How *M.Spl* and *M.HpaII* decide which base to methylate

Sha Mi and Richard J. Roberts*

Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA

Received June 24, 1992; Revised and Accepted August 20, 1992

ABSTRACT

The *HpaII* methylase (*M.HpaII*) recognizes the sequence CCGG and methylates the inner cytosine residue. The *MspI* methylase (*M.Spl*) recognizes the same sequence but methylates the outer cytosine residue. Both methylases have the usual architecture of 10 well-conserved motifs surrounding a variable region, responsible for sequence specific recognition, that is quite different in the two methylases. We have constructed hybrids between these two methylases and studied their methylation properties. A hybrid containing the variable region and C-terminal sequences from *M.Spl* methylates the outer cytosine residue. A second hybrid identical to the first except that the variable region derives from the *M.HpaII* methylates the inner cytosine residue. Thus the choice of base to be methylated within the recognition sequence is determined by the variable region.

INTRODUCTION

5-methylcytosine methylases (m5C-methylases) are found in organisms ranging from bacteria to mammals (1). Their function is to transfer a methyl group from S-adenosylmethionine (SAM) to the C5 position of a cytosine residue contained within a specific double stranded DNA sequence. Two classes of m5C-methylases are known: the mono-specific methylases, which recognize and modify a single recognition sequence and the multi-specific methylases, which recognize and modify several different sequences. In prokaryotic systems the mono-specific enzymes are usually involved in restriction/modification systems to protect host DNA from cleavage by the corresponding restriction endonucleases (2). However, the methylases are sometimes found in the absence of a companion restriction enzyme and their biological role in such circumstances is less clear. More than thirty m5C-methylases have been sequenced.

The mono-specific and multi-specific m5C-methylases share an overall common architecture, yet they also contain interesting differences. Five highly conserved sequence motifs about 10 to 20 amino acids long and five moderately conserved motifs have been observed (3,4). Motif I, close to the N-terminus, has been suggested to interact with S-adenosylmethionine based on its homology to similar regions in other families of protein, DNA and RNA methylases (5,6). Motif IV is the enzyme catalytic center, which contains a cysteine residue that is involved in

covalent bond formation with the target cytosine residue during catalysis. This mechanism was originally proposed by Santi (7), based upon similarities to the mechanism employed by thymidylate synthase, and has recently been proved directly (8). Specific DNA sequence recognition by the methylases is mediated by the so-called variable region which lies between conserved motifs VIII and IX. In the multi-specific methylases, mutations in the variable region can disable the methylation of one DNA sequence while permitting the methylation of others (9). In the mono-specific methylases, studies of hybrids between two methylases of different specificities, have demonstrated that the variable region is necessary and sufficient for sequence recognition (10).

The *MspI* and *HpaII* methylases both recognize the same DNA sequence CCGG. However, they methylate different cytosine residues within that sequence. *M.Spl* methylates the outer cytosine (11), while *M.HpaII* methylates the inner cytosine (12). Although both methylases share the common motifs that are found in all m5C-methylases they differ in the details of these motifs and also have quite dissimilar variable regions and N-terminal and C-terminal regions. To determine how these different methylases select which base is to be methylated we have constructed hybrids between *M.Spl* and *M.HpaII* in the hope of recovering functional methylases that would target the alternative cytosine residue. From the structure of such hybrids we have discovered which portion of the methylase is responsible for the choice of the base to be methylated.

MATERIALS AND METHODS

Bacterial strains and plasmids

Escherichia coli K-12 strains: ER1727 (*McrA*[−]*McrB*[−], containing F' *lacI*^q) (13), ER1700 (*McrA*⁺*McrB*[−]) (14), ER1969 (*McrA*[−]*McrB*⁺) (14) and ER1996 (*McrA*[−]*McrB*[−]) were kindly provided by Dr. E. Raleigh from New England Biolabs. ER1996 has a similar genotype to ER1969 except that it is *e14*[°] Δ (*mcrC-mrr*)114::IS10 and contains Δ (*argF-lac*)U169 rather than Δ (*lacZ*)r1. The protein expression plasmid pUHE25-2, containing a T7 promoter and an ampicillin resistance marker, was provided by Dr. U. Deuschle. Plasmid pJ505 (10) was obtained from Dr. J.N. Nelson. Plasmid pQ8, containing the *MspI* methylase gene, was constructed by Dr. A. Bhagwat. Plasmid pHSP0-1, containing the *HpaII* methylase gene in the plasmid pUHE25-2, was constructed by Dr. S. Klimasauskas using a method identical to

