Volume 12 Number 13 1984 Nucleic Acids Research

Alteration of apparent restriction endonuclease recognition specificities by DNA methylases

Michael Nelson, Chantal Christ and Ira Schildkraut

New England Biolabs, Inc., 32 Tozer Road, Beverly, MA 01915, USA

Received 5 June 1984; Accepted 12 June 1984

ABSTRACT

An in vitro method of altering the apparent cleavage specificities of restriction endonucleases was developed using DNA modification methylases. This method was used to reduce the number of cleavage sites for 34 restriction endonucleases. In particular, single-site cleavages were achieved for Nhe I in Adeno-2 DNA and for Acc I and Hinc II in pBR322 DNA by specifically methylating all but one recognition sequence.

INTRODUCTION

Modification of DNA by site-specific DNA methylases has been observed to block DNA cleavage by restriction endonucleases (1-5). For example, DNA isolated from E. coli K12 is methylated at GATC sequences by the dam methylase and is resistant to cleavage by Mbo I (GATC), and resistant to cleavage by Cla I (ATCGAT) only at those Cla I recognition sequences which are preceded by guanine (GATCGAT) or followed by cytosine (ATCGATC). We describe how DNA methylases can be used in vitro to create new DNA cleavage specificities. A DNA methylase is selected whose recognition sequence overlaps only a subset of the recognition sites of a given restriction endonuclease: only those methylated subsets will be resistant to cleavage by that restriction endonuclease.

Two classes of useful overlaps can be described. The first class consists of an overlap between recognition sequences of a restriction endonuclease which recognizes a degenerate sequence, and a methylase which acts on only one of the subsets of the degenerate sequence. An example of this class is Hinc II restriction endonuclease in combination with M.Taq I methylase. Hinc II recognizes the degenerate sequence GTPyPuAC which represents the four sequences, GTCGAC, GTCAAC, GTTGAC, and GTTAAC. McClelland (6) has previously shown that M.Taq I recognizes the sequence TCGA, methylates the adenine residue, and blocks the cleavage of the sequence GTCGmAC by Hinc II. Therefore Hinc II sequences which contain the internal sequence TCGA will

be resistant to cleavage by Hinc II after methylation by M.Taq I, but all other Hinc II recognition sequences (GTCAAC, GTTGAC, and GTTAAC) will be cleaved. This resulting specificity can be represented as the sequence GTPyAAC, or equivalently, GTTPuAC.

The second class of overlap can be described as an overlap which occurs at the boundaries of the recognition sequences of a restriction endonuclease and a methylase. In this case, only some fraction of restriction endonuclease recognition sequences will be bounded by specific nucleotides, which together with the nucleotides of the endonuclease recognition sequence also define a methylase recognition sequence. An example of this class is the overlap of the recognition sequences of dam methylase (GATC) and Cla I restriction endonuclease (ATCGAT), described above.

It should be noted that overlapping methylase/endonuclease recognition sequences do not ensure that new specificities will result. The effect of methylation at residues other than the cognate methylation site is generally unknown. Several endonucleases are known to cleave DNA methylated within their recognition sequence. For example, BamH I, which recognizes GGATCC, is not blocked by the overlapping dam methylation at GGmATCC. Therefore, a restriction endonuclease recognition sequence that has been modified by a methylase with an overlapping recognition sequence must be tested for resistance to the restriction endonuclease cleavage.

The altered apparent specificities are generated using a two-step in vitro procedure: 1) methylation of DNA by a site-specific methylase, followed by 2) cleavage of the DNA by a restriction endonuclease with an overlapping recognition sequence. Using this method we demonstrate a single-site cleavage for Nhe I in Adeno-2 DNA, and single-site cleavages for Acc I and Hinc II in pBR322 DNA. Furthermore, 44 methylase/restriction endonuclease combinations were tested and determined to result in altered specificities.

