Cytosine methylated DNA synthesized by Taq polymerase used to assay methylation sensitivity of restriction endonuclease *Hin*fI

Joseph Colasanti and Venkatesan Sundaresan*
Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

Received May 30, 1990; Revised and Accepted December 4, 1990

ABSTRACT

We have studied the resistance of cytosine methylated DNA to digestion by the restriction endonuclease *Hin*fI, using a simple PCR procedure to synthesize DNA of known sequence in which every cytosine is methylated at the 5 position. We find that *Hin*fI cannot digest cytosine methylated DNA at the concentrations normally used in restriction digests. Complete digestion is possible using a vast excess of enzyme; under these conditions, the rate of *Hin*fI digestion for cytosine methylated DNA is at least 1440-fold slower than for unmethylated DNA. The presence of an additional methylated cytosine at the degenerate position internal to the recognition sequence does not appear to increase the resistance to *Hin*fI digestion. We also tested *Hha*II, an isoschizomer of *Hin*fI, and found that it is completely inactive on cytosine methylated DNA. The procedure we have used should be of general applicability in determination of the methylation sensitivities of other restiction enzymes, as well as studies of the effects of methylation on gene expression in direct DNA transfer experiments.

INTRODUCTION

Many restriction endonucleases have been shown to be incapable of digesting DNA if a cytosine residue within the recognition sequence is methylated at the 5 carbon atom of the pyrimidine ring (reviewed in ref. 1). The effect of 5-methylation of cytosine residues on restriction enzyme activity has been previously studied using hemimethylated ϕX174 DNA synthesized *in vitro* (2), or using bacteriophage XP12 DNA in which every cytosine is replaced by 5-methylcytosine (m^5C; ref. 3). The restriction endonuclease *Hin*fI recognizes the sequence GANTC (reviewed in ref. 4) and has been shown to completely digest hemimethylated ϕX174 DNA (2). Further, it appears that *Hin*fI will also digest the fully methylated DNA from phage XP12, although at a rate that is 10 times slower than an unmethylated substrate DNA (1,3). However, because the sequence of XP12 is not known, and there is no m^5C-insensitive isoschizomer of *Hin*fI, it is not possible to determine if the XP12 DNA is completely digested, or whether there are some *Hin*fI sites within the XP12 sequence that remain undigested. 5-methylation of cytosines is frequent in many higher eukaryotes, and is often correlated with the inactivation of genes, or of transposable elements (reviewed in ref. 5). It has been shown that the inactivation of the *Mu1* transposon in maize correlates with resistance of the transposon sequence to *Hin*fI digestion (6,7,8), suggesting that either *Hin*fI will not digest certain sequences containing m^5C, or that the modified base is other than m^5C, for example 6-methyladenine (m^6A; ref. 1). For these reasons it is of interest to test the activity of *Hin*fI on a cytosine methylated substrate of known DNA sequence.

MATERIALS AND METHODS

Enzymes. *Hin*fI was obtained from Promega Corporation (Madison, WI, U.S.A.) and was supplied at a concentration of 50 units/μl where 1 unit is defined as the amount of enzyme required to completely digest 1 μg of phage λ DNA in 1 hour at 37°C. *Hha*II was a gift of Dr. Hamilton Smith (Johns Hopkins University), and was supplied at a concentration of 1000 units/μl. *Sau*3AI and *Mbo*I were purchased from New England Biolabs (Beverly, MA, U.S.A.).

Synthesis of substrate DNA. The plasmid pUC19 (9) was used as the template for synthesis by the Polymerase Chain Reaction (PCR; ref. 10). Two 20 bp oligonucleotides L(5'CTGCAGGCATGCAAGCTTGG3') and R (5'CTCGCCTTGATCGTTGGGAA3') at positions 435–454 and 1995–2014 respectively of the pUC19 sequence (9) were used to amplify a 1.58 kbp sequence (Fig. 1). The pUC19 plasmid was linearized by BamHI digestion and approximately 0.1 to 40 ng linear pUC19 DNA (see below) was used in a 100 μl reaction containing a standard buffer of 50 mM KCl, 10 mM Tris HCl pH 8.3, 1.5 mM $MgCl_2$, 0.01% gelatin with 0.4 μM of each primer and 0.8 units of AmpliTaq polymerase (Perkin-Elmer). Deoxyribonucleotide triphosphates dATP, dGTP, dTTP and dCTP or dm^5CTP were purchased from Boehringer Mannheim Biochemicals, and used at a final concentration of 0.2 mM. The reaction conditions for PCR synthesis were 95°C 40″, 63°C 2′, 72°C 5′ (see below) for 20 cycles. Amplification of template DNA using dm^5CTP instead of dCTP was inefficient and