* To whom correspondence should be addressed

that described for *M.HhaI* (10). Plasmid pMSP, containing the *M.SplI* gene in plasmid pUHE25-2, was prepared similarly using primers 15 and 2. Plasmid pNK627, which is a pACYC184 derivative carrying the *lacI^q* gene inserted within its chloramphenicol resistance gene, was obtained from Dr. N.Kleckner (Harvard University). Plasmids were prepared by the alkaline lysis method (15) and purified on Qiagen columns. Transformation of *E.coli* cell lines by plasmid DNA was as described by Hanahan (16).

Enzymes and reagents

All restriction enzymes, T4 DNA ligase, plasmid pBR322, Vent DNA polymerase and bacteriophage lambda phage DNA were obtained from New England Biolabs and used according to the manufacturer's specifications. Deoxycytidine-5'-[α -³²P]-triphosphate and deoxyguanosine-5'-[α -³²P]-triphosphate (800 Ci/mmole) were from New England Nuclear. Spleen phosphodiesterase II and micrococcal nuclease were from Worthington. The DNA sequencing kit was from United States Biochemicals. Columns for plasmid DNA purification were obtained from Qiagen. PEI-F cellulose plates were from J.T.Baker. Cytosine was obtained from ICN.

PCR primers

The following primers, used for the polymerase chain reaction (PCR) and DNA sequencing, were synthesized by the Cold Spring Harbor Oligonucleotide Core Facility:

- 1: 5' GAAAGAATTTATATTGTAGGCTTCCTAAATCAAATATTC 3'
- 2: 5' TTCCACTTCTTGAAGCTTTTAAATTGATTT 3'
- 3: 5' CGATGCATGCGAGATGTGTTA 3'
- 4: 5' GCCTACAATATAAATCTTTC 3'
- 5: 5' CGGACGTCCAGCTGAGAT 3'
- 6: 5' AAATCGTGAAGGATCCGTCTTTT 3'
- 7: 5' AAAAGACGGATCCCTTCACGATTT 3'
- 8: 5' AGGTTTACCATCATCTTTAGTTATCAATATACT 3'
- 9: 5' CAGTATATTGATACTAAAGATGATGGTAAACCT 3'
- 10: 5' AGTTATCTCTTTAAATTAAGAAAACATAAAGAA 3'
- 11: 5' TTCTTTATGTTTTCTTAATTTAAAGAGATAACT 3'
- 12: 5' AAGGATGGAGAAACAGTATTCGTAATAATGACCCCT 3'
- 13: 5' AGGGGTCATTTACGAATACCTGTTTCTCCATCCTT 3'
- 14: 5' TAGATTTTGCATGCAAATGAAACCT 3'
- 15: 5' CTTTAGATTCAAATTGTGAGCGG 3'
- 16: 5' ATCAACAGGAGTCCAAGCTCAGCTCAG 3'
- 17: 5' AACACCAGTGCTTTTATGGAAAGCTACTAAGTA 3'
- 18: 5' TACTTAGTAGCTTTCCATAAAAGCACTGGTGTT 3'
- 19: 5' TAGATTTTGCATGCAAATGAAACCT 3'
- 20: 5' AACACCAGTGCTTTTATGAAAAGCTACTAAGTAAAAACG 3'
- 21: 5' TTGGTGTGCCACAAAATAGA 3'

Construction of plasmids expressing hybrid methylases

The PCR strategy for generating hybrid methylases used the splicing by overlap extension method (17). Briefly, amplification was carried out for 20 or 25 cycles, each consisting of 1 min denaturation at 94°, 1 min annealing at 45° to 60°, and 2 min chain elongation at 72°. After the last cycle, samples were chain elongated for an additional 20 min to complete the synthesis. Five percent of the PCR product was analyzed by 0.9% agarose gel electrophoresis and the remainder was agarose gel purified and digested with *SphI* and *HindIII* restriction endonucleases for cloning into plasmid pUHE25-2, previously cleaved at these sites. The *SphI* site, located at the 5'-terminus of the methylase gene was engineered into the sequence from the 5' PCR primer. The *HindIII* site is part of the 3' end of the methylase sequence.