MATERIALS AND METHODS

Methylases and restriction endonucleases were isolated and purified in this laboratory. All DNA's except Adeno-2 DNA were prepared in this laboratory. Adeno-2 DNA was a gift from R. Roberts. S-adenosylmethionine-HCL (SAM) is from Sigma. The SAM is stored at -20°C as a 30 mM solution in sulfuric acid(0.005M):ethanol, (9:1, v:v). One unit of methylase is defined as the amount of enzyme required in one hour in a 10 ul reaction volume to completely protect one ug of lambda DNA from digestion with excess cognate restriction endonuclease. In order to determine if methylation

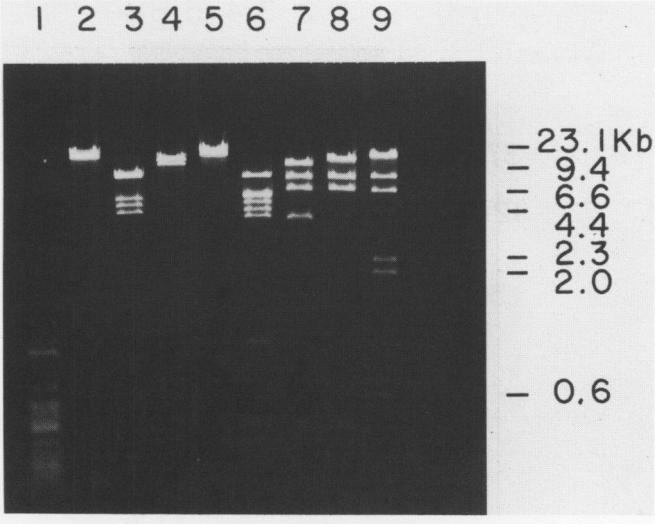

Figure 1. The effect of methylating Adeno-2 DNA with M.Alu I methylase followed by cleavage with Nhe I. Adeno-2 DNA was treated with Alu I (lane 1), Nhe I (lane 3), or untreated (lane 5). Adeno-2 DNA was methylated with M.Alu I (lanes 2 and 4) and treated with Alu I (lane 2) or Nhe I (lane 4). DNA in lanes 6-8 was treated identically to that in lanes 3-5, respectively, except that the DNA was also cleaved by Sal I in order to more easily visualize the Nhe I cleavages. Lane 8 is the molecular weight standard, Hind III digested Lambda DNA.

at a specific site resulted in resistance to cleavage by a restriction endonuclease whose recognition sequence overlapped that site, the following method was employed: 1-2 ug of DNA were incubated with 1-5 units of purified methylase in 10 ul methylase assay buffer (generally 50 mM Tris-HCl, pH 7.5, 10 mM EDTA, 5 mM 2-mercaptoethanol, and 80 uM S-adenosylmethionine) at 37°C for 1-3 hours. M.Taq I methylations were incubated at 65°C. The methylation reactions were terminated by heat treatment for twenty minutes or phenol extraction. 40-60 ul of restriction endonuclease assay buffer with excess magnesium (20 mM) was added to each reaction. The reaction was then incubated with 5-10 units of restriction endonuclease for one hour. The resulting DNA fragments were analyzed by agarose gel electrophoresis.

RESULTS

Figure 1 demonstrates the effect on Adeno-2 DNA of first methylating with M.Alu I methylase (AGCT) followed by cleavage with Nhe I. Nhe I is a type II restriction endonuclease which recognizes the sequence GCTAGC(7). M. Alu I

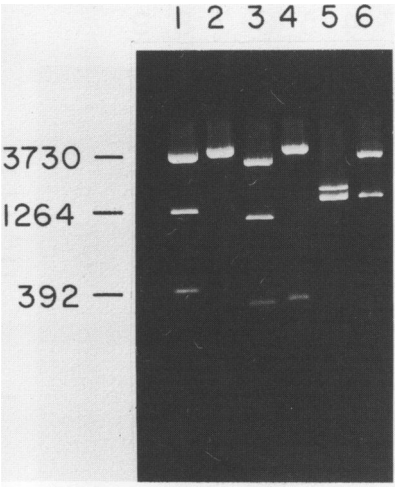

Figure 2. The effect of methylating pBR322 DNA with M.Taq I followed by digestion with Hinc II or Acc I. Lane 1 is the molecular weight standard of PhiX174 DNA digested with Hpa I. Pst I digested (linearized) pBR322 DNA is the substrate for lanes 2 - 6. DNA with no further treatment (lane 2). DNA digested with Hinc II (lane 3). DNA methylated by M.Taq I followed by digestion with Hinc II (lane 4). DNA digested with Acc I (lane 5). DNA methylated by M.Taq I followed by digestion with Acc I (lane 6).

blocks three (AGCTAGC) of the four Nhe I sequences in adeno-2 DNA (positions 10799, 25733, and 31497). The cleavage specificity of Nhe I on M.Alu I methylated adeno-2 DNA is a single cleavage at map position 20701, or 57.6% of the genome.

