Volume 4 Number 12 December 1977 **Nucleic Acids Research**

Specificity of Hpa II and Hae III DNA methylases

Michael B. Mann and Hamilton O. Smith

Department of Microbiology, The Johns Hopkins University, School of Medicine, Baltimore, MD 21205, USA

Received 31 October 1977

ABSTRACT

The methylases M•HaeIII and M•HpaII recognize the tetranucleotide sequences (5') G-G-$\overset{*}{C}$-C / C-$\overset{*}{C}$-G-G (5') and (5') C-$\overset{*}{C}$-G-G / G-G-$\overset{*}{C}$-C (5') respectively, in DNA, and transfer a methyl group from S-adenosylmethionine to the 5-position of cytosine on each strand as indicated by the asterisks. Restriction endonuclease R•HaeIII does not cleave the methylated sequence G-G-$\overset{m}{C}$-C / C-C-G-G but can cleave G-G-C-$\overset{m}{C}$ / C-C-G-G in which methylation is introduced on the unnatural external cytosine positions. Similarly, R•HpaII does not cleave C-$\overset{m}{C}$-G-G / G-G-C-C but can cleave $\overset{m}{C}$-C-G-G / G-G-C-C.

INTRODUCTION

Restriction endonucleases R•HaeIII from Hemophilus aegyptius (1) and R•HpaII from Hemophilus parainfluenzae (2) recognize the tetranucleotide sites (5') G-G$\overset{\downarrow}{}$C-C / C-C$\overset{\uparrow}{}$-G-G (5') (3) and (5') C$\overset{\downarrow}{}$C-G-G / G-G-C$\overset{\uparrow}{}$C (5') (4), respectively, in DNA and cleave both strands at the positions indicated by the arrows. We report here the identification of the corresponding DNA methylases, M•HaeIII and M•HpaII, and we describe their methylation specificities.

As a substrate, we used a polymer constructed by blunt-ended ligation of the self-complementary decadeoxynucleotide (5'-3') pC-C-G-G-A-T-C-C-G-G. This tandemly repeating polymer contains 2 tandem HpaII sites and a HaeIII site with the arrangement (5') . . . C-C-G-G(HaeIII)pC-C-G-G . . ., at each junction . . . G-G-C-CpG-G-C-C . . . (HpaII HpaII) between the monomeric units. By introducing ^{32}P solely at these junction sites, a substrate was obtained which proved particularly suited for determining the methylation specificity of both enzymes.

MATERIALS AND METHODS

Enzymes. T4 polynucleotide kinase was from P-L Biochemicals. Pan-

creatic DNase, bacterial alkaline phosphatase, spleen phosphodiesterase, and venom phosphodiesterase were from Worthington Biochemical Corp. T4 DNA ligase and R·BamI were from BioLabs. R·HaeIII and R·HpaII were purified by standard procedures (1, 2). One unit of restriction enzyme is the amount required to yield a complete digest of phage λDNA in 1 hour at 37°.

Purification of M·HaeIII and M·HpaII methylases. Identical procedures were used for both enzymes. Hemophilus aegyptius (ATCC 11116) and Hemophilus parainfluenzae (from J. Setlow) cells were grown into late stationary phase in Difco brain heart infusion broth containing 10 μg/ml hemin and 2 μg/ml NAD. 50 gm of harvested cells, suspended in 50 ml of 10 mM Tris-Cl (pH 8)-5 mM 2-mercaptoethanol, were disrupted by sonication. Cell debris was removed by centrifugation at 100,000 x g for 3 hours. The supernatant was loaded onto a DEAE Sephadex A-25-120 column (200 ml bed volume) and eluted with 0.3 M NaCl in TGESH buffer (T=10 mM Tris-Cl, pH 8; G=5% glycerol; E=1 mM EDTA; SH=5 mM 2-mercaptoethanol). The protein-containing fractions were precipitated in 70% ammonium sulfate. Precipitate was dissolved in 50 ml of PGESH (same as TGESH, but buffered with 10 mM sodium phosphate, pH 7.5) buffer and chromatographed on a Sephadex G-50 column (750 ml bed volume) with PGESH buffer. The excluded fraction was loaded onto a single-stranded DNA-agarose column (75 ml bed volume, equilibrated with PGESH buffer) and eluted with 1 liter of a 0 to 1 M NaCl gradient in PGESH buffer. Fractions were assayed for DNA methylase activity as described below. Both methylases eluted at approximately 0.65 M-0.75 M NaCl. Fractions containing activity were pooled and concentrated 20-fold by ultrafiltration through a PM-10 Amicon filter. Enzyme was stored at -20° in 50% glycerol.