Table 1 shows a summary of the PCR products produced by the different plasmid templates and primers. In each case the complete structures of the final constructs were confirmed by DNA sequencing.

Cell crude extract preparation

One ml of an overnight culture of *E.coli* ER1727 cells harboring a recombinant plasmid was diluted to 1000 ml in LB medium with ampicillin (100 μ g/ml) and grown at 37°C until an OD₆₀₀ of 0.8. IPTG (400 μ M) was added to the culture which was then shaken for an additional 2 to 3 hours. The cells were harvested by centrifugation for 10 min at 10000 g. The cell pellet was resuspended in 10 ml of lysis buffer (10 mM Na-phosphate, 30 mM NaCl, 0.25% Tween-20, 10 mM β -mercaptoethanol and 10 mM Na₂EDTA, pH 7.0) and sonicated at 20 Hz for 1 min on ice. Streptomycin sulfate (2 mg/ml) was added to the lysed cells, the final salt concentration was adjusted to 500 mM, and the lysate was shaken at 4°C for 30 min. The lysate was then centrifuged for 10 min at 10,000 g and the supernatant adjusted to 70% with ammonium sulfate. After shaking at 4°C for 30 min, the suspension was centrifuged at 10,000 g for 10 min. The pellet was dissolved in 1 ml of 10 mM Tris.HCl (pH 7.9), 10 mM Na₂EDTA and aliquoted into small volumes for storage at -70°C.

Methylation specificity and activity assays

a. In vivo Mcr restriction assay: The method is a modification of that described by Raleigh et. al (19). Briefly, plasmids modified by various methylases were transformed into *McrA*⁺*McrB*⁻, *McrB*⁺*McrA*⁻ or *McrA*⁻*McrB*⁻ strains of *E.coli*. These strains all carry the tetracycline resistant plasmid pNK627 expressing the *lacI^q* gene. This serves to suppress expression of the methylase gene from a T7 promoter regulated by a *lac* operator in the incoming plasmid. Suppression of methylase expression is necessary to avoid the deleterious effects of overexpression of the methylase gene on cell viability, which would mask the effects of methylation. The transformation ratio was calculated as the ratio of the number of transformants obtained in *Mcr*⁺ versus *Mcr*⁻ strain. For experiments to test the effects of induction the recipient cells were induced prior to the preparation of competent cells.

b. Sensitivity to restriction endonucleases: Plasmid DNAs, carrying wild type or hybrid methylase genes, were isolated using the alkaline lysis method, further purified on Qiagen columns

Table 1. Hybrid Methylase Construction

Hybrid	Reaction 1		Reaction 2		Reaction 3
	Template DNA	Oligos	Template DNA	Oligos	Oligos
HM1	pJ505	3,4	pQ8	1,2	3 + 2
HM2	pJ505	3,6	pQ8	2,7	3 + 2
HM5	pHSP0-1	15,8	pMSP	16,9	15 + 16
HM6	pHSP0-1	15,11	pHM1	16,10	15 + 16
HMH	pHM1	3,13	pJ505	5,12	3 + 5
MHM	pMSP	15,17	HM2	16,18	15 + 16
MH1	pMSP	19,20	pJ505	5,21	19 + 5

The hybrid methylases were assembled using the polymerase chain reaction to splice by the overlap extension method (17). Each hybrid methylase was constructed in two steps: reactions 1 and 2 were used to generate the individual fragments, which were then fused in reaction 3 to generate the hybrid. DNA templates and primers for each reaction are shown. The oligonucleotide numbers correspond to the sequences indicated in the Methods.

and the DNA digested with an excess of *R.HpaII* or *R.MspI* (2 hr/20 units/0.4 μ g DNA) restriction endonucleases. The digestion products were analysed by 1% agarose gel electrophoresis in the presence of ethidium bromide.

