DNA Methyltransferase: The enzyme studied here exhibited a primary specificity for the CG dimer. Since prior methylation of the cytosines in both strands at the central CG dimer effectively reduced the substrate activity to near background levels, the preferred sites of methylation are likely to be these two cytosine residues. All of the evidence presented here suggests that each of the two C·G base pairs that form the central CG site act as equivalent

methylation acceptors even though the cytosine residues are on different strands.

The enzyme does, however, exhibit a weak tendency to label cytosines (C·G base pairs) outside this dimer in the absence of hemimethylated or mispaired cytosines. In studies of the hemimethylated duplex, labelling of either cytosine at the second CG site near the right end of the molecule is excluded by the absence of label in the 16-mer or the 18-mer produced by MspI. In studies of both the hemimethylated and mispaired duplexes the absence of label in either of the two 6-mers produced by HaeIII also rules out labelling at that CG site. Thus it appears that the enzyme has a low affinity for this particular CG site, perhaps because it is near the end of the duplex.

The presence of a single m^5C on one strand at the central CG site had two apparent effects. First, it enhanced the innate selectivity of the enzyme for the centrally located CG site by directing methylation to the unmethylated cytosine in the dimer as shown in (I) above. Second it enhanced the reaction rate by nearly 46 fold over the unmethylated homologue, or 200 fold over the doubly methylated homologue. These findings are in agreement with existing evidence on the specificity and selectivity of the mammalian enzymes (15,18,24) and they provide a framework for the interpretation of the more novel results below.

Trinucleotide Methylation: Methylation at the CNG trinucleotide (where N can be any nucleotide) has been observed in human and plant DNA (7, 25). Woodcock et al. (26) have proposed the existence of a human enzyme with this specificity. The enzyme studied here can not have this as a primary specificity since most of the enzymatically labelled products can be cleaved by MspI which does not cleave when the outside C residue (on either strand) in the CCGG site is methylated. Further, the absence of label in the 16-mer produced by MspI digestion of enzymatically labelled hemimethylated DNA (Fig.1) clearly excludes one type of specificity for a hemimethylated trinucleotide with the present enzyme, since the trinucleotide product:

was not observed.

Fig. 5. Possible structures for the oligodeoxynucleotide containing a C·C mispair.

A.) A fully base paired structure in which the C·C pair is expected to be protonated. B.) A structure in which the CCGG site is regenerated by stacking an unpaired cytosine residue into symmetrical positions in each single strand. C.) A structure similar to that in B except that the cytosines are looped out of the helix instead of stacked into it. D.) A structure containing a region of local denaturation, that would permit the attack of single strands by *MspI*. The anomalous cytosine residues that are mispaired in structure A and otherwise accommodated in structures B through D are underlined.

Methylation at the C·C mispair in duplex DNA: The sites of methylation indicated in Table I for the duplex containing the C·C mispair are consistent with the observed labelling of restriction fragments (Figs. 1-3). Since the assignment of the labelled cytosine on the "bottom" strand is unequivocal, the C:C mispair acts as a methylation acceptor. The assignment on the "top" strand is not unequivocal. The absence of label from either of the 6-mers shows that the three cytosines at the right end of the fragment are not labelled. This finding combined with the lack of labelling in the 16-mer produced by *MspI* shows that labelling does not occur at the CG site near the right terminus of the oligonucleotide. Even so, we can not formally rule out labelling at any of the first four cytosines in the 17-mer produced by *MspI* (see Fig. 2).

The mechanism by which the C·C mispair enhances the rate of methylation remains to be elucidated. Circular dichroism has recently been used to show that a protonated C·C base pair can form between antiparallel DNA strands (27). This finding led to

the suggestion (27) that this type of base pair may in fact be accommodated within a standard Watson-Crick helix without the unfavorable syn conformation. If the structure in Fig. 5A is formed, then the mispair is expected to be protonated. Since the proposed enzyme mechanism involves nucleophilic attack on the cytosine ring by a group on the enzyme (28); the presence of the protonated C·C base pair could provide an explanation for the enhanced substrate activity.