* To whom correspondence should be addressed

required several modifications to synthesize larger quantities of DNA. First, a large amount of input template was used; i.e. 40 ng per reaction, as compared to 0.1 to 0.5 ng for a standard reaction. Second, an extension time of at least 5 min was required for significant accumulation of product. Yield of the amplified product was decreased when larger amounts of Taq polymerase or greater concentrations of nucleotides were used, although the parameters for maximum efficiency of amplification have not been extensively studied. In addition a time course experiment showed that, in reactions with m^5C, increasing the number of cycles does not lead to a corresponding increase of product. For the above reaction maximum yield of specific product was obtained between 15 and 20 cycles and a greater number of cycles caused increased amounts of non-specific products (not shown). Attempts to amplify the entire 2.7 kbp pUC19 sequence using appropriate primers were not successful, suggesting that there may be an upper limit to the length of the methylated DNA that can be synthesized. For the synthesis of 'demethylated' DNA, 0.1 ng of the methylated product was used in a standard PCR reaction with unmethylated cytosine. The input 1.58 kb methylated DNA was first eluted from an agarose gel to avoid contaminating unmethylated 2.7 kb template from the original amplification. Methylated template appeared to amplify as efficiently as unmethylated template.

Conditions for restriction digests. Three types of PCR derived DNA were used in restriction reactions: i) unmethylated DNA from a PCR synthesis with unmethylated nucleotides, ii) methylated product which incorporated 5-methylcytosine in the synthesis and iii) 'demethylated' DNA, where methylated DNA was used as template and unmethylated cytosine was used in the synthesis. Approximately 50 ng of each type of DNA was used in each restriction digest. For *Sau*3AI and *Mbo*I, 6–8 units of enzyme were used in each reaction for 2 h representing a 240 to 320 fold overdigestion. In the *Hin*fI and *Hha*II digestions 50 ng of DNA was digested with 2 units of enzyme for 2.5 h ('100-fold overdigestion'), or 40 units for 5 h ('4000-fold overdigestion'), or 80 units for 10 h ('16,000-fold overdigestion'). All digestions were performed in a 20 µl volume of a standard buffer of 100 mM NaCl, 10 mM Tris pH 7.4, 10 mM $MgCl_2$, 5 mM 2-mercaptoethanol and 100 mg/ml bovine serum albumin at 37°C. Dilutions of the enzyme stocks were performed in the reaction buffer.