c. *In vitro* methylation assays: a) Methylation of bacteriophage λ DNA: Crude cell extracts were prepared as described above. Bacteriophage λ DNA was incubated in the presence of [3 H-methyl]-SAM with various crude cell extracts prepared from ER1727 cells expressing hybrid methylases. The incorporation of [3 H]-methyl groups into DNA was quantitated using a filter binding assay (20). Briefly, the DNA was incubated with various crude cell extracts and [3 H]-methyl SAM in a buffer containing 50 mM Tris.HCl (pH 7.5), 10 mM Na₂EDTA, 5 mM β -mercaptoethanol at 30°C for 0 to 60 min. 20 μ l of the reaction mixture were spotted onto DE-81 filter paper, washed 5 times with 3 ml 0.2 M NH₄HCO₃, twice with distilled water and finally with 95% EtOH. The filters were dried and counted in Cytosint.

b) Methylation Specificity Assay: pBR322 DNA was first methylated by either *M.MspI* or *M.HpaII* in the presence of unlabeled SAM. Following phenol extraction and ethanol precipitation, the DNA was dried and resuspended into distilled water. The pre-methylated DNA was then incubated with either HM1 or HM2 cell crude extract in the presence of [3 H]-SAM and the extent of further methylation quantitated using the filter binding assay.

c) Nearest-neighbor analysis: The plasmid DNA, carrying either a wild type or a hybrid methylase, was digested with DNaseI (0.5 mg/ml) for 3 min. *E.coli* DNA polymerase I and [α^{32} p]dCTP or [α^{32} p]dGTP were added and incubation continued at 15°C for 15 min. The reaction mixture was passed through a Sephadex G-50 spin column to remove unincorporated triphosphate. The radiolabeled DNA was then digested with spleen phosphodiesterase and micrococcal nuclease for 4 hours

and the reaction mixture was analysed by TLC on PEI-cellulose in two dimensions. 32 P-labeled products were detected by autoradiography.

RESULTS

A series of hybrid methylases were constructed between *M.MspI* and *M.HpaII* so that the variable regions and their flanking sequences were derived from different methylases (Figure 1). Hybrids are named as HM1–6 when the N-terminal segment is derived from *M.HpaII* and the C-terminal segment is derived from *M.MspI*. Constructs named MH, MHM or HMM follow the same logic. Except for hybrids HM5 and HM6, the junctions between *M.MspI* and *M.HpaII* sequences were chosen so that they lay within conserved motifs VIII or IX, which flank the variable region. Two key hybrids are HM1 and HM2. HM1 contains the N-terminus of *M.HpaII*, up to and including motif VIII, joined to the variable region and C-terminus of *M.MspI*. HM2 contains the N-terminus of *M.HpaII* up to motif IX, including the complete variable region. The C-terminus of HM2 is from *M.MspI* and includes motif IX and all downstream sequences. The only difference between HM1 and HM2 is the variable region: HM1 contains the *MspI* variable region and HM2 contains the *HpaII* methylase variable region.

Hybrid methylase specificity

Two tests were employed to measure the methylase activity of the hybrids. First, plasmid DNA was prepared from cultures either uninduced or induced for methylase production. The DNA was then challenged *in vitro* for its sensitivity to digestion by either *R.HpaII* or *R.MspI*. DNA from the uninduced cultures is expected to be unmethylated or poorly methylated. After induction, if the hybrid methylase is active, it should provide protection against one or both of the cognate restriction enzymes. The results are shown in Figure 2. Among the hybrid methylases

Figure 1. Structure and properties of hybrid methylases. The uppermost line indicates the ten conserved motifs found in most prokaryotic 5mC methylases (3,4). The solid bars represent the highly conserved motifs and the open bars represent the moderately conserved motifs; the variable region is located between motifs VIII and IX. Sequences derived from *M.HpaII* are indicated by open boxes and those from *M.MspI* are shown as hatched boxes. The activity of each hybrid methylase *in vitro* was measured by digestion of the encoding plasmid DNA with the cognate restriction endonuclease. Plasmid DNAs for this test were produced either with or without IPTG induction of the methylase. The extent of methylation is indicated – for no protection, + for partial protection and ++ for complete protection. The specificity of the methylases is shown in the right column.