However the generation of a 13-mer and a 17-mer by MspI cleavage suggests that there may be an equilibrium in solution between one or more of the alternative substrate structures shown in Fig. 5. In the structure shown in Fig. 5B, the unpaired cytosines are stacked into the helix at symmetrically placed sites on each strand. One example of a structure of this type is known and has recently been crystallized (29). The CCGG recognition site is regenerated in the structures given in Fig. 5B and Fig. 5C. MspI cleavage would produce the observed fragments from either structure, but it is not clear why HpaII would not be able to cleave at the same site. A looped structure like that shown in Fig. 5D is not energetically favorable. If such a structure were to form transiently, it could explain the differential cleavage by HpaII and MspI, since MspI (30) does cleave single strands with low efficiency while HpaII does not (31).

In any case this finding raises an interesting question regarding the recognition of the site by the DNA methyltransferase. Recognition of structure A would be consistent with an enhancement of the hypothetical (28) nucleophilic attack by the proton on the C·C mispair. Recognition of structure B would suggest that methylation of a normally paired CG dimer in DNA is promoted by an anomalous structure or perhaps an abasic site in its vicinity.

Although the mechanism by which the C·C mispair enhances the reaction rate may be multifaceted, the enhancement itself is particularly interesting in light of the results with restriction enzyme cleavage. Thus recognition by HpaII was abolished, and recognition by MspI was diminished by the same mispair-induced structure(s) that enhance methylation by the human DNA methyltransferase.

The exact structural preferences exhibited by the enzyme remain to be worked out. However, it is clear that the DNA methyltransferase responds to secondary structure in DNA in addition to methylation state. Preliminary indications are that the catalogue of structural anomalies that are recognized by the enzyme may be quite large. In addition to the rate enhancement associated with the C·C mispair, poly(dIdC):poly(dIdC) has been reported to enhance the rate of methylation by the human enzyme over that observed with poly(dGdC)·poly(dGdC), possibly as a consequence of its "more open structure" (32). We have recently found that the cytosine in the A·C mismatch is also methylated by the enzyme with nearly a ten fold rate enhancement (in preparation), while the T·G mispair induces the enzyme to label an adjacent cytosine in the strand opposite the T in the mispair (14). Since the response of the enzyme depends on the nature of the anomaly the potential information content of the process is high.

Structural perturbations like those studied thus far are likely to be associated with a variety of naturally occurring phenomena. These phenomena include the formation of replication and recombination intermediates in DNA, in addition to the occurrence of replication errors or damage in DNA. Whether or not de novo methylation of cytosines near such structural perturbations in DNA occurs in vivo remains to be seen. It is, however, consistent with recent reports of significant level of methylation at cytosines outside the CG dimer (7,26) in human DNA, and with a report (33) that a tandem duplication can provide a signal for DNA methylation in Neurospora.

ACKNOWLEDGEMENTS

We wish to thank Dr. Bruce Kaplan and his laboratory for help and advice in the synthesis and use of the deoxyoligonucleotides. And we would like to thank Drs. L. Sowers, M. Goodman, R.E. Dickerson, and J. Rossi, for helpful suggestions. This work was supported by grants from the Council for Tobacco Research (1571A), and the National Institutes of Health (GM-32863). S.S.S. is a member of the City of Hope Cancer Center CA-335172.