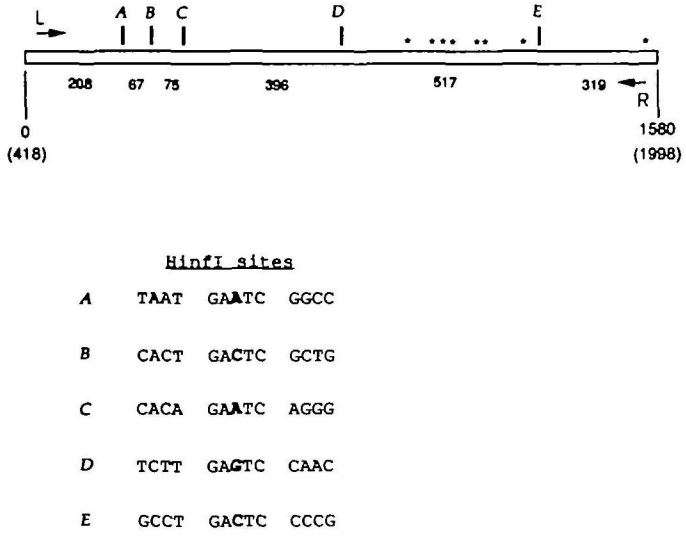

Figure 1. Amplified region of pUC19 showing the location of *Hin*fI/*Hha*II sites (*A*–*E*). The sequence of each site and flanking regions are indicated below. The variable nucleotide of each site is indicated in bold. The sizes of the *Hin*fI/*Hha*II restriction fragments, in base pairs, are indicated below the line. The asterisks above the line show the relative position of *Sau*3AI/*Mbo*I restriction sites. Primers L and R (see text) used in the amplification are represented by arrows. The numbers at the ends of the bar represent the distance from one primer site to the other and the numbers in parenthesis show the published coordinates for these sites on pUC19 (ref. 9).

RESULTS

To verify that the PCR reaction using dm^5CTP produced a pUC19 sequence in which C was replaced by m^5C, we digested the PCR product with *Sau*3A which is sensitive to m^5C methylation of its GATC recognition sequence and with *Mbo*I, an isoschizomer of *Sau*3A which is not affected by m^5C methylation (1,4). If the 1580 bp PCR product is digested by *Sau*3A or *Mbo*I, we expect to see disappearance of the 1580 bp DNA fragment and the appearance of a 939 bp fragment and several smaller fragments. As shown in Fig. 2, the methylated DNA (lane 4) was completely resistant to *Sau*3A digestion (lane 5) but not to Mbo1 digestion (lane 6). The unmethylated DNA synthesized using normal dCTP (lane 1) was digested by both enzymes (lanes 2 and 3). This confirmed that the Taq polymerase incorporated the methylated cytosine nucleotides correctly into the DNA.

We then repeated the enzyme digestions using the same substrates and the enzyme *Hin*fI, or the isoschizomer *Hha*II (4). Both *Hin*fI and *Hha*II are sensitive to m^6A, and the corresponding specific adenine methylase has been identified for *Hha*II but not for *Hin*fI (reviewed in ref. 1). The results are shown in Fig. 3. The unmethylated PCR product was completely digested by moderate excess (100-fold overdigestion) of either

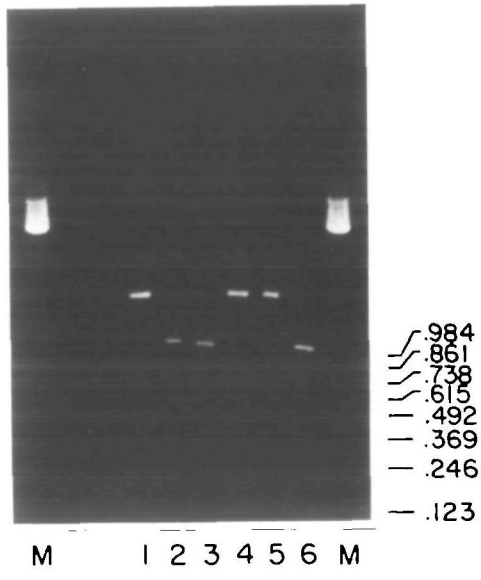

Figure 2. *Sau*3A and *Mbo*I digests of unmethylated and methylated 1580 bp PCR products electrophoresed on 1.8% agarose gel and stained with Ethidium Bromide. Lanes 1–3, unmethylated DNA: undigested (lane 1), digested with *Sau*3A (lane 2), or *Mbo*I (lane 3). Lanes 4–6, m^5C methylated DNA: undigested (lane 4), digested with *Sau*3A (lane 5), or *Mbo*I (lane 6). Lanes M are DNA size-markers (123 bp ladder).

