Probing DNA – protein interactions *in vitro* with the CpG DNA methyltransferase

Stefan Kochanek, Doris Renz and Walter Doerfler*
Institute for Genetics, University of Cologne, Weyertal 121, D-5000 Cologne 41, Germany

Received March 5, 1993; Revised and Accepted April 16, 1993

ABSTRACT

A sensitive method was devised to monitor the *in vitro* binding of nuclear proteins from HeLa cells presumably to the major groove of DNA. Upon the incubation of DNA with nuclear extracts, the complexed DNA was incubated with the CpG DNA methyltransferase from *Spiroplasma species*. Subsequently, the DNA was repurified, and the location of the methylated cytidine residues was determined by the hydrazine reaction of the DNA sequencing method. By using as DNA substrate the VAI (virus associated) region of human adenovirus type 2 (Ad2) DNA or specific Alu sequences associated with a number of human genes, it was documented that those segments of DNA that were protected by bound proteins against the reaction with DNase I also escaped *in vitro* methylation by the CpG DNA methyltransferase. This new footprinting method provides a sensitive indicator for *in vitro* DNA – protein interactions which are specific for the major groove of DNA.

INTRODUCTION

Specific DNA–protein interactions are at the core of many reactions important in biology. The methodological repertoire available to date for the detailed analysis of these interactions is still limited. In *in vivo* or *in vitro* footprinting experiments, the enzyme bovine DNase I or the reagent dimethylsulfate (DMS) combined with DNA sequencing technology have been widely used to assess the precise locations of proteins binding to a DNA sequence ladder (1).

DNase I preferentially interacts with the minor groove of double-stranded DNA (2, 3). DNase I cleavage rates are largely dependent on the width of the minor groove (4) and on the bendability or stiffness of DNA (5). Many of the regulatory proteins affecting gene transcription contact DNA bases in the major groove of the double helix. It is thus conceivable that the interference by proteins bound in the major groove of DNA with the access of DNase I targeted towards the minor groove is indirect and might not accurately reflect the actual binding pattern.

DMS on the other hand methylates guanine residues in position 7 in the major groove of the DNA double strand (6). This reaction is directly inhibited by proteins with tight contacts to guanine residues in the major groove. A small molecule like DMS can survey only short regions, hence the method is not very sensitive.

The method described in this report utilizes the CpG DNA methyltransferase (7) which methylates all 5′-CG-3′ dinucleotides in a DNA sequence (8). DNA sequences protected by proteins bound at specific sequence motifs will escape methylation by the enzyme from *Spiroplasma species* and will present as non-methylated cytidines in a chemical sequencing experiment. In contrast, unprotected sequences will be methylated and thus will fail to produce signals in the sequencing ladder. The interference with the CpG DNA methyltransferase reaction seems to provide a sensitive method to determine directly DNA–protein interactions in the major groove and will complement the conventionally used techniques. The method has been tested with proteins from HeLa cell nuclear extracts bound *in vitro* to the VAI region of Ad2 DNA or to Alu elements associated with several human genes.

MATERIALS AND METHODS

DNA templates used for *in vitro* footprinting experiments with the CpG DNA methyltransferase

The plasmid pVAI-II contained a 759 bp fragment from Ad2 DNA carrying the VAI and VAIL segments cloned into pUC18 (8). The plasmid pM14 BglI3 contained a 2.6 kb BglII fragment from the tissue plasminogen activator gene cloned into pBR322 (9). The plasmid pAngio2 contained a 1686 bp PvuII–BglII fragment from the upstream region of the human angiogenin gene (10) cloned into the pBluescript KS+ vector (11). The plasmid pAlu ori contained a 1943 bp HindIII-EcoRI fragment from the α1 globin gene and part of the downstream region cloned into the psVOd vector (12, 13).

