Volume 11 Number 11 1983 Nucleic Acids Research

The sequence GGC^{m}CGG is resistant to *Msp*I cleavage

M.Busslinger*, E.deBoer, S.Wright, F.G.Grosveld and R.A.Flavell+

Laboratory of Gene Structure and Expression, NIMR, Mill Hill, London NW7 1AA, UK, *Institut für Molekularbiologie II, University of Zürich, Zürich, Switzerland and +Biogen Research Corp., 241 Binney Street, Cambridge, MA 02142, USA

Received 22 March 1983; Accepted 11 May 1983

ABSTRACT

MspI essentially fails to cut the sequence GGC^{m}CGG at enzyme concentrations which give total digestion of CCGG, C^{m}CGG and GGCCGG sites. This result explains why certain sites in mammalian DNA are resistant to both MspI and HpaII and shows that this results from an idiosynchracy of MspI rather than a novel form of DNA methylation at this site in mammalian cells.

INTRODUCTION

The analysis of DNA methylation in eukaryotic cells has been greatly aided by the use of restriction enzymes which are sensitive to the presence of 5 methyl Cytosine (^{m}C) residues in their recognition site. Thus, Bird and Southern (1) showed that HhaI (GCGC) and HpaII (CCGG) would not cut Xenopus rDNA if the internal C residue was methylated and that some, but not all of these sites were methylated in genomic rDNA. The failure of these enzymes to cut at a given site indicated the presence of ^{m}C in that sequence or the absence of that site in the DNA because of a polymorphic difference in primary DNA sequence. It became easier to discriminate between these two possibilities for CCGG sites when Waalwijk and Flavell (2) showed that MspI would cut at CCGG, irrespective of ^{m}C at the internal C residue; MspI could, therefore, be used to show the presence of that site and HpaII to diagnose the presence of ^{m}C residues.

An apparent exception to this rule was fortuitously observed in a study of the methylation of the human β-related globin genes (3). They showed that MspI failed to cut at two CCGG sites present in the 5' flanking regions of the Gγ- and Aγ-globin genes, respectively. Since this site was cut after cloning the γ-globin DNA in a phage vector (4), this effect had to be explained by modification of the DNA in human cells. Furthermore, since it was shown (5) that MspI could not cut the sequence ^{m}CCGG, it was suggested that the sites in the human γ-globin genes were modified at the external C residue (3, 4). Others have also noted such sites (8). An alternative explanation would be,

however, that those exceptional MspI sites are a subset of the CCGG sites which are uncut when methylated at the internal C residue as a result of some unique property of that site. We show here that these unique sites have the consensus sequence GGCCGG and that when the internal C residue is methylated (GGC^{m}CGG) MspI cuts this site with great difficulty. That MspI does not cut at these sites is therefore an interesting enzymological artefact and does not necessarily indicate $^{m}C^{m}CGG$ sites in mammalian DNA.

MATERIALS AND METHODS

Enzymes and reagents

Restriction enzymes, E. coli DNA polymeraseI, HpaII methylase and T4 DNA ligase were purchased from N.E.Bio Labs. α-^{32}P-dATP (3000Ci/mMol) and α-^{32}P-dCTP (300 Ci/mMol) was from the Radiochemical Centre, Amersham, 5-methyl dCTP from P-L Biochemicals, dATP, dCTP, dTTP from Boehringer Mannheim and dextransulfate from Pharmacia.