Methylase assays were carried out in reaction mixtures (20 μl) containing 8 μg salmon sperm DNA, 4 μM [methyl-^{3}H]S-adenosylmethionine, 50 mM Tris-Cl (pH 7.5), 20 mM 2-mercaptoethanol, 10 mM EDTA, and 1-5 μl of enzyme fraction. Incubation was at 37° for 1 hour. A 10 μl sample was then chromatographed on a 1 cm x 5 cm polyethyleneimine thin-layer strip with 1 M HCl and the lower half of the strip was counted to determine acid-precipitable radioactivity remaining at the origin (5). One unit of methylase activity incorporated 1 pmole of methyl groups into DNA in 1 hour.

Synthesis of tandemly repeating decamer polymer. The decadeoxynucleotide (5'-3') CpCpGpGpApTpCpCpGpG was kindly provided by Saran Narang,

Chandler P. Bahl and Ray Wu. The synthesis of this BamI adapter (6) and confirmation of the sequence (7) are published. The decamer exists as a duplex under usual reaction conditions because of its self-complementary structure and high GC content.

The decamer, which contained a 5'-OH group, was ^{32}P-labeled at the 5'-terminus in a reaction mixture (62.5 µl) containing 225 pmoles of decamer molecules, 50 pmoles of [γ-^{32}P]ATP (1000 Ci/mmole), 50 mM Tris-Cl (pH 9.5), 10 mM MgCl , 5 mM dithiothrietol, 0.1 mM EDTA, 5% glycerol, and 715 units of T4 kinase. Incubation was at 37° for 10 min after which 15 nmoles of unlabeled ATP (5 µl) and 7.5 units of additional T4 kinase (2.5 µl) were added and incubation continued for 30 min at 37° to complete the phosphorylation. The reaction was terminated by addition of 0.2 ml of 0.2 M NH_4HCO_3 and 2 µl of 0.5 M EDTA. The mixture was extracted twice with chloroform and then chromatographed on a 1 cm x 14 cm Sephadex G-25 (fine) column with 0.2 M NH_4HCO_3. The radioactive decamer peak (1.5 ml) moved well ahead of unincorporated ATP and non-volatile salts. After repeated lyophilization over KOH and phosphorus pentoxide, the residue was dissolved in 5 µl of water.

Blunt-end ligation of the [5'-^{32}P] decamer was carried out in a reaction mixture (10 µl) containing the 5 µl of decamer, 1 µl of 10X ligase cocktail, (0.2 M Tris-HCl (pH 7.6), 0.1 M MgCl, 0.1 M dithiothrietol, 4 mM ATP), and 4 µl of T4 ligase (92 units/ml). Incubation was at 23° for 1 h during which 70% (and 88% in a second preparation) of the radioactivity became resistant to alkaline phosphatase. The polymer was diluted to 40 µl with 10 mM Tris-Cl (pH 7.6), 1 mM EDTA (TE buffer) and stored at -20°. One 1 µl (2 x 10 cpm) contained approximately 5 pmoles of polymerized decamer units.

Pure 50-mer (5 tandem decamer units), labeled only at its 5'-termini, was made by the following procedure. Unlabeled polymer was made by the above protocol except for omission of [γ-^{32}P]ATP. After removal of terminal 5'-phosphoryl groups with phosphatase, it was ^{32}P-labeled at the 5'-termini in a standard kinase reaction. The reaction mixture was then phenol extracted, the DNA precipitated twice with ethanol, dried, and finally dissolved in 50 µl of TE buffer. The product was fractionated by electrophoresis on a 20% polyacrylamide gel (no urea). Individual polymer bands were located by autoradiography, excised, and the DNA recovered by the methods of Maxam and Gilbert (8). The 50-mer band, obtained in 50 µl of TE buffer, contained $\sim$3000 cpm/µl.