Figure 2 demonstrates the effect on pBR322 DNA of first methylating with M.Taq I methylase followed by cleavage with either Hinc II or Acc I. The pBR322 substrate was first linearized with Pst I to more easily visualize the cleavage by HincII and Acc I. Of the two Hinc II sequences in pBR322, the site in the tetracycline resistance gene, position 651 (GTCGAC), is blocked by M.Taq I methylation. This results in a single cleavage by Hinc II in the ampicillin resistance gene, at position 3907(GTCAAC). Similarly M.Taq I blocks the Acc I site at position 651(GTCGAC), generating a single Acc I cleavage at position 2246 (GTATAC).

In addition to the three altered specifities demonstrated above, Tables 1 and 2 list all the methylase/endonuclease combinations we have tested, which have generated new cleavage specificities. DNA banding patterns for some of these new cleavage specificities are depicted in Figure 3. Table 1 lists those combinations which overlap at the boundaries of the recognition

TABLE 1. DNA CLEAVAGE SPECIFICITIES GENERATED AS A RESULT OF METHYLATION AT THE BOUNDARIES OF OVERLAPPING RECOGNITION SEQUENCES OF A RESTRICTION ENDONUCLEASE AND A METHYLASE

RESTRICTION ENDONUCLEASE	METHYLASE[a]	SUBSET BLOCKED BY METHYLATION[b]
AhaII [GPuCGPyC]	M.HpaII [CmCGG]	CCGGCGPyC
AluI [AGCT]	M.PstI [CTGCmAG]	AGCTGCAG
AvaII [GG(AT)CC]	M.HpaII [CmCGG]	CCGG(AT)CC
BamHI [GGATCC]	M.MspI [mCCGG]	CCGGATCC
BglI [GCCN$_5$GGC]	M.HaeIII [GGmCC]	GGCCN$_5$GGC
BstXI [CCAN$_6$TGG]	M.HaeIII [GGmCC]	GGCCAN$_6$TGG
ClaI [ATCGAT]	Dam [GmATC]	GATCGAT
DdeI [CTNAG]	M.AluI [AGmCT]	AGCTNAG
EcoRV [GATATC]	M.TaqI [TCGmA]	TCGATATC
FnuDII [CGCG]	M.HhaI [GmCGC]	GCGCG
HinfI [GANTC]	M.HphI [TmCACC][c]	GANTCACC
HinfI [GANTC]	M.TaqI [TCGmA]	TCGANTC
HphI [GGTGA]	Dam [GmATC]	GGTGATC
MboI [GATC]	M.ClaI [ATCGmAT]	ATCGATC
MboI [GATC]	M.TaqI [TCGmA]	TCGATC
MboII [GAAGA]	Dam [GmATC]	GAAGATC
MspI [CCGG]	M.BamHI [GGATmCC]	CCGGATCC
MspI [CCGG]	M.HaeIII [GGmCC]	GGCCGG
NaeI [GCCGGC]	M.HaeIII [GGmCC]	GGCCGGC
NcoI [CCATGG]	M.HaeIII [GGmCC]	GGCCATGG
NheI [GCTAGC]	M.AluI [AGmCT]	AGCTAGC
NruI [TCGCGA]	Dam [GmATC]	GATCGCGA
PstI [CTGCAG]	M.AluI [AGmCT]	AGCTGCAG
SacII [CCGCGG]	M.HaeIII [GGmCC]	GGCCGCGG
Sau3AI [GATC]	M.HphI [TmCACC]	GATCACC
TaqI [TCGA]	M.ClaI [ATCGmAT]	ATCGAT
TaqI [TCGA]	Dam [GmATC]	GATCGA
XbaI [TCTAGA]	Dam [GmATC]	GATCTAGA
XmnI [GAAN$_4$TTC]	M.TaqI [TCGmA]	TCGAAN$_4$TTC

a) Nomenclature for methylases from Smith and Nathans, 1973 (8) Methylase specificities from Smith and Kelly, 1984 (3)

b) Sequences blocked by methylation are non-palindromic, and only one strand, 5'-3', is written. For example, the M.Pst I - Alu I sequence that is blocked could also be written CTGCAGCT.

c) Methylation site from Nelson and Feehery, unpublished observations.