enzyme (lanes 2–4). By contrast, the methylated substrate was completely resistant to *Hin*fI digestion under the same conditions (lane 6). We then further increased the amounts of enzyme and the time of digestion in order to determine the extent of resistance of the m^5C pUC19 DNA to *Hin*fI. We were able to get partial digestion at 4000-fold overdigestion with *Hin*fI (lane 7) but not with *Hha*II (lane 9), and complete digestion at 16000-fold overdigestion with *Hin*fI (lane 8); *Hha*II digestion at the latter concentration resulted in non-specific degradation presumably due to star activity (not shown). A comparison of lanes 7 and 8 in Fig. 3 shows that one of the final products of *Hin*fI digestion, a 319 bp fragment, is overrepresented in the partial digest suggesting that the corresponding methylated site is less resistant to *Hin*fI (see Discussion). To rule out the possibility of mutations in the methylated PCR product that could be errors generated by the Taq polymerase (10), we re-amplified the methylated PCR product using normal dCTP. This 'de-methylated' substrate was then digested with *Hin*fI and *Hha*II. As seen in Fig. 3 (lanes 10 and 11), both *Hin*fI and *Hha*II are able to digest this DNA under conditions of 100-fold overdigestion, making it unlikely that the resistance of the methylated substrate was due to mutations introduced by the Taq polymerase.

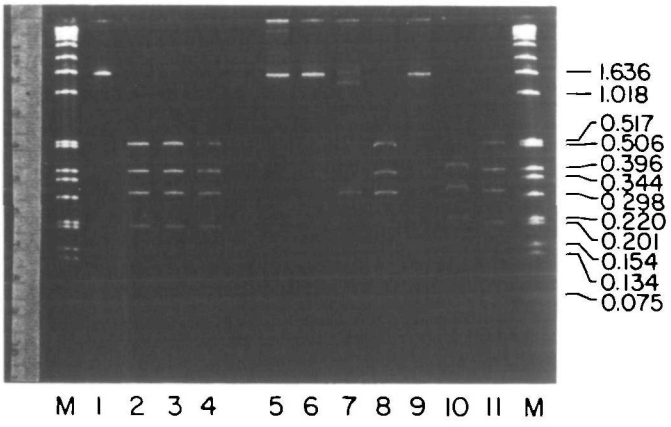

Figure 3. *Hin*fI and *Hha*II digests of unmethylated, methylated and demethylated 1580 bp PCR products electrophoresed on a 6% polyacrylamide gel and stained with Ethidium Bromide. Lanes 1–4, unmethylated DNA: undigested (lane 1), digested to 100 fold excess with *Hin*fI (lane 2), 4000 fold excess *Hin*fI (lane 3), 100 fold excess *Hha*II (lane 4). Lanes 5–9, m^5C methylated DNA: undigested (lane 5), digested 100 fold excess with *Hin*fI (lane 6), 4000 fold excess *Hin*fI (lane 7), 16000 fold excess *Hin*fI (lane 8), 4000 fold excess *Hha*II (lane 9). Lanes 10 and 11 show the 'demethylated' DNA (see text) digested to 100 fold excess with *Hin*fI (lane 10) or *Hha*II (lane 11). Lanes M are DNA size markers (kilobase ladder + *Hin*fI digested pBR322 DNA).

The rates of *Hin*fI digestion of methylated and unmethylated DNA were estimated for the highest enzyme:substrate ratio used, i.e. 80 units of enzyme/50 ng DNA (Fig. 4). For cytosine methylated DNA we estimate that, after 4 hours of incubation, more than one half of the *Hin*fI restriction sites are cleaved (Fig. 4A,lane 3). For estimating the rate of digestion of unmethylated DNA under the same conditions, the *Hin*fI enzyme was pre-incubated at 37° for 4 hours in the reaction mixture without DNA in order to approximate the activity of the enzyme at the point of the methylated DNA restriction where the reaction is more than 50% complete. Under these conditions the unmethylated DNA is digested very quickly, with more than one half of the *Hin*fI sites restricted in less than 10 seconds. While a precise estimation of the rate of digestion of unmethylated DNA for this extremely high enzyme-substrate ratio cannot be easily made, it is apparent that the amount of digestion of unmethylated DNA in 10 seconds is greater than the amount of digestion of cytosine methylated DNA in 4 hours (14,400 seconds) *i.e.*, the rate of digestion for cytosine methylated DNA is at least 1440 times slower than that for unmethylated DNA.