M13 cloning and in vitro DNA methylation

The human Aγ-globin gene was isolated on a 3300bp long HindIII fragment from cosmid clone HG25 (9) and ligated into the HindIII site of the replicative form of phage M13mp8 (10). The human α1-globin gene was inserted into M13mp8 as a 1500bp long PstI fragment isolated from DNA clone pRBα1 (11). These ligated DNAs were used to transfect E.coli JM103 and single stranded DNA of the recombinant phages named Mγ1 and Mα1, respectively, was isolated according to Messing and Vieira (10). These phage DNAs were used as templates for the in vitro synthesis of hemimethylated DNA essentially as described by Stein et al. (6) (Fig. 1). An M13-specific DNA primer (SP16) was heat-denatured and added to a reaction mixture (100μl) containing 2μg of single stranded phage DNA, 66mM Tris-HCl (pH7.5), 6.6mM MgCl2, 10mM dithiothreitol and 50μM each of rATP, dATP, dGTP, dTTP and 5 methyl dCTP (5-methyldeoxycytidine-5'-triphosphate). The complementary strand was synthesized at 30^{o}C for 1hr. with E.coli DNA polymeraseI (25 units) in the presence of T4 DNA ligase (400 units). The DNA was then extracted with phenol and separated on a 0.8% low-melting agarose gel in the presence of EtBr (0.5μl/ml). The covalently closed circular DNA molecules were isolated from the agarose by phenol extraction and ethanol precipitation and used directly for transformation of L-cells. Methylation of DNA with HpaII methylase was performed according to the instructions of the manufacturer (N.E.Biolabs) and methylation with Bacillus phage methylase essential as in Jentsch et al.(5). Under the conditions used (1hr. incubation) only the outer C residue of CCGG

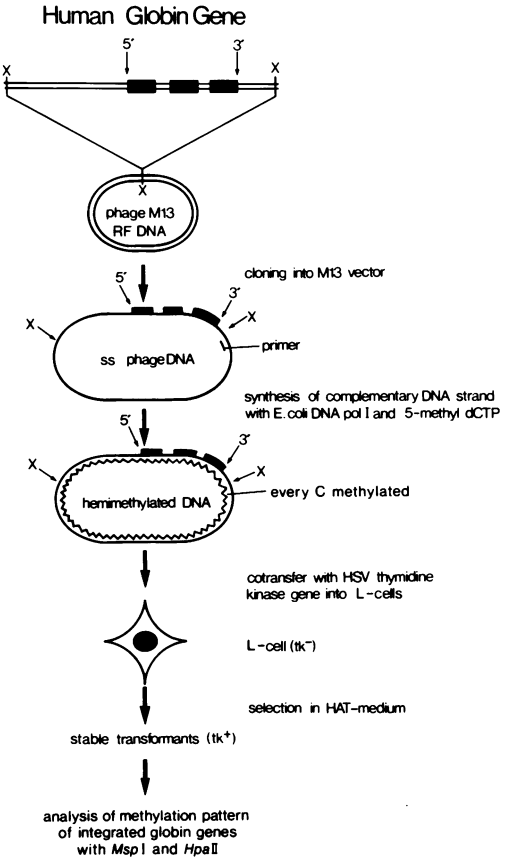

Fig. 1. Procedure used to study the inheritance of DNA methylation.
X denotes the restriction enzyme used for M13 cloning, i.e. PstI for the α1-globin gene and HindIII for the Aγ-globin gene; the resulting phages are named Mα1 and Mγ1, respectively. For experimental details see Methods section.

is methylated. This will be described in detail elsewhere (U. Gunthert et al., in preparation).

Cell Culture and Transformation

Ltk$^-$ mouse cells obtained from R. Axel were maintained in Dulbecco's modified Eagle's (DME) medium supplemented with 5% new born calf serum. These cells were transformed with the plasmid pTKLM176 which contains the gene of herpes simplex virus type I inserted into pBR322 (T. Lund and A. Mellor, unpublished data). 40ng of pTKLM16 DNA, together with 100-200ng of hemimethylated Mα1 DNA and 20μg of salmon sperm DNA were added as calcium phosphate coprecipitate to each petri dish containing 10^6 L-cells. In some experiments

unmethylated replicative form DNA of Mα1 or Mγ1 were used as cotransforming DNAs instead of hemimethylated DNA. Transformants (tk^+) were selected in DME medium supplemented with 5% new born calf serum and hypoxanthine, aminopterin and thymidine (HAT). Individual colonies were picked and grown into mass cultures.