Gel electrophoresis. Electrophoresis on 20% polyacrylamide gels with or without 7 M urea, recovery of DNA from gel bands, and autoradiography of gels followed the methods set forth by Maxam and Gilbert (8).

RESULTS AND DISCUSSION

Construction of a tandem decamer polymer containing HaeIII and HpaII sites. The self complementary deoxyoligonucleotide decamer (5'-3') C-C-G-G-A-T-C-C-G-G (in duplex form) was [5'-^{32}P] labeled using T4 kinase and [γ-^{32}P]ATP followed by blunt-ended ligation into a tandemly repeating polymer series using T4 ligase as described in Materials and Methods. Electrophoretic analysis on a 20% polyacrylamide gel (without urea) showed a series of bands at positions corresponding to lengths (in base pairs) of 20, 30, 40, ..., >100 (Figure 1A). (The position of unpolymerized [^{32}P]decamer had been previously determined relative to the bromphenol blue marker.) Denatured polymer was analysed in the adjacent gel tract (Figure 1B). Bands at positions of 10, 15, 20, 25 ... correspond

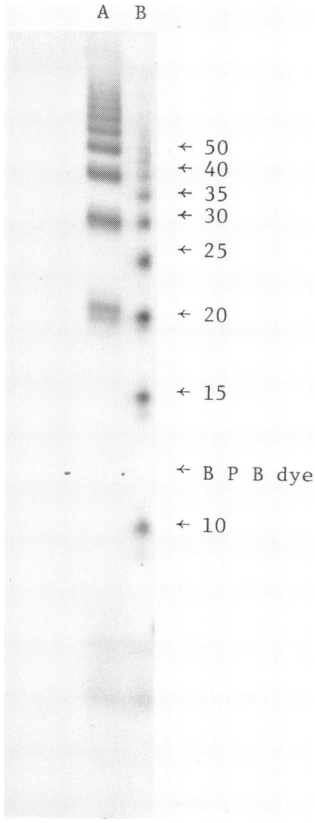

Figure 1. Gel electrophoresis of decamer ligation products.

The 5'-labeled self-complementary decadeoxynucleotide pC-C-G-G-A-T-C-C-G-G (as duplex) was blunt-end ligated using T4 ligase as described in Materials and Methods. The polymerization products were then analysed on a 20% polyacrylamide gel (no urea) and autoradiographed.

A) Native.

B) Denatured and loaded in 0.1 M NaOH. The numbers indicate duplex chain length.

to various hairpin structures formed by folding at either of the 2 possible points of symmetry, i.e., between each A and T or at the joints between decamers. Some electrophoretic heterogeneity was present in each major band position and could be due largely to incomplete phosphorylation. Since all the bands were shown to be sensitive to HaeIII, HpaII, and BamI cleavage, it did not appear to reflect primary sequence heterogeneity and did not interfere with the experiments described here.

The sequence at the decamer joints is predicted from the known decamer sequence to consist of 2 tandem HpaII sites and a HaeIII site straddling the labeled phosphates as shown

```
                                  HaeIII
(5') . . . C-C-G-G-A-T-C-C-G-GpC-C-G-G-A-T-C-C-G-G . . .
     . . . G-G-C-C-T-A-G-G-C-CpG-G-C-C-T-A-G-G-C-C . . . (5')
                       HpaII    HpaII
```

Several sorts of evidence confirmed this. Nucleotide analysis showed that essentially 100% of the radioactivity could be recovered as 5'-dCMP after digestion with pancreatic DNase and snake venom phosphodiesterase. Seventy percent of labeled phosphate groups were internal and not accessible to alkaline phosphatase, but after 30 min treatment with HaeIII restriction endonuclease, 92% became accessible. Finally, digestion of polymer yielded the predicted restriction enzyme cleavage products. To demonstrate this, an aliquot of purified 50-mer containing only 5'-terminal ^{32}P was digested separately with each of the three restriction endonucleases: BamI, HpaII, and HaeIII. Radioactive products were analysed by gel electrophoresis and autoradiography as shown in Figure 2. The 3 enzymes released labeled products migrating at relative positions consistent with single-strand oligonucleotide lengths of 3, 7, and 10 as predicted from the sequence

```
                    BamI      HpaII HaeIII
                     ↓          ↓     ↓
(5'-3') *pC-C-G-G-A-T-C-C-G-G-C-C-G-G-A-T-C-C-G-G
```

Experiments (not shown) verified that the R•HaeIII digestion product co-migrated with a sample of unpolymerized [^{32}P]decamer.

