Volume 9 Number 3 1981 Nucleic Acids Research

# In vitro methylation of DNA with Hpa II methylase

Adinah Quint and Howard Cedar

Department of Molecular Biology, Hebrew University - Hadassah Medical School, Jerusalem, Israel

Received 10 October 1980

ABSTRACT

The enzyme Hpa II methylase extracted and partially purified from Haemophilus parainfluenza catalyzes the methylation of the tetranucleotide sequence CCGG at the internal cytosine. The enzyme will methylate this sequence if both DNA strands are unmethylated or if only one strand is unmethylated. Conditions have been developed for producing fully methylated DNA from various sources. In vitro methylation of this site protects the DNA against digestion by the restriction enzyme Hpa II as well as the enzyme Sma I which recognizes the hexanucleotide sequence CCCGGG. These properties make this enzyme a valuable tool for analyzing methylation in eukaryotic DNA where the sequence CCGG is highly methylated. The activity of this methylase on such DNA indicates the degree of undermethylation of the CCGG sequence. Several examples show that this technique can be used to detect small changes in the methylation state of eukaryotic DNA.

## INTRODUCTION

Although the presence of 5 methylcytosine in eukaryotes was noted many years ago, it is only recently that this modification has been studied in detail. It is clear that methyl moieties are not randomly distributed on the DNA. The dinucleotide sequence CpG is preferentially methylated and constitutes over 90% of all methyl groups (1). As determined by several independent techniques, highly repeated satellite sequences are preferentially methylated and contain over 50% of the 5 methylcytosines in various animal cells (2-4). The location of methyl groups with respect to chromosomal structure is also asymmetrical, with over 75% of these groups concentrated within the nucleosome core (3,5). Recent evidence shows that the methylation pattern in particular genes is tissue specific, suggesting that this modification may play a role in the regulation of gene expression (6-8).

Much of the progress in this field can be attributed to the use of restriction enzymes to assay the presence of methyl groups at particular eukaryotic DNA sequences. The isoschizomers Hpa II and Msp I are particularly use-

ful in this regard. Although both enzymes recognize the same tetranucleotide, CCGG, only Msp I cuts this sequence if the internal C is methylated (9-11). This particular sequence is over 70% methylated in animal cell DNA (11). Although these enzymes are useful for detecting the presence of methyl groups, it is in many instances more useful to measure the sites which remain unmethylated. It has been suggested, in fact, that it is the absence of methylations at certain sites which may play a role in gene expression (6,7,12,13). The enzyme Hpa II methylase is the enzyme responsible for methylating the CCGG sites in the bacteria Haemophilus parainfluenza. In this paper we show that this enzyme can be used for methylating in vitro unmethylated CCGG sites and therefore provides a tool for studying the effect of methylation on DNA transcription and replication and a probe for measuring the degree of methylation in eukaryotic DNA.

## MATERIALS AND METHODS

Preparation and purification of Hpa II methylase. Frozen Haemophilus parainfluenza cells were obtained from New England Biolabs. The methylase was extracted according to the method of Mann and Smith (14) omitting chromatography on DNA agarose columns. Following purification on Sephadex G-50 the active fractions were concentrated against 50% glycerol, 10 mM Tris-HCl pH 7.9, 1 mM EDTA and 5 mM mercaptoethanol and stored at -20°C.

Methylation in vitro. The incorporation of methyl groups into DNA was always measured under standard assay conditions. A typical reaction mixture of 20 µl contained 0.1 - 3 µg DNA, 2 units Hpa II methylase, 50 mM Tris-HCl, pH 7.9, 5 mM DTT and 1 µCi $^{3}$H-S-Adenosyl methionine (SAM) (13.5 Ci/mmole, New England Nuclear Corp.). The reaction mixture was incubated at 37°C for 1 hr and the reaction stopped by the addition of 0.5 ml TNE (10 mM Tris-HCl, pH 7.9; 0.4 M NaCl; 1 mM EDTA) containing 0.2% SDS. Proteins were digested with 50 µg/ml Proteinase K (Merck) for 30 min. NaOH was then added to a concentration of 0.5 N NaOH and the mixture incubated 10 min at 60°C. The remaining labeled DNA was precipitated by TCA, collected on Whatman GF/C filters and counted in a liquid scintillation system (3). One unit of Hpa II methylase is defined as that amount which yields 1 pmole of DNA methylation in 1 hr using trout DNA as template. In vitro methylation of large quantities of DNA was carried out in 50 mM Tris-HCl, pH 7.9, 5 mM DTT, 5µM S-adenosylmethionine for 1 hr at 37°C using a ratio of 1 enzyme unit per µg DNA. The resulting methylated DNA was deproteinized by extraction with phenol and

chloroform-isoamyl alcohol (24:1) and precipitated with ethanol.