Restriction and Hybridization Analysis

DNA of transformed cells was isolated as described by Stein et al.,(6). 10-50μg DNA was digested with a 5 to 10 fold excess of restriction endonuclease in the buffer recommended by the suppliers. After addition of the enzyme, a sample (0.5μg) of the DNA digest was removed and incubated with 0.5μg of phage λDNA, which allowed us to monitor the digestion. DNA fragments were electrophoresed on 1% or 1.5% agarose gels in Loening E buffer and then blotted onto nitrocellulose filters. The filter strips were hybridized overnight at 65°C to ^{32}P-labelled Aγ-globin, α1-globin or M13 DNA probes in 2 X SSC, 10% dextran sulfate, 0.1% SDS, and 10 X Denhardt's solution containing 20μg/ml of sheared mouse liver DNA. Post-hybridization washes were carried out at 65°C in 1 X SSC, 0.1% SDS (M13 DNA probes) or in 0.1 X SSC, 0.1% SDS (globin DNA probes) for 1-2 hrs. All DNA probes were labelled by nick-translation with α-^{32}P-dATP or α-^{32}P-dCTP to a specific activity of 10^8 cpm/μg. Complete digestion with the methylation-sensitive enzymes MspI and HpaII was controlled by adding a small amount of λDNA as internal control to the digestion mixture. The restriction pattern of this λDNA was made visible by hybridization of nick-translated λDNA to the blots previously hybridized with globin of M13 DNA probes.

RESULTS AND DISCUSSION

MspI does not cut certain CCGG sites in methylated human globin genes

To determine whether the MspI resistant cleavage sites have a specific DNA sequence, we carried out a search for such sites in the human α- and γ-globin genes and in the M13 viral cloning vector. The two globin genes were cloned into M13 and the single stranded DNA of the recombinant phages (Mα1 and Mγ1) was used as template for DNA synthesis using DNA polymeraseI, a short oligonucleotide as primer and as substrates dATP, TTP, dGTP and d^mCTP instead of normal dCTP. In this way, a duplex DNA methylated in one strand is produced. Stein et al. (6) showed that the unmethylated template strand of this hemimethylated DNA is methylated in vivo when this DNA is introduced into L cells by DNA mediated gene transfer. The methylated residues are predominantly or exclusively at ^mCpG sequences. The methylation pattern