No labeled pC was released by HpaII even though the complete tetranucleotide site is present at the 5'-terminus, and in other experiments, we were unable to demonstrate cleavage at either site on labeled duplex decamer after extensive incubation with HpaII endonuclease. This enzyme apparently requires one or more additional nucleotides on both sides of the recognition site before cleavage can occur.

Methylation specificities of M•HaeIII and M•HpaII. The labeled polymer was used to determine the specificity of each DNA methylase according to the following rationale: If M•HaeIII methylates the cytosine adjacent

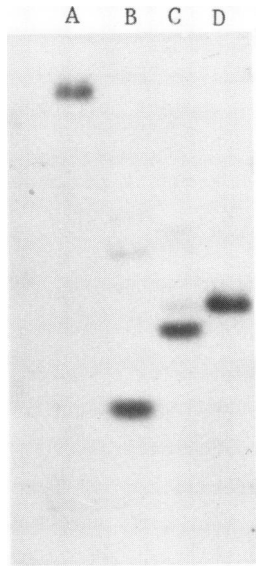

Figure 2. BamI, HpaII, and HaeIII cleavage products of 5'-terminally labeled 50-mer.

Unlabeled decamer was ligated into a series of polymerization products. 5'-terminal radioactivity was then introduced followed by fractionation on a 20% polyacrylamide gel and recovery of the 50-mer band as described in Materials and Methods. Cleavage reaction mixtures (20µl) contained 2µl of [5'-^{32}P]50-mer (~6000 cpm), 20mM Tris-Cl (pH 7.5), 7mM $MgCl_2$, 7mM 2-mercaptoethanol and 5 units of restriction enzyme. Incubation was for 15 min at 37° after which 5µl 50% glycerol was added and the samples were loaded on a 20% polyacrylamide-7M urea gel. Electrophoresis was at 600 V for 5 hours. An autoradiogram was then obtained.

A) No enzyme treatment.
B) R•BamI digestion.
C) R•HpaII digestion.
D) R•HaeIII digestion.

to the labeled phosphate, then 5'-nucleotide analysis should yield ^{32}P-labeled m^5dCMP. Methylation of the distal cytosine would yield only unlabeled m^5dCMP. Similarly, methylation with M•HpaII will yield some ^{32}P-labeled m^5dCMP if the first cytosine in the site is methylated and totally unlabeled m^5dCMP if the second one is methylated.

Polymer containing ^{32}P at each ligated joint was methylated in a reaction mixture containing [methyl-^{3}H]S-adenosylmethionine and either M•HpaII or M•HaeIII. After pancreatic DNase and venom phosphodiesterase treatment, the 5' mononucleotides were chromatographed to separate m^5dCMP. In the case of M•HpaII, no peak of ^{32}P-label was found in the [^{3}H]m^5dCMP peak (Figure 3A) indicating that the methylated cytosine is distal to the ^{32}P-groups, i.e., pC-$\overset{m}{C}$-G-G. With M•HaeIII (Figure 3B), about 50% of the ^{32}P-label moved as m^5dCMP indicating that the proximal cytosine is methylated, i.e., G-Gp$\overset{m}{C}$-C. The presence of [^{32}P]dCMP indicates that the reaction was not complete, a not unexpected result considering the sub-optimal concentration of high specific activity [methyl-^{3}H]S-adenosyl-methionine used to ensure adequate incorporation of counts. This particular experiment did not, however, rule out simultaneous methylation of the distal cytosine.