Figure 2. Hybrid methylase specificity. The specificity of each hybrid methylase was determined by the sensitivity of the encoding plasmid DNA to digestion with either *R.HpaII* or *R.MspI*. Plasmid DNAs were prepared either with (+) or without (–) IPTG induction to over express the hybrid methylases. The plasmid DNAs, modified by various methylases *in vivo*, were purified and digested with an excess of the indicated endonuclease. The digestion products were analyzed by 1% agarose gel electrophoresis in the presence of 0.5 µg/ml ethidium bromide.

constructed only HM1 and HM2 showed signs of protection against one or other of the cognate restriction enzymes. These hybrids displayed a specificity that was identical to that of the parent methylase contributing the variable region. Thus HM1 DNA from induced cultures is almost completely resistant to cleavage by *R.HpaII* and *R.MspI*. This pattern of resistance is also shown by its parent *M.MspI* (Figure 2). Similarly HM2 shows the pattern of resistance displayed by its parent, *M.HpaII*, although the hybrid methylase does not give as complete protection as either HM1 or *M.HpaII*. The HMH and MHM hybrids, which contain exact swaps of the variable regions, appeared inactive in this test as did MH1, the reciprocal of HM1. None of the other hybrids shown in Figure 1 were active.

The *in vitro* restriction assay (Figure 2) measures steady state levels of methylation, but only detects the specific methylation against which the DNA is challenged. We therefore employed a second test to measure methylase activity. The *Mcr* systems of *E.coli* have a broad specificity (22) and could detect methylation at sites other than the CCGG expected. The *McrA* system restricts DNA methylated by *M.HpaII*, but does not affect DNA methylated by *M.MspI* or other cytosine methylases. In contrast, *McrB* restricts DNA methylated by many m5C-methylases, including *M.MspI*, but not *M.HpaII*. We therefore measured the ability of each hybrid to transform either *McrA*⁺ or *McrB*⁺ strains of *E.coli*. The results are shown in Table 2. Only HM1 and HM2 showed clear evidence of restriction by these *Mcr* systems. The results indicate that hybrid HM1 is restricted in an *McrB*⁺ strain approximately 1000 fold (wildtype *MspI* is restricted in *McrB*⁺ strains about 10,000 fold), but not restricted in an *McrA*⁺ strain. HM2 is restricted in an *McrA*⁺ strain approximately 5 fold (wildtype *HpaII* is restricted in a *McrA*⁺ strain about 3000 fold), but not restricted in an *McrB*⁺ strain. This test provides confirmation that HM1 has the methylation specificity of *M.MspI* and HM2 has the specificity

Table 2. Hybrid methylase transformation ratios in *Mcr* strains

Methylase	A ⁺ /A [–] B [–]	B ⁺ /A [–] B [–]
None (pUHE25-2)	1	1
<i>M.MspI</i>	1	1.0 × 10 ^{–4}
<i>M.HpaII</i>	3.5 × 10 ^{–4}	0.95
HM1 uninduced	0.58	0.91
HM1 induced	0.46	1.0 × 10 ^{–3}
HM2 uninduced	0.76	0.81
HM2 induced	0.20	0.90

The hybrid methylases and their parental wild type genes, cloned in pUHE25-2, were transformed into either *McrA*[–]*McrB*[–], *McrA*⁺*McrB*[–] or *McrA*[–]*McrB*⁺ strains of *E.coli* carrying pNK627 (*lacI*^q). The ratio of the number of transformants obtained in an *McrA*⁺ or an *McrB*⁺ strain divided by the number of transformants obtained in the *McrA*[–]*McrB*[–] strain is shown in column one. pUHE25-2 is the expression vector containing no methylase gene insert. The *M.HpaII* and *M.MspI* parental methylase genes serve as a positive control. Restriction was monitored either with induction of the incoming methylase gene using IPTG or in the absence of induction.

of *M.HpaII*. The poor efficiency of the HM2 hybrid displayed during the *in vitro* restriction assay is also apparent by this *in vivo* assay.