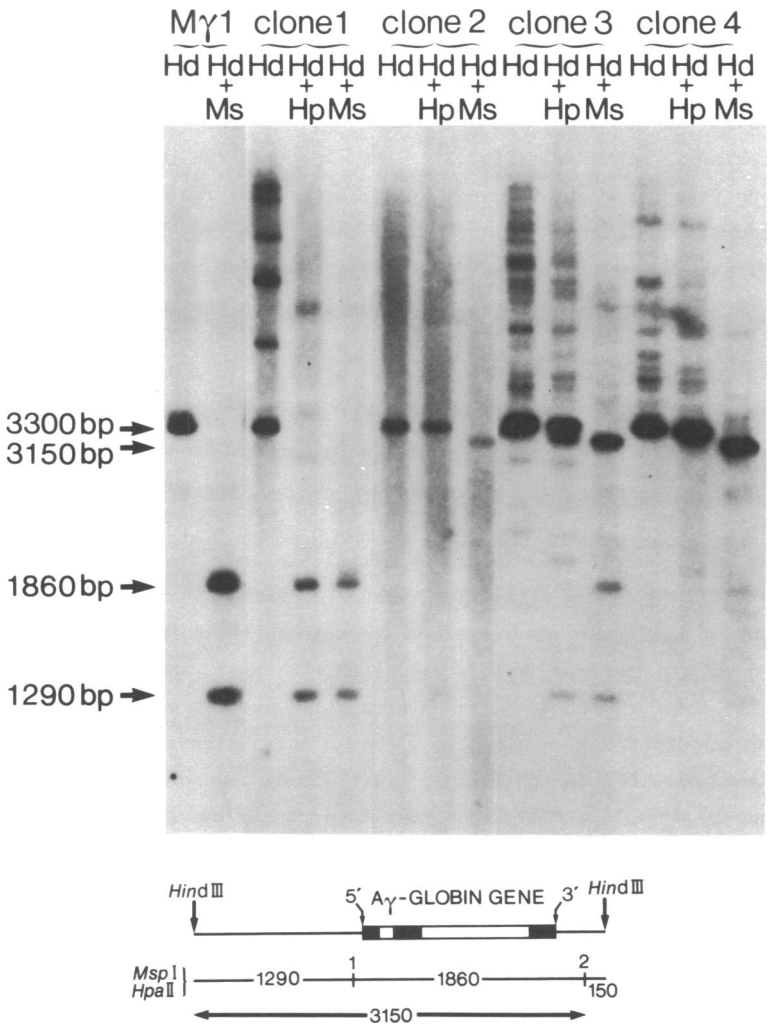

Fig. 2. MspI does not cut the 5' CCGG site of the Aγ-globin gene.
High molecular weight DNA of L-cells transformed with unmethylated (clone 1) or hemimethylated (clones 2, 3, 4) Mγ1 DNA was digested with the restriction enzymes indicated: HindIII (Hd), HpaII (Hp) and MspI (Ms). DNA fragments were electrophoresed on a 1% agarose gel, transferred to nitrocellulose and hybridized to the nick-translated 3300bp HindIII fragment of the Aγ-globin gene. Mγ1 denotes the M13 clone containing the γ-globin HindIII insert and was used as a size marker. The relevant restriction map of the γ-globin gene is shown in the bottom part with the sizes of the corresponding restriction fragments indicated.

thereby established is inherited faithfully from one cell generation to the next and can therefore be studied in transformed cell lines. DNA from L cells (clone 2, 3, 4) transfected with hemimethylated Mγ1 DNA was cleaved with HindIII (to release the human DNA insert from the vector(, HindIII plus HpaII or HindIII plus MspI (Fig. 2). As expected, the two CCGG sites present on this fragment are uncut by HpaII, showing that both sites are methylated at the internal C residues of the CCGG site. MspI cleaves the CCGG site present on the 3' side of the γ-globin gene, but fails to cut the site to the 5' side of the gene, resulting in a DNA fragment of 3150bp (Fig. 2). Since the same site is uncut (3) in human DNA from most tissues (Fig. 2), it follows that the 'inheritance' of this type of modification in this transformation system is similar to that of other ^{m}CpG's and that we can use this method to find other such sites. When unmethylated γ-globin DNA was used in the transformations (clone 1), both CCGG sites were cut with MspI and HpaII.

We next checked the cloned human α-globin gene (Mα1) for CCGG sites uncut by MspI. One such site was detected (number 16 in Fig. 3). DNA from transformed L cells was cut with PstI (to release the 1.5kb α-globin DNA insert) and either HpaII or MspI. Digestion with HpaII shows essentially no cleavage of the PstI fragment in DNA from cells (clones 6, 7, 8) transformed with hemimethylated α-globin genes, although total cleavage of this fragment was obtained with DNA for cell lines transformed with unmethylated α-globin DNA (clone 5). All 16 CCGG sites therefore retain the internal ^{m}C residue. Since there are so many CCGG sites in the PstI fragment cloned in Mα1, cleavage with MspI yields small fragments, most of which run off the gel showing in Fig. 3. However, two MspI fragments can be seen of 490bp and 460bp, respectively. The latter fragment results from cleavage at site 16. This site is partially resistant to MspI digestion resulting in the larger 490bp fragment. Digestion of DNA from unmethylated α-globin DNA (either cloned DNA or DNA from transfected cells) gives only the 460bp fragment. That the difference between methylated and unmethylated DNA is real is shown by mixing cloned α-globin DNA with the DNA from the L cells transfected with hemimethylated α-globin genes. Here two poorly resolved bands are found instead of a single band (Fig. 3).