Preparation of Nuclei and DNA. DNA was prepared from frozen tissues by grinding in liquid air. The thin powder was suspended in TNE containing 0.2% SDS and the DNA purified as previously described. Chicken erythrocyte nuclei were prepared from whole cells by homogenization in RSB (10 mM Tris, pH 7.9; 10 mM NaCl; 3 mM $MgCl_2$) containing 0.5% NP-40 and successive washing in this same buffer without detergent. For digestions with micrococcal nuclease, nuclei were resuspended in 0.25 M sucrose 10 mM Tris-HCl, pH 7.4; 25 mM KCl; 5 mM $MgCl_2$; 0.5 mM $CaCl_2$ at a DNA concentration of 1 mg/ml. Nuclei were incubated for 10 min at 37°C with 0.5 - 8 µg/ml micrococcal nuclease and the percent digestion was determined by solubility of $A_{260}$ material in 1 M PCA, 1 M NaCl. After digestion the remaining DNA was purified by phenol and chloroform extractions and precipitated with ethanol. Bacterial plasmid pTK (15) which consists of pBR322 containing the Herpes thymidine kinase gene was obtained from Dr. R. Axel.

## RESULTS AND DISCUSSION

The methylation of DNA is usually assayed by following the incorporation of ($^3$H)SAM into DNA. Although quantitative, this method does not provide information about the specificity of methylation. In order to determine the pattern of methylation catalazed by the Hpa II methylase, plasmid pTK DNA was exhaustively methylated and analyzed by restriction enzyme analysis. As shown in figure 1 this enzyme specifically methylates the internal C of the sequence CCGG. Whereas unmethylated DNA could be cut by both Hpa II and Msp I, the methylated DNA was resistant to Hpa II digestion, although the pattern obtained with Msp I was invariant. This methylation was limited to CCGG sites, since other restriction enzymes which are inhibited by either 6mAd or 5mCyt in their recognition site were able to cut Hpa II methylated DNA normally (data not shown). Even the Hpa I site (GTTAAC) which can be modified by the methylation of adenine was not methylated in our in vitro reaction. Specificity can also be determined by quantitative analysis. The plasmid pBR322 has been sequenced and shown to contain 26 sites for the enzyme Msp I (Hpa II) (16). When this DNA is methylated to saturation levels using ($^3$H)SAM, the saturated number of methyl additives corresponds to the number of available Msp I sites.

It has been noted that the enzyme Sma I (CCCGGG), whose recognition site is a subset of the Hpa II sites, is protected by methylation at the in-