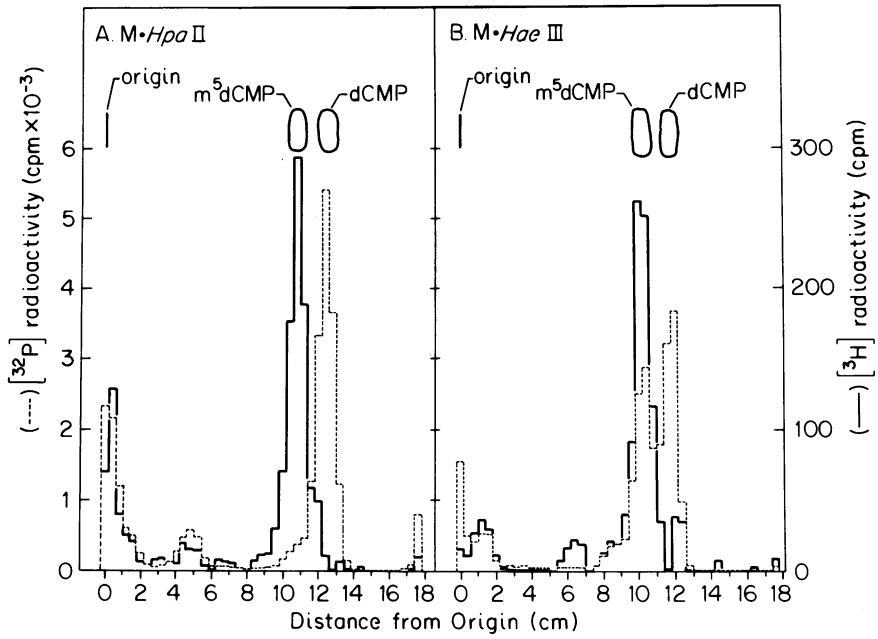

Figure 3. Chromatography of mononucleotides obtained from the tandemly polymerized decamer after methylation by either M•HpaII or M•HaeIII.

Polymerized [^{32}P]decamer (approximately 1pmole) was methylated with 10 units of methylase in 50mM Tris-HCl (pH 7.6), 20mM β-mercaptoethanol, 10mM EDTA, 20μM [methyl-^{3}H]S-adenosylmethionine (12.25Ci/mM). Incubation was for 3 h at 37°. The reaction mixtures were phenol extracted, ethanol precipitated, and then digested with pancreatic DNase and snake venom phosphodiesterase to yield 5'-nucleotides. The products were mixed with unlabeled dCMP and m^{5}dCMP markers and chromatographed on cellulose thin-layer sheets (20 cm x 20 cm) in solvent containing 60g ammonium sulfate, 100 ml 0.1M $NaPO_4$, (pH 7.0), and 2 ml n-propanol. Marker spots were located by short wave UV light. The chromatograms were then sectioned into 4mm zones and counted for ^{3}H and ^{32}P radioactivity.

To determine whether the distal cytosine was being methylated by M•HaeIII, an additional aliquot of the M•HaeIII methylated polymer was hydrolyzed to pyrimidine tracts using diphenylamine-formic acid. From the sequence of the ligated decamer, the only expected labeled product would be [methyl-^{3}H, 5'-^{32}P]d-pm^{5}CpCp. Further treatment with bacterial alkaline phosphatase would yield [methyl-^{3}H]d-m^{5}CpC and free [^{32}P] orthophosphate. If the 3'- (distal) cytosine were being methylated, then further digestion with snake venom phosphodiesterase should yield

[methyl-^{3}H]d-m^{5}CMP as well as the expected [methyl-^{3}H]m^{5}CdR; however, none of the former was found (Table I). Likewise, digestion with spleen phosphodiesterase should yield some [methyl-^{3}H]m^{5}CdR, but again, none could be detected (Table I). These data extend the previous result by confirming that M•HaeIII methylates only the interior cytosine residue in the sequence (5') . . . G-G-C-C . . . (3').

Table I. Analysis of [methyl-^{3}H]compounds from the tandemly polymerized decamer after methylation by M•HaeIII

Compound	Rf	Percent of total ^{3}H-radioactivity		
		control	venom phosphodiesterase treated	spleen phosphodiesterase treated
[methyl-^{3}H]d-m^{5}CMP	0.04	1	0	72
[methyl-^{3}H]d-m^{5}CpC	0.52	99	0	28
[methyl-^{3}H]m^{5}CdR	0.84	0	100	0