M13 DNA also contains MspI-resistant CCGG Sites

Finally, we screened M13 vector DNA for CCGG sites which are resistant to cleavage by MspI. Two such sites are found at residues 2552 and 7007, respectively. To do this we purified the 818bp DNA fragment which flanks the

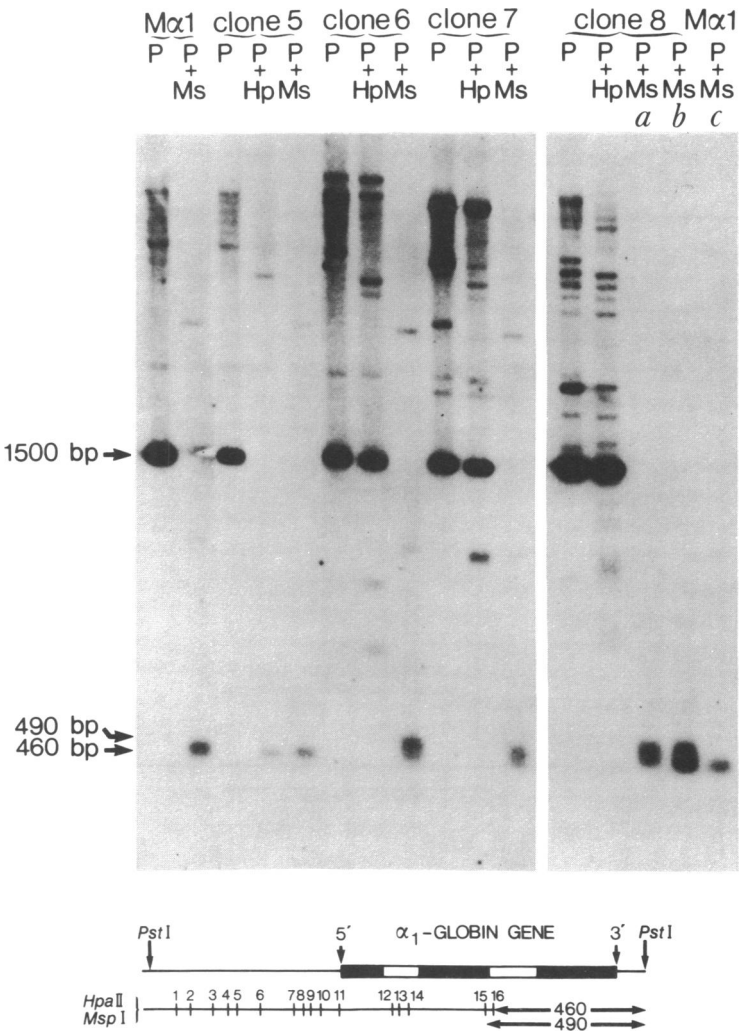

Fig. 3. Cleavage in the human α1-globin gene with MspI.
High molecular weight DNA of L-cells transformed with unmethylated (clone 5) or hemimethylated (clone 6, 7, 8) Mα1 DNA was digested with the restriction enzymes indicated: PstI (P), HpaII (Hp) and MspI (Ms). DNA fragments were separated on 1.5% agarose, transferred to nitrocellulose and hybridized to the nick-translated 1500bp long PstI fragment of the α-globin gene. Mα1, the M13 recombinant containing the α-globin PstI insert, was used as a size marker. The HpaII restriction map of the α-globin gene, shown in the bottom part, was determined by partial restriction mapping of the cloned PstI fragment. The HpaII sites are numbered from left to right with the size of the largest restriction fragment indicated. In one case (clone 8) we added MspI digested cloned Mα1 to the cut L cell DNA (in lane b) to show the resolution of the two bands.