Labeling DNA with [32P]

The templates were [32P] labeled either at their 5′ termini by using γ-[32P] ATP and T4 polynucleotide kinase (14) or at their 3′ termini by using [32P]-labeled deoxyribonucleoside triphosphates (dNTPs) for the fill-in reaction with the Klensin fragment (15) of the *E. coli* DNA polymerase (16). Subsequently, the labeled DNA was cleaved with a suitable restriction endonuclease to generate fragments labeled at only one end. The fragments were then purified by electrophoresis on agarose or

* To whom correspondence should be addressed
polyacrylamide gels and eluted. The pVAI DNA was labeled at the HindIII sites and then cut with KpnI. The construct pM14 BglI3 was labeled at the NcoI sites and subsequently cleaved with StuI. pAngio2 was labeled at the NcoI sites and then cleaved with SacI inside the polylinker of the vector. The construct pAlu ori finally was labeled at the BclII sites and cleaved with EcoRI.

**Footprinting reaction with the Spiroplasma species CpG DNA methyltransferase**

The *in vitro* footprinting experiments with the CpG DNA methyltransferase (M. SssI) (2 units/reaction) (7) were performed under conditions of cell-free transcription omitting the ribonucleotide triphosphates, as described (11). The CpG DNA methyltransferase was purchased from New England Biolabs (NEB). The terminally labeled templates (20,000 to 100,000 cpm, Cerenkov) were incubated in a total volume of 10 μl of 10 mM Hepes (N-2-hydroxyethyl)perazine-N'-2-ethanesulfate), pH 7.9, 40 mM KCl, 5 mM MgCl₂, 5 mM creatine phosphate, 400 ng pUC18 DNA and 25 μg of protein from crude nuclear extracts (17) prepared from exponentially growing HeLa cells. After 15 min of incubation at 30°C, 0.5 μl of freshly diluted S-adenosylmethionine (SAM, final concentration 160 μM) and 2 units of the CpG DNA methyltransferase were added, and incubation was continued at 30°C for 45 min.

In control reactions, end labeled DNA was incubated under the same conditions, except that either the DNA methyltransferase was omitted or the template DNA was not complexed with nuclear proteins prior to methylation, and pUC18 DNA as unspecific competitor was omitted.

The reactions were terminated by making the solution 1% SDS, 20 mM EDTA, 250 mM NaCl, and by adding 2 μg of proteinase K and 5 μg of sonicated salmon sperm DNA. After 15 min of incubation at 37°C, the samples were phenol-chloroform extracted and ethanol precipitated. Subsequently, the samples were treated with hydrazine in the presence of 1.5 M NaCl (18) to distinguish between methylated and unmethylated cytosine residues. In detail, the DNA samples were redissolved in 5 μl of H₂O. Subsequently, 15 μl of 5 M NaCl, then 30 μl of hydrazine (Merck) were added. The reaction was allowed to proceed at 20°C for 8 to 15 minutes. Under these conditions hydrazine will modify only the non-methylated cytosine residues which will appear as bands on the sequencing gels. The 5-methyldeoxycytidine residues will not react with hydrazine, and signals on the sequencing gels will fail to be generated in

![Image](image.png)