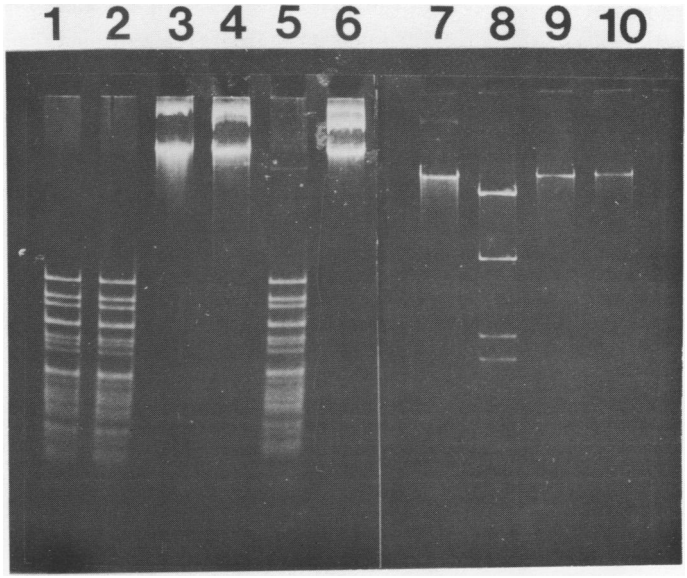


Fig. 1. The specificity of Hpa II methylase. The bacterial plasmid pTK was methylated to completion in vitro with Hpa II methylase and purified using the conditions described in Materials and Methods. These DNA samples were then digested by restriction enzymes and analyzed by agarose gel electrophoresis. Methylated pTK digested with Hpa II (4) or Msp I (5) and undigested (6). For comparison unmethylated pTK was digested with Hpa II (1) or Msp I (2) or left undigested (3). These samples were analyzed on a 2% agarose gel. In order to study digestion by Sma I both unmethylated and methylated plasmid pTK were first digested with Hind III to produce linear plasmid molecules. These samples were run on a 1.5% agarose gel either undigested (lanes 7 and 9) or after digestion with Sma I (lanes 8 and 10). $^{32}$P-RF ØX DNA was included as an internal control in every case.

ternal cytosine (7). It can be seen in figure 1 that modification of the CCGG sites of plasmid pTK also prevents digestion by the enzyme Sma I. When taken together the facts all suggest that this enzyme preparation is specific to CCGG sites and does not contain other methylase activities.

The data of figure 1 show that the sequence CCGG can be qualitatively methylated in vitro. In figure 2 it is demonstrated that the level of methylation at this site is extremely high. Either 1 ng or 10 ng of methylated pTK DNA was digested with the enzyme Hpa II and electrophoresed on a gel containing 2% agarose. This DNA was blotted onto a nitrocellulose membrane filter according to the method of Southern and then hybridized with a high specific activity $^{32}$P nick translated pTK probe (17). This method

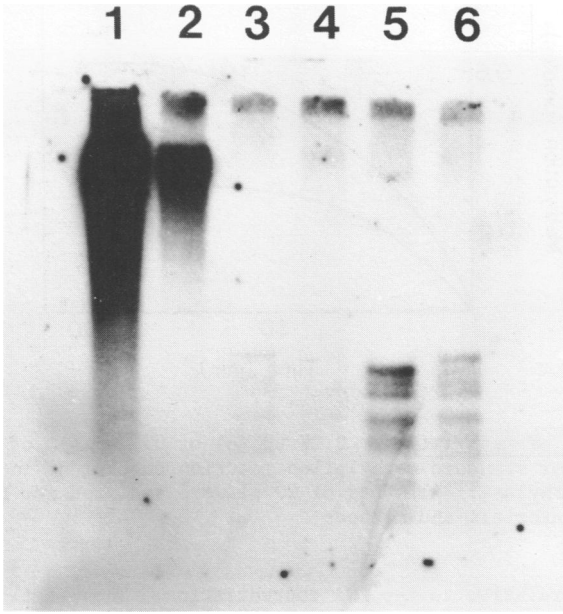


Fig. 2. Quantitation of DNA methylase reaction. 10 ng (lane 1) or 1 ng (lane 2) of methylated pTK (lane 1) together with 30 pg, 10 pg, 200 pg and 100 pg unmethylated pTK (lanes 2-4) were digested with Hpa II and run on a 2% agarose gel. The DNA was transferred to a nitrocellulose filter using the method of Southern blotting (17), hybridized with a nick translated pTK probe ($10^8$ cpm/μg) and autoradiographed using Agfa RP2 X-ray film.

can detect the presence of 10 pg of pTK DNA on the nitrocellulose filter. As shown in figure 2, not more than 30 pg of the methylated pTK was completely digested by Hpa II demonstrating that Hpa II methylase is a highly efficient enzyme for methylating CCGG sequences.

Since this enzyme is potentially useful in probing the methylation pattern of animal DNA it was subjected to a standard enzyme analysis in order to provide information about the kinetics of this reaction. When DNA methylation was followed as a function of time (figure 3) saturation of methylation was obtained at 45-60 minutes independent of the DNA concentration. In these experiments the amount of enzyme was in excess and the number of available methylation sites determines the extent of methylation. No further incorporation was observed after the addition of enzyme at 60 min.