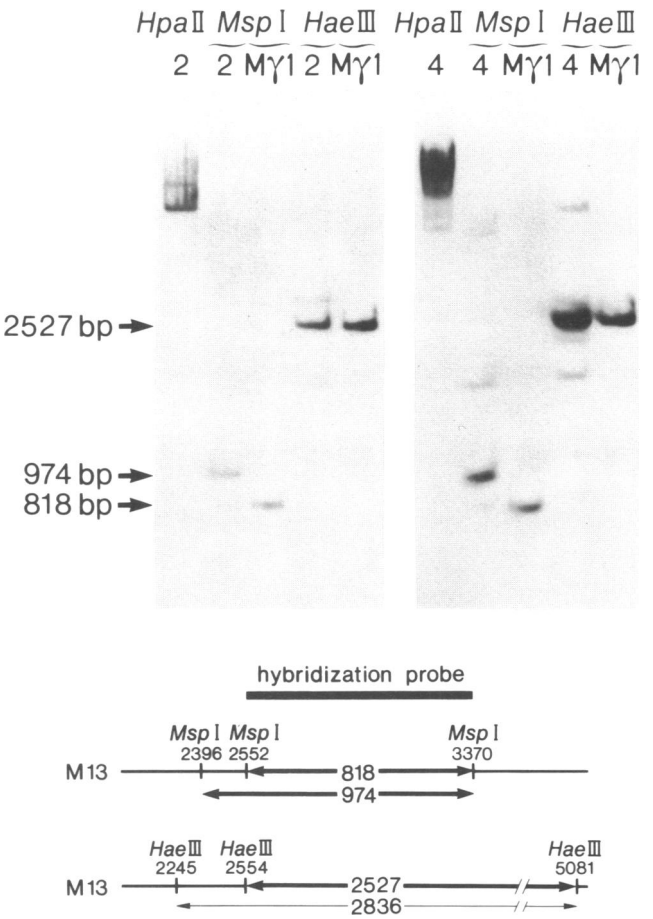

Fig. 4. Failure of MspI to cut a CCGG site of phage M13.
DNA from L cells transformed with hemimethylated Mγ1 (clones 2 and 4 of Fig. 2) was cut with HpaII, MspI or HaeIII, electrophoresed on an agarose gel and hybridized to the 818bp MspI fragment shown in the scheme.

site at residue 2552 from M13 RF DNA and used this as a probe in our Southern blots of DNA from L cells transformed with hemimethylated Mγ1 DNA (clones 2 and 4). The 818bp fragment is however not seen upon MspI cleavage. Instead, a fragment of about 970bp is seen which results from the failure of MspI to cut at site 2552 (Fig. 4). We have performed the similar experiments with the MspI site at residue 7007 with identical results (not shown).

We have aligned the DNA sequences of the four MspI-resistant sites in Fig. 5 (13, 14, 15). It can be seen that they have in common the sequence

*Msp*I #1 ($^{A}\gamma$) TGAGGCCAGGGGCCGGCGGCTGGCTA
*Msp*I #16 (α_1) GTGAGCGGCGGGCCGGGAGCGATCTG
*Msp*I 2552(M13) CCATTAGCAAGGCCGGAAACGTCACC
*Msp*I 7007(M13) GGGTGAGAAAGGCCGGAGACAGTCAA

Fig. 5. Alignment of the MspI sequences uncut by MspI in 'methylated' DNA.

(Pu)GGCCGG. Since the consensus sequence of the site contains overlapping HaeIII(GGCC) sites, this permitted a direct test of the methylation status of the two C residues of this sequence. If both C residues are methylated, then the site will be resistant to cleavage by HaeIII (blocked by GG^{m}CC) as well as MspI (blocked by ^{m}CCGG). This GG↓CCGG site at position 2554 of M13 is however cut by HaeIII in the DNA from the L cells transformed with hemimethylated DNA despite the fact that it is resistant to MspI; a 2557bp HaeIII fragment is seen rather than the 2836bp fragment which would be expected if the site were uncut (see Fig. 4).

Modification with HpaII Methylase Blocks MspI Cleavage at (Pu)GGCCGG

Since the four sites have a common DNA sequence and because of the HaeIII results, we were concerned that the failure of MspI to cut at this site could derive simply from the presence of the sequence GGC^{m}CGG, that is, a site with only a single ^{m}C residue. We therefore used HpaII methylase to methylate the sequence of Mγ1 and then digested the DNA with MspI, HpaII and HaeIII. MspI cleaves the unmethylated DNA much more readily than the methylated DNA. In addition, MspI only partially cuts at the MspI site at the 5' side of the γ-globin gene at enzyme amounts that give complete cleavage at other sites; at extremely high enzyme doses (200X excess) about 50% cleavage is found. Similar results are found for the M13 GGCCGG sites. In contrast, MspI cuts the CCGG sites 3' to the γ-globin gene and the remaining CCGG sites in the M13 vector to completion. As expected, HpaII fails to cut all CCGG sites (Fig. 6a). This result suggests that either MspI is blocked by the single ^{m}C residue, or the HpaII methylase modifies both C residues.