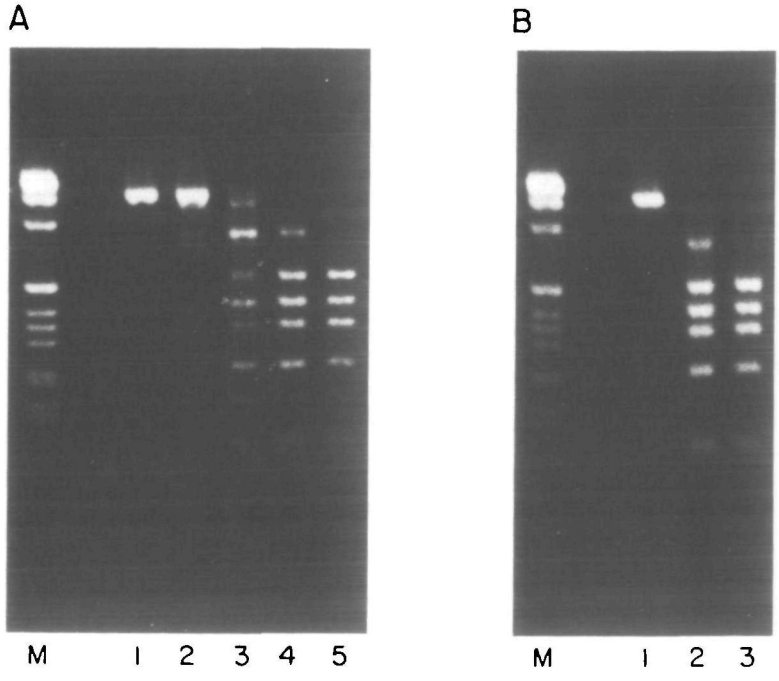

Figure 4. Comparison of the rates of digestion of methylated and unmethylated DNA by *Hin*fI at the maximum enzyme:substrate ratio (80 units enzyme to 50 ng. DNA). A, Lanes 1–5 show restriction of cytosine methylated DNA after 0, 2, 4, 8 and 16 hours of digestion, respectively. B, Unmethylated DNA that is uncut (lane 1) or digested for 10 seconds (lane 2) and 30 seconds (lane 3) with *Hin*fI pre-incubated for 4 hours at 37°C.

DISCUSSION

We have shown that it is possible to synthesize long (1.58 kbp) DNA sequences in which m^5C is substituted for C, and that they are correctly discriminated by the isoschizomers *Sau*3A and *Mbo*I. There is no detectable digestion of m^5C pUC19 by *Hin*fI at 100-fold excess, and complete digestion is achieved only at 16000-fold excess. A comparison of the rates of digestion for the methylated and unmethylated substrates at the highest enzyme-substrate ratio used suggested that *Hin*fI digests cytosine methylated DNA at a rate that is more than 1440 times slower than the rate for unmethylated DNA. We did not observe significant star activity of *Hin*fI under our reaction conditions even in the most extreme overdigestion i.e. 10 hour incubation with a *Hin*fI concentration of 4 units/μl, 4% glycerol and 100 mM NaCl. Under these conditions very little star activity is observed in assays using human genomic DNA (11).