M•HaeIII methylated, polymerized [^{32}P]decamer was prepared as described in Figure 3 through the ethanol precipitation step. The material was hydrolyzed with diphenylamine-formic acid (11), followed by treatment with bacterial alkaline phosphatase. The labeled deoxynucleoside diphosphate was purified by chromatographing on a 10 cm thin-layer strip of PEI-cellulose in 0.1 M LiCl (12). Counts were eluted from the strip by washing with 0.2 M HCl. The eluate was lyophilyzed to dryness. The sample was split three ways. One aliquot (control) received no further treatment. A second was digested with approximately 2.5 µg acid-treated snake venom phosphodiesterase (13) in a 10 µl volume containing 0.1 M Tris-HCl (pH 8.9), 0.1 M NaCl, and 15 mM $MgCl_2$. (The acid treatment reduced 5'-nucleotidase activity to a negligible level.) The third aliquot was digested with approximately 10 µg of spleen phosphodiesterase in a 10 µl volume containing 50 mM sodium succinate•HCl (pH 6.5). Incubation was for 30 min at 37°C. Each reaction mixture was chromatographed on a 10 cm PEI-cellulose strip in 0.1 M LiCl with the appropriate absorbance markers. Marker regions were excised and counted for tritium. Areas between marker regions contained no detectable radioactivity. Counting error was ≤3%. Although the spleen phosphodiesterase reaction was not complete, the fact that no [methyl-^{3}H]m^{5}CdR was produced was considered to be consistent with the conclusion drawn in the text.

Effect of methylation on subsequent restriction. Methylation in a restriction site can be classified as either normal or abberant: the former referring to methyl groups introduced by the specific methylase for that site, and the latter referring to methylation at other positions. Inspection of the overlapping HpaII and HaeIII sites in the polymer shows that normal methylation of either site with the respective specific methylase introduces abberant methylation in the other site. We were thus able to study the effects of abberant methylation on restriction.

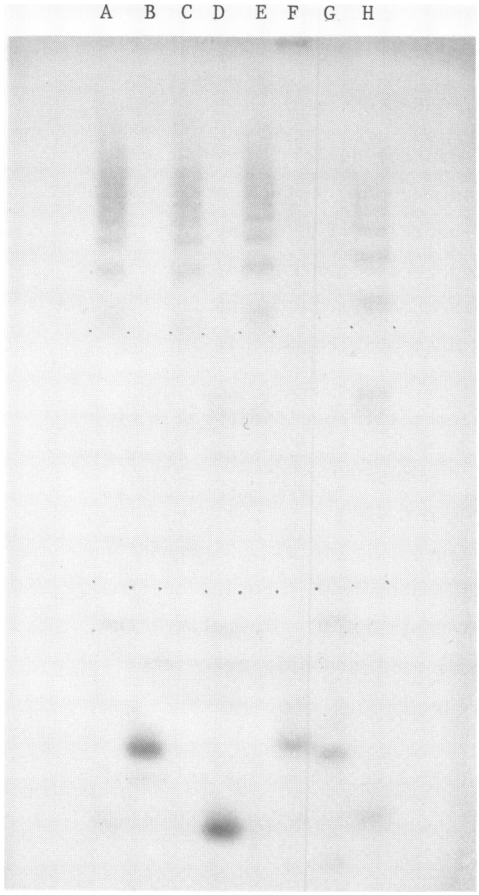

Figure 4. Effect of methylation on subsequent restriction.

Polymerized [^{32}P]decamer was methylated using either M•HpaII or M•HaeIII in a reaction mixture (20 µl) containing 50 mM Tris-Cl (pH 7.5), 20 mM 2-mercaptoethanol, 10 mM EDTA, 80 µM S-adenosylmethionine, 20 pmoles of [^{32}P]polymer, and 20 units of methylase. Incubation was for 3 hours at 37°. The reaction mixture was then divided into 4 aliquots and digested with restriction enzymes in reaction mixtures (20 µl) constituted as in the legend to Figure 2. After 15 min incubation at 37°, the samples were lyophilized, dissolved in 20 µl of formamide, heated 2 min at 100°, mixed with dye markers and loaded on a 18 cm x 40 cm 20% polyacrylamide gel (no urea). Electrophoresis was carried out at 400 V until the bromphenol blue marker was at 23 cm after which an autoradiogram was obtained. Slots A-D, M•HpaII methylated; shots E-H, HaeIII methylated. A and E, no restriction; B and F, R•BamI digested; C and G, R•HpaII digested; D and H, R•HaeIII digested. A number of digestion products are present which presumably represent folded structures, the analysis of which is beyond the scope of this paper.