**Figure 1.** Interference by protein binding in the *in vitro* CpG DNA methyltransferase reaction with the VAI segment of Ad2 DNA after complex formation with nuclear proteins from HeLa cells. (a) Map and summary of data. (i) Map of the Ad2 DNA and of its VAI and VAILI regions; the nucleotide (nt.) numbers refer to the published sequence of Ad2 DNA (21). (ii) Summary of the results of *in vitro* footprinting experiments in the VAI region using DNase I (open bars) or the CpG DNA methyltransferase. Open squares designate unmethylated (protected), closed squares methylated (unprotected) 3'-CG-3' sequences. Symbols above the horizontal line pertain to the upper, symbols below the line to the lower DNA strand. The regulatory regions A and B of the VAI region and the location of the transcribed VAI RNA are also indicated. (b) Autoradiogram of a chemical sequencing experiment performed after the *in vitro* footprinting reaction. The experimental procedures were described under Materials and Methods. The autoradiogram presents results obtained with the bottom strand. The individual lanes contained HindIII-cut DNA templates subsequently treated as follows: 1: DMS (G) reaction; 2: Hydrazine (C) reaction of the DNA *in vitro* methylated by the CpG DNA methyltransferase. 3: Hydrazine reaction of mock-methylated DNA in the absence of SAM. The DNA samples in lanes 1–3 were not complexed with proteins. 4: Template complexed with nuclear proteins from HeLa cells, subsequently treated with CpG DNA methyltransferase followed by the hydrazine (C) reaction (*in vitro* footprint with that enzyme). 5: Control reaction, similar to 4, but without prior incubation with nuclear extract; 6: Template complexed with nuclear proteins from HeLa cells, subsequently treated with DNase I (DNase I *in vitro* footprint). 7: Control DNase I reaction without the incubation with nuclear extracts. Horizontal arrows on the left of the autoradiogram pointed to C residues methylated by the CpG DNA methyltransferase in the presence of nuclear extracts. Circles designated cytidine residues that were protected by protein binding and remained unmethylated. Note bands present in lane 4, but missing in lane 5. A and B referred to the control boxes of the VAI region (see a), the vertical arrow to the site of transcriptional initiation. Asterisks indicated hyperreactive sites in response to nuclease activity in the crude nuclear extracts.
these positions (19). After the reaction with hydrazine, the samples in a total volume of 100 µl were treated for 30 min with 1 M piperidine (Sigma) at 90°C followed by ethanol precipitation and two subsequent lyophilization steps. In parallel experiments, standard in vitro DNase I footprinting experiments were performed (20). DNA fragments generated were resolved by electrophoresis on 5% polyacrylamid sequencing gels containing 7 M urea. The dried gels were autoradiographed to X-Omat AR films for 10 to 20 h.

RESULTS AND DISCUSSION

In the experiments presented, several different DNA templates were first incubated with nuclear proteins from HeLa cells and were then immediately exposed to the CpG DNA methyltransferase from Spiroplasma species (7). The DNA methyltransferase could modify only those 5′-CG-3′ dinucleotides in a sequence that were not protected in the major groove by proteins that had bound to DNA during the first incubation. The DNA templates used in these experiments were listed under Materials and Methods. The results presented were those obtained with the VAI region of Ad2 DNA (Fig. 1) or with the Alu element associated with the human tissue plasminogen activator gene (Fig. 2). The principle of this method was similar to DNase I or DMS footprinting, except that an enzyme was applied as probe that recognized specific dinucleotides and had its target in the major groove of the double helix.

Fig. 1a depicted the map of the VAI region of Ad2 DNA (part i) and included all 5′-CG-3′ dinucleotides in the sequence (part ii). Proteins in nuclear extracts from HeLa cells protected some, but not all of the 5′-CG-3′ dinucleotides from enzymatic methylation (Fig. 1b, lane 4). In the positions designated by arrows, deoxycytidine (C) residues were 5′-CG-3′ methylated as they did not or incompletely react with hydrazine (lane 2). In these positions, C bands were missing or were very weak. Other 5′-CG-3′ dinucleotides were protected and did not become methylated. The hydrazine reaction then produced C bands (circles in Fig. 1b, lane 4). As a control, the standard footprinting reactions with DNase I were also performed (lanes 6, 7). The results (lane 6) as compared to the reaction with unprotected naked DNA (lane 7) revealed that a very similar region in the VAI gene was protected against DNase I and DNA methyltransferase. Additional controls (lanes 1–3, 5) were explained in the legend to Fig. 1b. The schematic in Figure 1a, part ii documented that DNA methyltransferase and DNase I protected regions colocalized to the same area of the regulatory element B on both strands in the VAI segment. At the B box of the VAI region, the polymerase III transcription factor TFIIIC was reported to bind (22–24).