To exclude the latter possibility we cleaved methylated DNA with HaeIII and asked whether the GGC^{m}CGG site of M13 was cut by HaeIII. If the sequence was GG^{m}C^{m}CGG, HaeIII cannot cut this site (12). HaeIII, however, cleaves the *in vitro* methylated DNA to completion (Fig. 6b). To establish beyond doubt that the sequence $\begin{matrix} GG^{m}CCG\ G \\ CC\ GGC^{m}C \end{matrix}$ is not cut by HaeIII we used DNA methylase of Bacillus phage SPR19 which gives this modification pattern on short incubation times (U. Gunthert, unpublished). This methylated sequence is not cut by HaeIII (not shown).

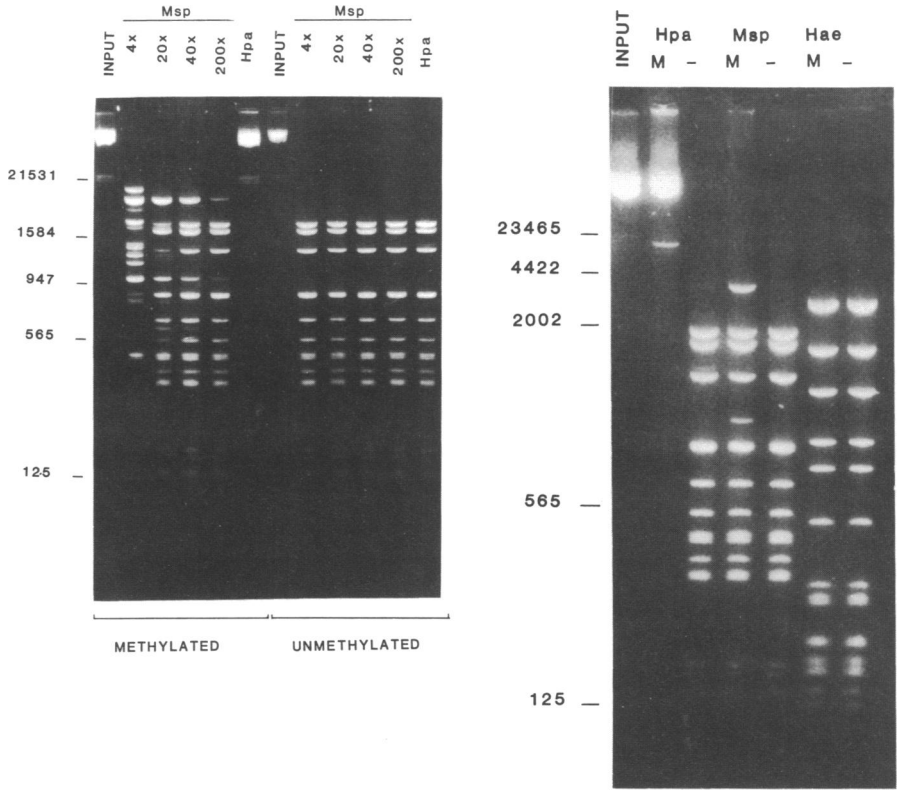

Fig. 6. Cleavage of Mγ1 DNA methylated in vitro with HpaII methylase.
Panel a, the left hand five lanes show the MspI digestion of methylated Mγ1 DNA up to 200 units/μg of DNA. Lane 6 shows the HpaII digest of methylated DNA. The right hand five lanes show the same MspI digests of unmethylated Mγ1 DNA. The last lane shows the HpaII digest of unmethylated DNA.
Panel b, shows the input DNA followed by HpaII, MspI and HaeIII digests of methylated (M) and unmethylated (-) Mγ1 DNA.