Direct evidence showing that HM1 has the precise specificity of *M.MspI* was provided by nearest neighbor analysis of the base methylated. HM1 plasmid DNA was labelled in a nick translation reaction containing [α -³²P]-dCTP and then subjected to complete digestion to 3'-mononucleotides using micrococcal nuclease and spleen phosphodiesterase. In this experiment the label is transferred from the 5'-position of the labelling nucleotide to the 3'-position of its neighbour. Upon analysis by two dimensional thin layer chromatography a spot corresponding to m5CMP was observed (Figure 3A). However, if the DNA was labelled with [α -³²P]-dGTP there was no corresponding signal (Figure 3B). The HM1 methylation pattern was the same as the *MspI* methylation pattern, but with a less intense methylated m5C spot. This result shows that HM1 transfers the methyl group to the outer cytosine residue in the CCGG recognition sequence. The specificity of HM2 was the same as *M.HpaII* methylase. The m5C spot was only detected when the DNA was labelled by [α -³²P]-dGTP but not [α -³²P]-dCTP (Figure 3). In this case the inner cytosine in the CCGG recognition sequence becomes methylated.

Hybrid methylase activity

Among the hybrid methylases that we prepared only HM1 and HM2 showed methylase activity, although their efficiency as methylases appeared considerably reduced from that of either parent. They gave only partial protection from digestion by their cognate restriction endonucleases and responded weakly in the *in vivo* *Mcr* transformation assays. These results could have been due to a lower level of expression of the hybrid genes or some inherent instability in the protein products. This was tested by subjecting a crude cell extract from each of the hybrid methylases to SDS-PAGE. In all cases, following induction, the hybrid proteins were clearly visible and the amounts were comparable to those of the wild type methylases (data not shown). This indicated that neither poor expression nor protein instability was the cause of the poor activity.

We monitored the methylation activity of HM1 and HM2 by measuring the time course for transfer of [³H]-labelled methyl groups from the methyl donor S-adenosylmethionine (SAM) to the cytosine residue in the target sequences within bacteriophage

Figure 3. Analysis of the base methylated by HM1 and HM2. Nearest neighbor analysis was used to identify the methylated cytosine in the CCGG recognition sequence. The samples in Panel A were labelled with [$\alpha^{32}\text{P}$]-dCTP. 5mCMP is detected if the outer cytosine in the CCGG sequence is methylated. The samples in Panel B were labelled by [$\alpha^{32}\text{P}$]-dGTP. 5mCMP will be detected if the inner cytosine in the CCGG sequence is methylated.

λ DNA (data not shown). Both hybrids showed methyltransferase activity, albeit much weaker than their parental methylases. Although HM1 was approximately twice as efficient as HM2 its initial rate of methylation was only about 10% that of the wild type *M.HpaII*.

To test the specificity of the reactions catalysed by these hybrid methylases *in vitro* we measured their ability to transfer methyl groups to pBR322 DNA that had been premethylated with either *M.HpaII* or *M.MspI*. The results are summarized in Table 3. HM1 was able to transfer methyl groups to DNA that had been premethylated with *M.HpaII*, but gave essentially background levels of incorporation into DNA premethylated with *M.MspI*. HM2 gave equivocal results, probably because it was a much poorer methylase than even HM1. It should be noted that even the wild type enzymes performed much more poorly on DNA that had previously been methylated at the non-cognate site when compared with unmethylated pBR322 DNA.

We also carried out DNA binding studies, using a mobility shift assay. The preliminary results, using crude extracts, suggested that all of the hybrid methylases have a much lower affinity for the specific DNA sequence than the wild type methylase (data not shown) and precluded further experiments to test for DNA binding specificity.