DNA methylation increases in a linear fashion when using varying amounts of DNA substrate (figure 4), and reaches a plateau when the amount of enzyme

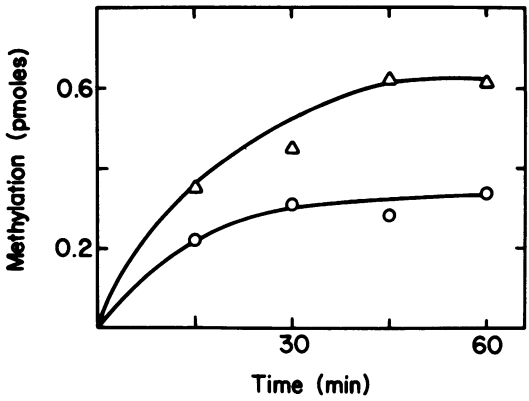


Fig. 3. Kinetics of methylation. 0.25 µg (o) or 0.5 µg (Δ) of chicken DNA was incubated under standard methylation reaction conditions in 100 µl using 2 units Hpa II methylase. Aliquots of 20 µl were analyzed for DNA methylation as described in Materials and Methods.

becomes limiting relative to the DNA concentration. The reaction is also dependent on the amount of enzyme used (figure 5). The data shown in figures 4, 5, and 6 define the conditions required for methylating DNA. Methylation for a period of 1 hr at the proper ratio of enzyme to DNA should guarantee

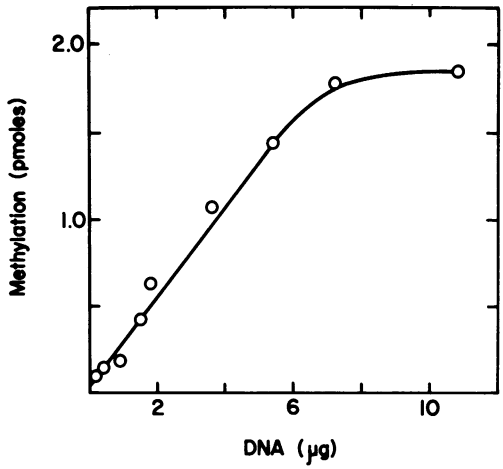


Fig. 4. Methylation as a function of substrate DNA. Various quantities of trout DNA were incubated in a standard reaction mixture of 50 µl for 1 hr at 37°C using 2 units of Hpa II methylase. The extent of methylation was determined as described in Materials and Methods.

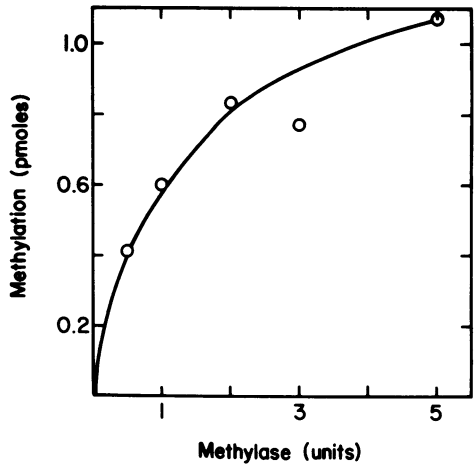


Fig. 5. Methylation as a function of enzyme. 3.6 µg of trout DNA was incubated for 1 hr under standard reaction conditions with various amounts of Hpa II methylase. One unit of enzyme is defined as the amount of enzyme which catalyzes the incorporation of 1 pmole/hr of 5-adenosylmethionine using 10 µg trout DNA under standard assay conditions.

complete methylation of any DNA sample.

Using the kinetic information obtained above, we developed a sensible strategy for determining the number of unmethylated CCGG sites in any DNA

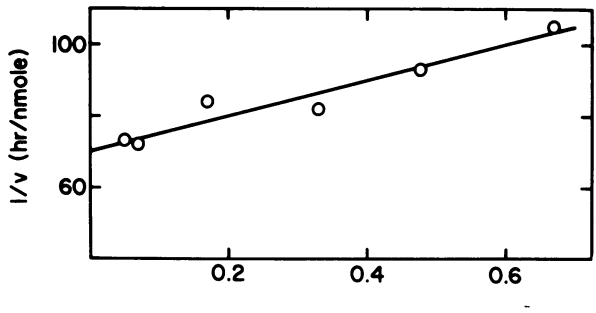


Fig. 6. Determination of the Km for S-adenosylmethionine in the Hpa II methylase reaction. 10 µg of E. coli B DNA was incubated for 1 hr under standard reaction conditions with 2 units of Hpa II methylase and various concentrations of S-adenosylmethionine. Under these conditions the incorporation of S-adenosylmethionine increased linearly with time and was therefore considered the initial rate of reaction. The Km was calculated to be 1.4 µM.