Labeled polymer was methylated using M•HpaII to give the methylated substrate . . . C-C^{m}-G-G-C-C^{m}-G-G . . . containing normally methylated HpaII
. . . G-G-C-C-G-G-C-C . . .
sites and an abberantly methylated HaeIII site. Aliquots were then treated with restriction enzymes and analyzed by gel electrophoresis (Figure 4, Slots A-D). R•HpaII was unable to cleave the modified sites (slot C), but R•BamI and R•HaeIII gave complete cleavage (slots B and D). Thus,

. . . C-C^{m}-G-G . . .
. . . G-G-C-C . . . sites are protected against HpaII restriction, but

. . . G-G-C-C^{m} . . .
. . . C-C-G-G . . . sites are not protected against HaeIII restriction.

Similarly, prior methylation of polymer by M•HaeIII protected against R•HaeIII cleavage (slot H), but not against R•BamI and R•HpaII cleavage

(slots F and G) (Figure 4, slots E-H). Thus, . . . G-G-$\overset{m}{C}$-C . . . is protected against HaeIII cleavage, but . . . $\overset{m}{C}$-C-G-G . . . is sensitive to R•HpaII cleavage.

. . . C-C-G-G . . .
. . . G-G-C-C . . .

Interpreted in terms of protein-DNA contact during "recognition", these results suggest that the restriction enzymes, R•HpaII and R•HaeIII, make close contact to the major groove 5-position of cytosine only for the cytosines that are normally methylated by the corresponding modification enzymes. In both cases, these positions are adjacent to the cleavage point so that simple steric hindrance is a tenable hypothesis for the mechanism of protection.

A final deduction from our data concerns the specificity of a hypothetical BamI methylase. We have shown that when the external cytosines of the BamI recognition site, (5') G-G-A-T-C-$\overset{m}{C}$, are methylated as
C-C-T-A-G-$\overset{m}{G}$ (5')
indicated, BamI cleavage still occurs (Figure 4, Slot B). Observations from Pirrotta (9) and Wilson and Young (10) show that internal adenine methylation does not protect against BamI restriction. Therefore, the modifying enzyme (if it is a DNA methylase) most probably methylates the internal cytosines.

ACKNOWLEDGEMENTS

We are indebted to Drs. Narang, Bahl, and Wu for making this work possible by supplying the synthetic decamer. The work was supported by NSF grant no. PCM74-19985 and NIH grant no. CA 16519. M.B.M. was a recipient of a Public Health Service Research Fellowship Award no. AI01731.

REFERENCES

1 Middleton, J.H., Edgell, M.H., and Hutchinson, C.A.III (1972) J. Virol. 10, 42-50.
2 Sharp, P.A., Sugden, B., and Sambrook, J. (1973) Biochemistry 12, 3055-3063.
3 Old, R., Murray, K., and Roizes, G. (1975) J. Mol. Biol. 92, 331-339.
4 Garfin, D.E. and Goodman, H.M. (1974) Biochem. Biophys. Res. Commun. 59, 108-116.
5 Kelly, T.J. and Smith, H.O. (1970) J. Mol. Biol. 51, 393-409.
6 Stawinski, J., Hozumi, T., Narang, S.A., Bahl, C.P., and Wu, R. (1977) Nuc. Acids Res. 4, 353-371.
7 Bahl, C.P., Marians, K.J., Wu, R., Stawinski, J., and Narang, S.A. (1976) Gene 1, 81-92.
8 Maxam, A.M., and Gilbert, W. (1977) Proc. Nat. Acad Sci. USA 74, 560-564.
9 Pirrotta, V. (1976) Nucl. Acids Res. 3, 1747-1760.
10 Wilson, G.A. and Young, F.E. (1975) J. Mol. Biol. 97, 123-125.
11 Maniatis, T., Kee, S.G., Efstratiadis, A., and Kafatos, F.C. (1976) Cell 8, 163-182.

12 Randerath, K. and Randerath, E. (1976) In Methods in Enzymology (Grossman, L. and Moldave, K., eds.) vol. 12, part A, pp. 323-347, Academic Press, New York.
13 Sulkowski, E. and Laskowski, M. (1971) Biochem. Biophys. Acta 240, 443-447.