DISCUSSION

In earlier work when preparing hybrids between *M.HpaII* and *M.HhaI* we had found that almost half of the hybrid methylases retained activity although in every case they were considerably weaker than their wild type parents (10). In our present studies, with *M.HpaII* and *M.MspI*, the success rate for the production of active methylases was much less but still we were able to prepare two key hybrids that allowed us to assess unequivocally the role of the variable region in the choice of the base to be methylated. As before it seems that these hybrid methylases retain greater activity when the variable region plus the flanking motif

Table 3. Hybrid Methylase activity *in vitro*

Methylase	pBR322/ <i>M.HpaII</i>	pBR322/ <i>M.MspI</i>	pBR322
none (pUHE25-2)	431	444	573
<i>M.MspI</i>	11,408	1,576	105,102
<i>M.HpaII</i>	1,821	10,078	113,375
HM1	1,430	633	1,364
HM2	986	693	1,426

Crude extracts were prepared from equal amounts of *E.coli* carrying plasmids encoding the various methylases. Samples from these extracts were used to methylate plasmid pBR322. An extract from cells carrying plasmid pUHE25, the expression vector alone, serves as a negative control. pBR322/*HpaII*.M. indicates that plasmid pBR322 has been exhaustively methylated with wild type *M.HpaII* in the presence of unlabeled SAM. pBR322/*MspI*.M. indicates that plasmid pBR322 has been exhaustively methylated with wild type *M.MspI*. Methylation was monitored by the transfer of the radiolabelled methyl group from [^3H]-SAM into DNA.

IX derive from the same parent. This suggests that there may be some direct structural interactions between the variable region and motif IX.

Throughout our experiments to prepare and characterize hybrid methylases we have been concerned about the most suitable methods to detect residual methylase activity. In our present experiments we have had the opportunity to compare three different assay methods. In the first method, plasmid DNA is prepared from a methylase-producing strain and tested *in vitro* for its sensitivity to restriction endonuclease digestion. In a second method, we have monitored the ability of the *E.coli* *Mcr* systems to restrict transformation by plasmids encoding methylases. Finally we have used crude cell extracts from strains encoding the hybrid methylases and checked for the transfer of methyl groups to recipient unmethylated DNA. It is clear that the last method is the least sensitive. We had anticipated that the *Mcr* assays would be the most sensitive, because of the continuous production of methylated DNA which should constantly provide

new targets for the *Mcr* systems. However, our results suggest that testing plasmid sensitivity to restriction enzyme digestion is preferable. Perhaps the sensitivity of this approach could be improved by overproducing the *Mcr* proteins.

It might appear surprising at first sight that the *in vitro* activity of the hybrid methylases HM1 and HM2 is so low (Table 3) in light of the extent of methylation that is shown in Figure 2. However, it must be remembered that the *in vivo* levels of methylated product which are being monitored in the experiment of Figure 2 represent the continuous production and methylation of plasmid DNA during cell growth whereas the *in vitro* activity represents a series of distinct time points under conditions which are probably not optimal for methylation.

By constructing hybrids between the *HhaI* methylase and the *HpaII* methylase we have previously shown (10) that the variable region lying between motifs VIII and IX determines the DNA sequence specificity of the *HpaII* methylase. In the present work we have constructed similar hybrids between the *HpaII* methylase and the *MspI* methylase to ask whether the variable region also determines the choice of base to be methylated. Two hybrids, HM1 and HM2, which differ only in the source of their variable region have allowed us to answer this question since HM1, which contains the variable region from *M.MspI*, specifically places the methyl group on the outer cytosine of the recognition sequence CCGG. Similarly the hybrid methylase HM2, which is identical to HM1 except that its variable region is from *M.HpaII*, catalyzes the transfer of the methyl group to the inner cytosine residue just as does *M.HpaII*. Thus, in this system and we suspect in other m5C-methylases, the base which becomes methylated depends critically upon the nature of the variable region. This would be consistent with the well-conserved architecture of the m5C-methylases and supports a structural model in which the 10 well-conserved motifs provide a common framework containing the catalytic cysteine. In this model the variable region can be viewed as a single domain that binds to a specific DNA sequence and holds it in a suitable conformation ready for reaction with the catalytic cysteine.

It is now clearly established both for the monospecific methyltransferases (10) and the multispecific methyltransferases (23) that the variable region is wholly responsible for the selection of DNA sequences to be methylated. In the case of the multispecific enzymes short segments within the variable region are implicated in specific DNA interactions (24). It seems likely that in the monospecific methyltransferases it is not the entire variable region that is responsible for DNA sequence recognition but rather some smaller part of it. It will be important to further dissect the variable region either by direct mutational analysis or by the preparation of further active hybrids so as to determine which portions interact directly with DNA.