sample. When several concentrations of DNA were methylated to completion in the presence of saturating amounts of enzyme, the degree of methylation was linearly dependent on the amount of DNA and thus on the number of unmethylated CCGG sites in that sample (figures 4 and 7). The slope of the curve obtained by plotting the degree of methylation as a function of DNA is an excellent relative guage of the number of methylatable sites. Figure 7 shows that this assay is accurate when used either on unmethylated DNA or on partially methylated DNA. To this end pTK plasmid was methylated to various degrees in vitro using unlabeled SAM; these partially methylated samples were then used as substrates for the methylation assay. As expected, the extent of methylation was always proportional to the degree of undermethylation of the template DNA.

In order to be useful this type of assay must be sensitive to unmethylated CCGG sites but unaffected by other parameters of the DNA. In table 1 it can be seen that this assay is independent of the size of the DNA over a

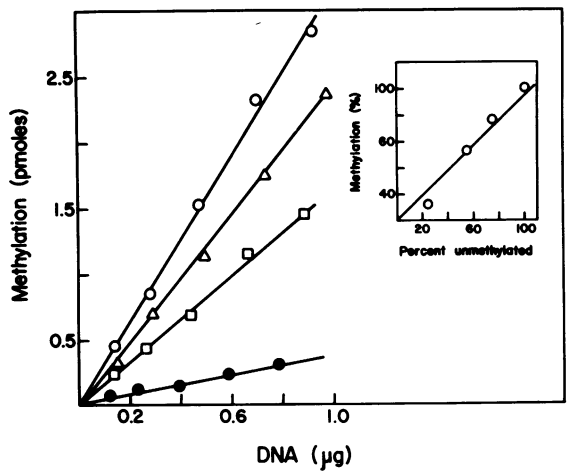


Fig. 7. DNA methylation is dependent on the number of unmethylated sites. 60 µg of pTK DNA were methylated using 40 units of Hpa II methylase for 0 time (o), 15 min (Δ), 30 min ( ) or 45 min (o) in the presence of low specific activity S-adenosylmethionine (0.5 Ci/mmole). The extent of methylation for each time point was determined and these methylated DNA samples were purified by deproteinization and ethanol precipitation. Each sample was then assayed in the standard methylation reaction using various quantities of DNA and 2 units of Hpa II methylase in the presence of high specific activity S-adenosylmethionine (20 Ci/mmole). The degree of methylation for each of these samples is a linear function of the amount of DNA and the slope of this curve serves as an accurate indicator of the methylation potential of each sample. In the insert the methylatability of these DNAs is plotted against the number of available unmethylated sites.

| DNA | Methylation (% of maximum) |
|---|---|
| 8.0 Kb | 100 |
| 4.0 Kb | 100 |
| 2.6 Kb | 100 |
| 0.6 Kb | 100 |
| 0.4 Kb | 70 |
| Hpa II digested | 2 |
| Single stranded | 33 |

Table 1. Effect of DNA on Hpa II methylation. Chicken DNA was digested by different restriction enzymes and the average size determined by reference to markers. Single stranded DNA was prepared by denaturation of DNA at 100°C for 10 min in 10 mM Tris-HCl, pH 7.4. Methylation was determined using the standard assay conditions, with 2 µg DNA and 2 units Hpa II methylase.

12 fold range. A decrease in the methylation efficiency was observed, however, in DNA smaller than 400 bp. When DNA is digested with the restriction enzyme Hpa II, which destroys the CCGG site, only a negligible degree of methylation was obtained, indicating that the reaction is specific only to this site. Single stranded DNA was a relatively poor substrate for methylation (table 1) and the specificity of this reaction has not yet been determined. It should be noted that the methylase activity is also highly dependent on the CCGG content of the DNA template.