TABLE 2. DNA CLEAVAGE SPECIFICITIES GENERATED BY METHYLATION AT A SUBSET OF RECOGNITION SEQUENCES OF DEGENERATE RESTRICTION ENDONUCLEASES

RESTRICTION ENDONUCLEASE	METHYLASE[a]	CLEAVAGE SPECIFICITY[b]
AccI [GT(AC)(GT)AC]	M.TaqI [TCGmA]	GT(AC)TAC
AhaII [GPuCGPyC]	M.HhaI [GmCGC]	GPuCGTC
AvaI [CPyCGPuG]	M.TaqI [TCGmA]	CPyCGGG
AvaI [CPyCGPuG]	M.HpaII [CmCGG]	CPyCGAG
AvaI [CPyCGPuG]	M.HpaII [CmCGG] + M.TaqI [TCGmA]	CCCGAG
BanII [GPuGCPyC]	M.AluI [AGmCT]	GPuGCCC
BanII [GPuGCPyC]	M.HaeIII [GGmCC]	GPuGCTC
BanII [GPuGCPyC]	M.HaeIII [GGmCC] + M.AluI [AGmCT]	GGGCTC
Bsp1286 [G(AGT)GC(ACT)C]	M.AluI [AGmCT]	G(GT)GC(ACT)C
Bsp1286 [G(AGT)GC(ACT)C]	M.HaeIII [GGmCC]	G(AGT)GC(AT)C
Bsp1286 [G(AGT)GC(ACT)C]	M.HaeIII [GGmCC] + M.AluI [AGmCT]	G(GT)GC(AT)C
HgiAI [G(AT)GC(AT)C]	M.AluI [AGmCT]	G(AT)GCAC
HincII [GTPyPuAC]	M.TaqI [TCGmA]	GTPyAAC
Sau96I [GGNCC]	M.HaeIII [GGmCC]	GGACC
ScrFI [CCNGG]	M.HpaII [CmCGG]	CCAGG

a) Nomenclature for methylases from Smith and Nathans, 1973 (8) Methylase specifities from Smith and Kelly, 1984 (3)

b) Cleavage specificities are interpreted as the following:
   1) If more than one base is indicated at a position within the sequence, the enzyme will recognize sequences with any of those bases at that site. For example Acc I will recognize the sequences GTATAC, GTCGAC, GTAGAC, and GTCTAC.
   2) Most sequences are non-palindromic, and only one strand, 5'-3', is written. For example, the M.TaqI - HincII specificity, GTPyAAC, can also be written GTTPuAC, and represents the sequences GTTAAC, GTCAAC, and GTTGAC.

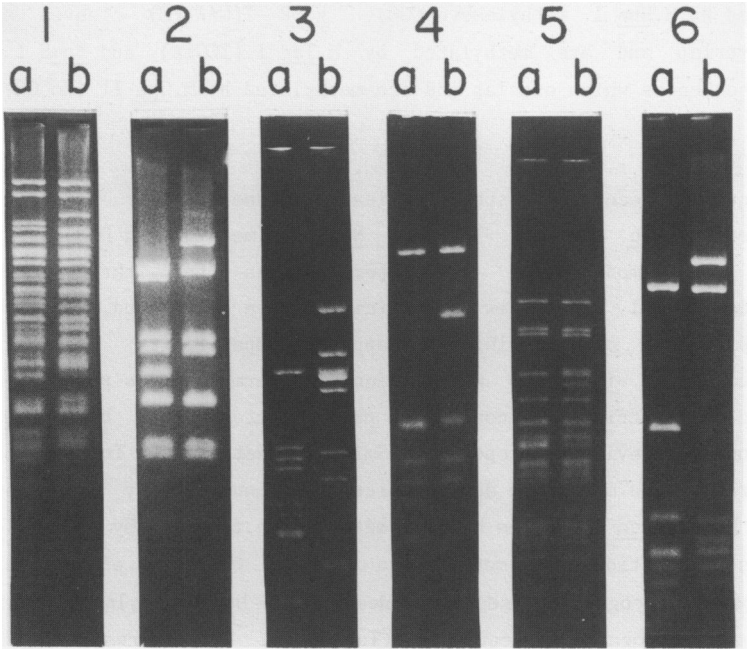

Figure 3. Altered DNA cleavage patterns resulting from methylation prior to restriction endonuclease cleavage. Adeno-2 DNA (lane 1) treated with Ava II alone (a) or M.Hpa II and Ava II (b). PhiX174 DNA (lane 2) treated with Dde I alone (a) or M.Alu I and Dde I(b). T7 DNA (lane 3) treated with FnuD II alone (a) or M.Hha I and FnuD II (b). pBR322 DNA (lane 4) treated with Hinf I alone (a) or M.Hph I and Hinf I (b). Lambda DNA (lane 5) treated with Mbo I alone (a) or M.Taq I and Mbo I (b). pBR322 DNA (lane 6) treated with Sau96 I alone (a) or M.Hae III and Sau96 I (b).