*To whom correspondence should be addressed

REFERENCES

- 1.) Bird, A.P. (1986) *Nature* 321, 209-213.
- 2.) Buslinger, M., Hurst, J. and Flavell, R.A. (1983) *Cell* 34, 197-206.
- 3.) Doerfler, W. (1983) *Ann. Rev. Biochem.* 52, 93-124.
- 4.) Wigler, M.H. (1981) *Cell* 24, 285-286.
- 5.) Riggs, A.D. (1975) *Cytogenet. Cell Genet.* 14, 19-25.
- 6.) Jaehner, D. and Jaenisch, R. (1985) *Nature* 315, 594-597.
- 7.) Bednarek, D.P., Mosca, J.D. and Raj, N.B.K. (1987) *J. Virol.* 61, 1253-1257.
- 8.) Sutter, D. and Doerfler, W. (1980) *Proc. Nat. Acad. Sci. U.S.A.* 77, 253-256.
- 9.) Jones, P.A. (1986) *Cancer Res.* 46, 461-466.
- 10.) Borello, M.G., Pierotti, M.A., Bongarzoni, I., Donghi, R., Mondellini, P., and Porta, G.D., (1987) *Cancer Res.* 47, 75-79.
- 11.) Razin, A. and Szyf, M. (1984) *Biochim. Biophys. Acta* 782, 331-342.
- 12.) Smith, S.S. and Thomas, C.A., Jr. (1981) *Gene* 13, 395-408.
- 13.) Hare, J.T., and Taylor, J.H., (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 7350-7354.
- 14.) Baker, D.J., Hardy, T.A., and Smith, S.S. (1987) *Biochem. Biophys. Res. Comm.* 146, 596-602.
- 15.) Zucker, K.E., Riggs, A.D. and Smith, S.S. (1985) *J. Cellular Biochem.* 29, 337-349.
- 16.) Hardy, T.A., Baker, D.J., Newman, E.M., Sowers, L.C., Goodman, M.F., and Smith, S.S. (1987) *Biochem. Biophys. Res. Comm.* 145, 146-152.
- 17.) Atkinson, D.E. (1977) *Cellular Energy Metabolism and its Regulation.* Academic Press, New York, New York. p.116.
- 18.) Bolden, A.H., Nalin, C.M., Ward, C.A., Poonian, M.S. and Weissbach A. (1986) *Molec. and Cellu. Biol.* 6, 1135-1140.
- 19.) Ohmori, H., Tomizawa, J. and Maxam, A. M. (1978) *Nucleic Acids Res.* 5, 1479-1485.
- 20.) Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* 74, 560-564.
- 21.) Sinha, N.D., Biernat, J., McManus, J., and Koester, H. (1984) *Nucleic Acids Res.* 12, 4539-4557.
- 22.) Tan, Z.-K., Ikuta, S., Huang, T., Dugaiczky, A., and Itakura, K. (1982) *Cold Spring Harbor Symposia on Quantitative Biology XLVII*, 383-391.
- 23.) Maniatis, T., Jeffery, A. and van deSande, H. (1975) *Biochemistry* 14, 3787-3794.
- 24.) Gruenbaum, Y., Cedar, H. and Razin, A. (1982) *Nature* 295, 620-622.
- 25.) Gruenbaum, Y., Naveh-Many, T., Cedar, H., and Razin, A. (1981) *Nature* 292, 860-862.
- 26.) Woodcock, D.M., Crowther, P.J., and Diver, W.P. (1987) *Biochem. Biophys. Res. Comm.* 145, 888-894.
- 27.) Gray, D.M., Cui, T. and Ratliff, R.L. (1984) *Nucleic Acids Res.* 12, 7565-7580.
- 28.) Santi, D.V., Garrett, C.E., and Barr, P.J. (1983) *Cell* 33, 9-10.

- 29.) Saper, M.A., Eldar, H., Mizuuchi, K, Nickol, J., Appella, E., and Sussman, J.L. (1986) J. Mol. Biol. 188, 111-113.
- 30.) Yoo, O.J. and Agarwal, K.L. (1980) J. Biol. Chem. 255, 10559-10562.
- 31.) Blakesley, R.W., and Wells, R.D., (1975) Nature 257, 421-422.
- 32.) Pedrali-Noy, G., and Weissbach, A. (1986) J. Biol. Chem. 261, 7600-7602.
- 33.) Selker, E.U. and Stevens, J.N. (1987) Molec. and Cellu. Biol. 7, 1032-1038.