The restriction enzyme Hpa II is prevented from digesting DNA at any particular site if either or both strands of the DNA are methylated at the internal cytosine (14). It was of interest to determine whether the *in vitro* reaction yielded full methylation of both cytosines of the CCGG site or hemimethylation of only one cytosine. To this end *in vitro* methylated DNA was digested with Msp I, leaving the internal C of the CCGG site free at the 5' end of each molecule. Digested DNA was labeled with $^{32}P$ at the 5' end of the molecules using polynucleotide kinase and the resulting DNA was digested to mononucleotides. The labeled mononucleotides were analyzed by TLC and the amount of 5mCyt and Cyt was determined (11). As shown in figure 8 all

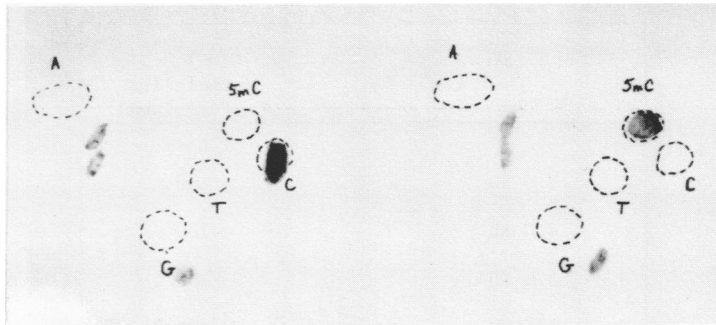


Fig. 8. Thin layer chromatography to determine methylation at Msp I sites. pTK DNA (A) or in vitro methylated pTK DNA (B) were cut with Msp I, treated with alkaline phosphatase and labeled at the 5' end using polynucleotide kinase in the presence of $\gamma^{32}$P ATP. Labeled DNA was digested to mononucleosides and analyzed by two dimensional thin layer chromatography (11).

CCGG sites in the in vitro methylated DNA were fully methylated, since no labeled Cyt was detected. If the DNA had been hemimethylated in vitro we would have obtained both labeled 5mCyt and Cyt in a ratio of 1:1.

The fact that Hpa II methylase produces fully methylated DNA in vitro suggests that this enzyme is capable of methylating hemimethylated sites. In order to test this possibility we prepared hemimethylated DNA by hybridizing unmethylated pTK plasmid with fully methylated plasmid. This template was then used in the methylation assay system described above. As shown in table 2, pTK was a good substrate for the methylation reaction, while fully methylated DNA was a rather poor substrate. After melting and self annealing of these substrates the results were unchanged. These substrates were then annealed to each other using a twenty-fold excess of methylated DNA in order to insure that all unmethylated strands hybridized to methylated partners. This hybrid DNA gave the same level of methylation as unmethylated DNA indicating that hemimethylated Msp I sites are an efficient substrate for this methylase.

Having extensively characterized the methylation reaction one can now use this methylase to ask questions about the methylation pattern of eukaryotic DNA. In particular this assay can be used to assay the number of unmethylated CCGG sites. Using the method described in figures 4 and 7 we assayed the methylatability of DNA from various trout tissues (table 3). The degree of in vitro methylation was relatively constant in most tissues, and was at a level of 1 methylation per 10,000 nucleotides. Analysis of CCGG sites in animal DNA indicates that this sequence appears approximately once

| DNA | Methylation (pmoles) |
|---|---|
| Me pTK | 0.06 |
| pTK | 0.34 |
| Me pTK (reannealed) | 0.06 |
| pTK (reannealed) | 0.38 |
| Me pTK - pTK hybrid | 0.41 |

Table 2. In vitro methylation of hemimethylated DNA. 20μg of fully methylated pTK DNA and 1.0μg unmethylated pTK DNA were linearized by treatment with the restriction enzyme Hind III. These DNAs were denatured by heating and reannealed in 20μl of 10mM Tris-HCl pH 7.8, 0.4M NaCl for 1 hr at 68°C. The resulting DNA was assayed for methylatability under standard reaction conditions. For comparison 20μg of methylated pTK DNA and 1.0μg unmethylated pTK DNA was tested for methylation in the same assay. In this assay background levels of 0.03 pmoles were obtained in the absence of added DNA template. This background has been subtracted from the raw data in order to obtain the numbers shown in the table.