Another DNA segment analyzed in the CpG DNA methyltransferase interference test as described was the Alu element located in intron 8 of the human tissue plasminogen activator gene (9). The map of this element was shown in Fig. 2. Some of the 5′-CG-3′ dinucleotides in this

---

Figure 2. In vitro DNA footprinting experiments using CpG DNA methyltransferase or DNase I with DNA from Alu element 25 in intron 8 of the human tissue plasminogen activator gene. (a) Map and summary of data. (i) Map of part of the tissue plasminogen activator gene (see reference 11). (ii) Summary of the data presented in (b). The symbols and the outlay of the scheme are the same as explained in the legend to Fig. 1a, part ii. (b) In vitro footprinting experiments with Alu element 25 in intron 8 of the human tissue plasminogen activator gene. The autoradiogram shown is derived from the top (lanes 1–7) or the bottom (lanes 8–13) strand of the DNA segment. Individual lanes contain NcoI-cleaved DNA templates subsequently treated as follows. Lanes 2 and 10: Template complexed with nuclear proteins from HeLa cells, subsequently treated with the CpG DNA methyltransferase, followed by the hydrazine (C) reaction (in vitro footprint with that enzyme). Lanes 1 and 11: Control reaction, similar to 2 and 10 but without prior incubation with nuclear extracts. Lane 3: Hydrazine (C) reaction of DNA in vitro methylated by the CpG DNA methyltransferase. Lanes 4 and 9: Hydrazine reaction of mock-methylated DNA in the absence of SAM. Lanes 5 and 8: DMS (G) reaction. Lanes 6 and 12: Template complexed with nuclear proteins from HeLa cells, subsequently treated with DNase I (DNase I in vitro footprint). Lanes 7 and 13: Control DNase I reaction without prior incubation with nuclear extract. Horizontal arrows on the left of the autoradiogram pointed to C residues methylated by the CpG DNA methyltransferase in the presence of nuclear extracts. Circles designated cytidine residues that were protected by protein binding and remained unmethylated. Note bands present in lanes 2 and 11, but missing in lanes 1 and 10. All other symbols were as designated in the legend to Fig. 1b.
Alu element did not become methylated in the course of the interference experiment, i.e. a DNA band was present in lanes 2 and 10 of Fig. 2b in the positions that were protected by HeLa nuclear proteins (circles in the autoradiogram), whereas in the control lanes 1 and 11 of Fig. 2b the corresponding band was missing. Other 5'-CG-3' dinucleotides, however, were not protected and did become methylated by the CpG DNA methyltransferase (arrows in Fig. 2b). Dnase footprinting reactions with DNA complexed to proteins (lanes 6 and 12) or with free DNA (lanes 7 and 13) were also performed. Additional control experiments (Fig. 2b, lanes 3, 4, 5, 8, and 9) were described in the legend to Fig. 2b. The data obtained with this Alu element again documented that the same B box region protected against modification by the CpG DNA methyltransferase was also resistant to Dnase I cleavage, at least to some extent (Fig. 2a, part i).

Very similar results were obtained with two other Alu elements located downstream of the α1 globin gene (12, 13) or upstream of the angiogenin gene (10). In both instances the B box of the Alu elements could not be methylated in the presence of HeLa nuclear proteins by the CpG DNA methyltransferase, paralleling the results of Dnase I footprinting experiments (data not shown).

In this study, protein-DNA interactions in the control regions of polymerase III-transcribed genes have been investigated. It is likely that the methylation protection assay will be applicable to polymerase I- or II-transcribed genes as well. The method described in this report might be generally applicable, in particular to the analysis of 5'-CG-3'-rich promoters and will complement conventional Dnase I footprinting techniques. Obviously, for promoter sequences lacking 5'-CG-3' dinucleotides the method cannot be used.

These results obtained with several DNA fragments document that the CpG-specific DNA methyltransferase from Spiroplasma species can be used as a sensitive tool to probe specific DNA-protein interactions. DNA sequences not complexed in vitro with nuclear proteins from HeLa cells are completely methylated by this enzyme. The regulatory motifs B both in different repetitive Alu elements and in the VAI gene of Ad2 DNA, are protected by bound HeLa nuclear proteins against the action of DNA methyltransferase. The CpG DNA methyltransferase likely senses changes in the major groove of the DNA double helix. In the RNA polymerase III transcribed Alu or VAI segments tested, the binding of the transcription factor TFIIIC at the regulatory region B (13, 23, 25) interferes with the DNA methyltransferase. Thus these sequences remain unmethylated (Figs. 1, 2).