At an intermediate level of digestion (4000-fold excess) we are able to visualize partial products (Fig. 3, lane 7), which are of interest in the determination of relative sensitivities of the different *Hin*fI sites. The sizes of the major DNA bands in this lane are approximately 1.6 kb, 1.3 kb, and 0.32 kb, representing the full length DNA and partial digestion at a terminal *Hin*fI site (site *E* in Fig. 1). The presence of the 1.3 kb product and absence of any visible 1.4 kb, 1.1 kb, or 0.2 kb bands suggests that site *E* is favoured over the other terminal site, site *A*. A comparison of sites *A* and *E* shows that site *E* includes a C at the internal degenerate position of the GANTC recognition sequence, while site *A* does not (Figure 1). Thus the presence of an additional m^5C within the GANTm^5C sequence does not increase the resistance of site *A* over site *E*. Further, an examination of the flanking DNA shows that site *E* is flanked by G-C rich sequences on both sides, while site *A* is flanked by an A-T rich sequence on one side and a G-C rich sequence on the other side. Therefore the presence of flanking m^5C residues also does not account for the increased resistance of site *A* over site *E*. There are two minor products in the partial digest (Fig. 3, lane 7) of sizes 0.8 kb and 0.5 kb approximately. These are likely to be due to further digestion of the 1.3 kb partial product at site *D*, which also contains an internal m^5C at the N position of GANTC (Fig. 1), again confirming that methylation of this C does not confer increased resistance to *Hin*fI. We conclude that the reason for the differential sensitivities of the *Hin*fI sites is not obviously related to m^5C methylation, but is due to other effects of the flanking DNA sequences. To summarize, 5-methylation of the terminal C in GANTC increases (by over 1440-fold) the resistance to digestion by *Hin*fI, but GAm^5CTm^5C is not more resistant than GAATm^5C, as might be expected if the enzyme does not interact with the residue at the degenerate N position of GANTC. We find similarly that the synthetic m^5C DNA is also resistant to specific cleavage by *Hha*II, the isoschizomer of *Hin*fI; however, in the case of *Hha*II, this resistance could not be overcome by increasing the extent of digestion. Therefore 5-methylation of the final C in GANTC absolutely interferes with its interaction with *Hha*II.

The above procedure for assaying the effects of m^5C methylation on restriction enzyme activity should be of general applicability. It has advantages over the use of XP12 DNA in that (i) the sequence of the methylated DNA is known, as are the sizes of the predicted restriction fragments (ii) Any sequence can be methylated in the context of its flanking DNA and (iii) the preparation of the methylated DNA is quick and easy. This method also has advantages over the use of synthetic methylated oligonucleotides in two respects. Firstly, the same methylated DNA can be used to test a large number of different restriction enzymes. For example, the 1.540 kbp cytosine methylated pUC19 sequence that we have synthesized using PCR contains recognition sequences for over 66 different restriction enzymes, not including isoschizomers (9). Secondly, prior knowledge of the recognition sequence of a restriction enzyme is not required for testing its sensitivity to methylation; it is only necessary than the enzyme can cleave the unmethylated substrate at least once. The same PCR procedure can also be used in principle to generate methylated sequences for studies on the effects of methylation on gene regulation in direct DNA transfer experiments (11,12).

ACKNOWLEDGEMENTS

We would like to thank Dr. S. Chandrasegaran and Dr. Hamilton Smith (Johns Hopkins University), for the *Hha*II restriction enzyme, Dr. Rich Roberts for advice, and Sadie Arana for preparation of the manuscript. This research was supported by grants from NSF (DCB-8702318) and Pioneer Hi-Bred International, and a post-doctoral fellowship to J.C. from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Nelson, M. and McClelland, M. (1989) *Nuc. Acids Res.* **17**, r389–r415.
2. Gruenbaum, Y., Cedar, H. and Razin, A. (1981) *Nuc. Acids Res.* **9**, 2509–2515.
3. Huang, L.-H., Farnet, C.M., Ehrlich, K.C. and Ehrlich, M. (1982) *Nuc. Acids Res.* **10**, 1579–1590.
4. Roberts, R.J. (1989) *Nuc. Acids Res.* **17**, r347-r388.
5. Holliday, R. (1987) *Science* **238**, 163–170.
6. Chandler, V.L. and Walbot, V. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9631–9635.
7. Walbot, V. (1986) *Genetics* **114**, 1293–1312.
8. Martienssen, R., Barkan, A., Taylor, W.C. and Freeling, M. (1990) *Genes Devel.* **4**, 331–343.
9. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103–119.
10. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R.G., Horn, G.T., Mullis, K.B. and Erlich, H.A. (1988) *Science* **239**, 487–491.
11. Kris, J., Kerrin, G., Forman, L. and Cotton, R. (1990) *Nuc. Acids Res.* **18** (in press).
12. Wigler, M., Levy, D. and Perucho, M. (1981) *Cell* **24**, 33–40.
13. Yisraeli, P., Frank, D., Razin, A. and Cedar, H. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4638–4642.