The results show that MspI fails to cut DNA at the sequence GGC^{m}CGG. It is theoretically possible that HpaII methylase generates the sequence $\begin{matrix} GGC^{m}C\ G\ G \\ CCG\ G^{m}C^{m}C \end{matrix}$ at this specific site even though the only modification detected up to now with the enzyme is $\begin{matrix} C^{m}C\ GG \\ G\ G^{m}CC \end{matrix}$. We consider this highly unlikely. The methylated strand of the M13 DNA sequence introduced into the animal cells in all our experiments was $\begin{matrix} GG^{m}C^{m}CGG \\ CC\ G\ GCC \end{matrix}$ and our data show that the external ^{m}C residue is lost after passaging through the cells ($\rightarrow \begin{matrix} GGC^{m}C\ GG \\ CCG\ G^{m}CC \end{matrix}$): it is implausible that this loss is accompanied by a transfer of the methyl group to the external C residue of the other DNA strand to give $\begin{matrix} GGC^{m}C\ C\ C \\ CCG\ G^{m}C^{m}C \end{matrix}$. Moreover, the fact

that we have observed the same phenomenon with both eukaryotic and bacterial methylase makes this explanation still more unlikely.

It is not clear why MspI has such difficulty cleaving this site when the internal C residue is methylated. ^{m}C residues have been implicated in the transition of DNA from the B to Z configuration (7) and it is possible that the site GGCCGG forms a non-B configuration (under MspI incubation conditions) when the internal C residue is methylated. Whatever the explanation, this phenomenon is clearly of some practical importance in the study of DNA methylation since such sites cannot be characterized with MspI. HaeIII and HpaII are, however, diagnostic for the presence of methyl group's sequence GGCCGG.

H. Cedar and his colleagues have independently examined this phenomenon (see this issue of NAR) and have also concluded that the internal ^{m}C residue prevents cleavage of these sites by MspI.

ACKNOWLEDGEMENTS

We are indebted to Dr. U. Gunthert for performing the experiment with the Bacillus phage DNA methylation. M. Busslinger was the recipient of a fellowship from the Schweizerischer National Fonds. We are grateful to H. Cedar for helpful discussions, to M. L. Birnstiel for providing the opportunity to M.B. to finish this work in his laboratory, and to F. Ochsenbein for graphical work. This work was supported by the British Medical Research Council.

REFERENCES

1. Bird, A. and Southern, E. M. (1978) J. Mol. Biol., 118, 27-48.
2. Waalwijk, C. and Flavell, R. A. (1978) Nucl. Acids Res., 5, 3231-3236.
3. van der Ploeg, L. H. T. and Flavell, R. A. (1980) Cell, 19, 947-958.
4. van der Ploeg, L. H. T., Groffen, J. and Flavell, R. A. (1980). Nucl. Acids Res., 8, 4563-4574.
5. Jentsch, S., Gunthert, U. and Trautner, T. A. (1981) Nucl. Acids Res., 9, 2753-2759.
6. Stein, R., Gruenbaum, Y., Pollack, Y., Razin, A. and Cedar, H. (1982) Proc. Natl. Acad. Sci. U.S.A., 79, 61-65.
7. Behe, N. and Felsenfeld, G. (1981) Proc. Natl. Acad. Sci. U.S.A., 78, 1619-1623.
8. Sanders, Haigh, L., Blanchard Owens, B., Hellewell, S. and Ingram, V. M. (1982) Proc. Natl. Acad. Sci. U.S.A., 79, 5332-5336.
9. Grosveld, F. G., Lund, T., Murray, E. J., Mellor, A. L., Dahl, H. H. M. and Flavell, R. A. (1982) Nucl. Acids Res., 10, 6715-6732.
10. Messing, J. and Vieira, J. (1982) Gene, 19, 269-276.
11. Lauer, J., Shen, C. K. and Maniatis, T. (1980) Cell, 20, 119-130.
12. Mann, N. B. and Smith, H. O. (1977) Nucl. Acids Res., 4, 4211-4221.
13. Slightom, J. L., Blechl, A. E. and Smithies, O. (1980) Cell, 21, 627-638.
14. Michelson, A. M. and Orkin, S. H. (1980) Cell, 22, 371-377.
15. van Weezenbeck, P. N. T. F., Hulsebos, T. J. M. and Schoenmakers, J. G. G. (1980) Gene, 11, 129-148.