Dimethylsulfate, on the other hand, reacts with G residues in the major groove of the DNA. Because it is a small molecule, only short range DNA-protein contacts can be assessed by this method. In vitro DMS footprinting of different Alu elements and of the VAI gene yields an enhancement of the reaction limited to the center of the B region (Kochanek et al., in preparation). In contrast, the enzymatic methylation reaction by CpG DNA methyltransferase, which also scans the major groove, can be inhibited by proteins binding over a longer stretch of DNA.

ACKNOWLEDGEMENTS

We thank C.-K.J. Shen, Department of Genetics, University of California, Davis, CA; S.J. Friezner Degen, Children's Hospital Medical Center, University of Cincinnati, Cincinnati, OH; B. Vallee and R. Shapiro, Harvard Medical School, Boston, MA for cloned DNA fragments containing the different Alu elements, and Petra Böhnn for expert editorial work. This research was supported by grants from the Deutsche Forschungsgemeinschaft through SFB274-TP1 and from the Fritz-Thyssen-Stiftung.

REFERENCES

1. Galas, D., and Schmitz, A. (1978) Nucleic Acids Res., 5, 3157–3170.
2. Drew, H.R. (1984) J. Mol. Biol., 176, 535–557.
3. Suck, D., and Oehler, C. (1986) Nature, 321, 620–625.
4. Drew, H.R., and Travers, A.A. (1984) Cell, 37, 491–502.
5. Lehmi, A., and Suck, D. (1991) J. Mol. Biol., 221, 645–667.
6. Siebenlist, U., and Gilbert, W. (1980) Proc. Natl. Acad. Sci. USA, 77, 122–126.
7. Renbaum, P., Abrahamov, D., Fainsod, A., Wilson, G.G., Rottem, S., and Razzin, A. (1990) Nucleic Acids Res., 18, 1145–1152.
8. Jittermann, R., Hosokawa, K., Kochanek, S., and Doerfler, W. (1991) J. Virol., 65, 1735–1752.
9. Friezner Degen, S.J., Rajput, B., and Reich, E. (1986) J. Biol. Chem., 261, 6972–6985.
10. Kurachi, K., Davie, E.W., Strydom, D.J., Riordan, J.F., and Vallee, B.L. (1985) Biochemistry, 24, 5494–5499.
11. Kochanek, S., Renz, D., and Doerfler, W. (1992) EMBO J., 12, 1141–1151.
12. Perez-Stable, C., Ayres, T.M., and Shen, C.-K.J. (1984) Proc. Natl. Acad. Sci. USA, 81, 5291–5295.
13. Perez-Stable, C., and Shen, C.-K.J. (1986) Mol. Cell. Biol., 6, 2041–2052.
14. Richardson, C.C. (1965) Proc. Natl. Acad. Sci. USA, 54, 158–165.
15. Kiesow, H., Overgaard-Hansen, K., and Petkar, S.A. (1971) Eur. J. Biochem. 22, 371–381.
16. Lehman, I.R., Bessman, M.J., Simms, E.S., and Kornberg, A. (1958) J. Biol. Chem., 233, 163–170.
17. Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983) Nucleic Acids Res., 11, 1475–1489.
18. Maxam, A.M., and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA, 74, 560–564.
19. Ohnori, H., Tomizawa, J., and Maxam, A.M. (1978) Nucleic Acids Res., 5, 1479–1485.
20. Toth, M., Muller, U., and Doerfler, W. (1990) J. Mol. Biol., 214, 673–683.
21. Roberts, R.J., Akaosiljov, G., Aleströen, P., Gelin, R.E., Gingeras, T.R., Scaky, D., and Petterson, U. (1986) Devel. Mol. Virol., 8, 1–51.
22. Lassar, A.B., Martin, P.L., and Roeder, R.G. (1983) Science, 222, 740–748.
23. Geiduschek, E.P., and Tocchini-Valentini, G.P. (1988) Annu. Rev. Biochem., 57, 873–914.
24. Yoshinaga, S.K., L’Etoile, N.D., and Berks, A.J. (1989) J. Biol. Chem., 264, 10726–10731.
25. Van Dyke, M.W., and Roeder, R.G. (1987) Mol. Cell. Biol., 7, 1021–1031.