sequences of the restriction endonuclease and the methylase. Table 2 lists methylase/endonuclease combinations in which the methylase overlaps one of the sequences of a restriction endonuclease which recognizes a degenerate sequence. In Table 2 the new apparent specificities can be represented by the conventional format; for example, the apparent specificity of the M.Taq I and Hinc II methylase/nuclease combination is GTPyAAC. In Table 1 this convention does not apply, and the cleavage sequence that is blocked by the methylase, rather than the sequences that are cleaved, is indicated. In all cases, the cleavage sequences generated are non-palindromic and the sequence of only one strand of DNA is indicated. It should be realized that the complimentary sequences are not listed and must be considered.

Some combinations which were tested did not generate new specificities. Sph I (GCATGC) cleaves at recognition sequences which overlap and have been

methylated by M.Hha I methylase(GmCGC). EcoR I(GAATTC) cleaves sequences which overlap and are methylated by M.Taq I (TCGmA), and BamH I (GGATCC) cleaves sequences which overlap and are methylated by M.Hpa II (CmCGG).

DISCUSSION

Unique new recognition specificities were generated for the restriction endonucleases Acc I, Hinc II, and Nhe I by methylating DNA in vitro. In addition to demonstrating new specificities for three restriction endonucleases, 41 other new specificities were determined. It is proposed that the method of altering the apparent specificity of restriction endonucleases by in vitro methylation be termed cross-protection. The methylase and restriction endonuclease pair is referred to by adopting the conventional abbreviation proposed by Smith and Nathans (8) for the methylase, and separating the methylase and restriction endonuclease by a hyphen. For example, M.TaqI-Acc I denotes M.Taq I methylation followed by Acc I cleavage.

Cross-protections by methylation at the boundary of a restriction endonuclease recognition sequence result in hemi-methylated restriction endonuclease recognition sequences (Table 1). Our results indicate, therefore, that hemi-methylation of restriction endonuclease recognition sequences is sufficient to block duplex cleavage by many restriction endonucleases, which is in agreement with the results of Gruenbaum et al. (9).

Although it is known that some restriction endonucleases (BamH I, Sau3A I, Msp I) require methylation at a particular nucleotide residue in order to be blocked, Tables 1 and 2 demonstrate that a significant fraction of the restriction endonucleases can be blocked by other than canonical methylation within their recognition sequence. In particular, Alu I, Ava I, Pst I, Hinf I, and Aha II, can be blocked by methylation at either one of two sites within the recognition sequence. This phenomenon has also been reported for EcoR I, Hha I, Hind III, Sal I, and Xho I (4).

Cross-protection is a method of specifically blocking a subset of cleavage sites of a restriction endonuclease, thereby altering the apparent recognition sequence of the restriction endonuclease. The 44 new cross-protection specificities presented here contribute significantly to the 103 known recognition specificities (10). Furthermore many other specificities should be realized when cross-protection is applied to other methylase/restriction endonuclease combinations. Cross-protection should prove to be a flexible and practical tool for the manipulation of DNA.

REFERENCES
1. Backman,K. (1980) Gene 11, 169-171.
2. Brooks,J.E., and Roberts,R.J. (1982) Nucl. Acids Res. 10, 913-934.
3. Smith,H.O. and Kelly,S.V. (1983) in DNA Methylation, Cedar,H., Riggs,A.D., and Razin,A., eds., Springer Verlag, in press.
4. McClelland, M. (1983) Nucl. Acids Res. 11, r169-r173.
5. Noyer-Wiedner,M., Jentsch,S., Pawleh,B., Gunthert,U. and Trautner,T.A. (1983) J. Virol. 46, 446-453.
6. McClelland, M. (1981) Nucl. Acids Res. 9, 6795-6804.
7. Comb,D., and Schildkraut,I. unpublished results.
8. Smith,H.O., and Nathans,D., (1973) J. Mol. Biol. 81, 419-423.
9. Gruenbaum,Y., Cedar,H., and Razin,A. (1981) Nucl. Acids Res. 9, 2509-2515.
10. Roberts,R.J. (1984) Nucl. Acids Res. 12, r167-r204.