ACKNOWLEDGEMENTS

We thank S.Klimasauskas for plasmids and advice on experiments, A.Krainer, S.Kulakauskas, S.Kumar and E.Richards for helpful comments on the manuscript and M.Ockler, J.Duffy and P.Renna for art and photography. This work was supported by grants from the NIH (GM 46127) and NSF (DMB-8917650). Centralized facility support for oligonucleotide synthesis was from a Cancer Center Support grant from the NCI (CA 45508).

REFERENCES

- Adams, R.L.P. (1990) *Biochem. J.* **265**, 309–320.
- Modrich, P. (1982) *CRC Crit. Rev. Biochem.* **13**, 287–323.
- Lauster, R., Trautner, T.A. and Noyer-Weidner, M. (1989) *J. Mol. Biol.* **206**, 305–312.
- Posfai, J., Bhagwat, A.S., Posfai, G. and Roberts, R.J. (1989) *Nucl. Acids Res.* **17**, 2421–2435.
- Ingrosso, D., Fowler, A.V., Bleibaum, J. and Clarke, S. (1989) *J. Biol. Chem.* **264**, 20130–20139.
- Klimasauskas, S., Timinskas, A., Menkevicius, S., Butkiene, D., Butkus, V. and Janulaitis, A. (1989) *Nucl. Acids Res.* **17**, 9823–9832.
- Wu, J.C. and Santi, D.V. (1985) In *Biochemistry and Biology of DNA Methylation*, (G.I. Cantoni and A. Razin, Eds) Alan R. Liss Inc., New York, pp 115–129.
- Chen, L., MacMillan, A.M., Chang, W., Ezaz-Nikpay, K., Lane, W.S. and Verdine, G. (1991) *Biochemistry* **30**, 11018–11025.
- Wilke, K., Rauhut, E., Noyer-Weidner, M., Lauster, R., Pawlek, B., Behrens, B. and Trautner, T.A. (1988) *EMBO J.* **7**, 2601–2609.
- Klimasauskas, S., Nelson, J.L. and Roberts, R.J. (1991) *Nucl. Acids Res.* **19**, 6183–6190.
- Walder, R.Y., Langtimm, C.J., Catterjee, R. and Walder, J.A. (1983) *J. Biol. Chem.* **258**, 1235–1241.
- Mann, M.B. and Smith, H.O. (1977) *Nucl. Acids Res.* **4**, 4211–4221.
- Dila, D., Sutherland, E., Moran, L., Slatko, B. and Raleigh, E.A. (1990) *J. Bacteriol.* **172**, 4888–4900.
- Kelleher, J.E. and Raleigh, E.A. (1991) *J. Bacteriol.* **173**, 5220–5223.
- Birnboim, H.C. and Doly, J. (1979) *Nucl. Acids Res.* **7**, 1513–1523.
- Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.
- Ho, S.N., Pullen, J.K., Horton, R.M., Hunt, H.D. and Pease, L.R. (1990) *DNA and Protein Engineering Techniques* **2**, 50–55.
- Laemmli, U.K. (1970) *Nature* **227**, 680–685.
- Raleigh, E.A., Trimarchi, R. and Revel, H. (1989) *Genetics* **122**, 279–296.
- Dubey, A.K., Mollet, B. and Roberts, R.J. (1992) *Nucl. Acids Res.* **20**, 1579–1585.
- Cedar, H., Solage, A., Glaser, G. and Razin, A. (1979) *Nucl. Acids Res.* **6**, 2125–2131.
- Raleigh, E.A. and Wilson, G. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9070–9074.
- Balganesh, T.S., Reinert, L., Lauster, R., Noyer-Weidner, M., Wilke, K. and Trautner, T.A. (1987) *EMBO J.* **6**, 3543–3549.
- Wilke, K., Rauhut, E., Noyer-Weidner, M., Lauster, R., Pawlek, B., Behrens, B. and Trautner, T.A. (1988) *EMBO J.* **7**, 2601–2609.