| DNA | Methylation (methyls/10,000 nucleotides) |
|---|---|
| Fish Spleen | 1.0 |
| Liver | 0.9 |
| Sperm | 0.9 |
| Gonads | 0.9 |
| Brain | 1.3 |
| Chicken Erythrocytes | 1.6 |

Table 3. Methylation of mammalian DNA samples. DNA from various sources were assayed for methylation at several concentrations using the technique described in figure 7. The results are expressed as the number of methylations per 10,000 nucleotides of DNA.

every 2200 nucleotides and is 80% methylated (H. Cedar, unpublished results). Thus we would have expected approximately one unmethylated CCGG sequence per 10,000 nucleotides, which is in keeping with the in vitro methylation assay. One tissue, brain, showed increased methylation, indicating that its CCGG sites may be undermethylated relative to other tissues.

The results obtained with other organisms were quantitatively similar but not identical to that of trout DNA. The fact that chicken DNA was methylated to a higher degree than trout DNA could either be due to a lower degree of natural methylation or to a relatively larger number of CCGG sites in this organism. It should be noted, in this regard, that this technique measures the number of unmethylated CCGG sites but gives no information as to what percentage of these sites are indeed methylated in vivo. The technique can, however, be used to compare DNAs from the same organism, since, in this case, the number of CCGG sites in all DNA samples should be the same.

It has been previously shown that the distribution of methyl groups with respect to chromatin structure is not random. When nuclei or chromatin are digested with micrococcal nuclease in order to release mononucleosomes and subnucleosomal particles, the methyl moieties of the DNA remain relatively resistant to digestion (3,5). These results suggest that these groups are preferentially localized in nucleosomal core particals and are therefore more resistant to micrococcal nuclease. We re-examined this possibility using the in vitro methylase assay. Nuclei were digested to various degrees with micrococcal nuclease and the remaining DNA was assayed for the presence of unmethylated CCGG sites using the technique shown in figures 4 and 7. The results shown in figure 9 demonstrate that micrococcal nuclease preferentially attacks unmethylated CCGG sites. After digestion of only 2.5% of the nuclear DNA over 20% of the unmethylated CCGG sites have been removed. That this specificity can be attributed to chromatin structure is demonstrated by the fact that these sites are not preferentially digested by micrococcal nuclease on naked DNA. These observations strengthen the claim that methyl groups are concentrated in nucleosome core particles.

The data in this paper show that the DNA methylase from Haemophilus parainfluenza can be used in studying the state of methylation of the sequence CCGG on various DNA templates. The enzyme is absolutely specific for this tetranucleotide and is capable of methylating this sequence even if one half of the site is already methylated. It should be noted that we have not conclusively proved that every unmethylated internal C of the sequence CCGG can be methylated by this enzyme. It is possible, for instance, that a methylated

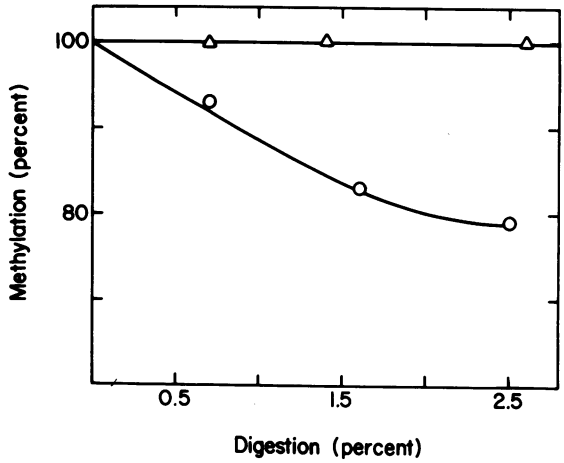


Fig. 9. Sensitivity of chromatin methyl moieties to digestion by micrococcal nuclease. Chicken erythrocyte nuclei (o) were digested with micrococcal nuclease to various extents and the remaining DNA was purified and assayed for methylatable sites as described in the legend to figure 7 using several concentrations of DNA. As a control, deproteinized erythrocyte DNA was also digested with micrococcal nuclease and analyzed as described above. The degree of micrococcal nuclease digestion was determined by measuring the solubilization of $OD_{260}$ absorbing material in PCA. It should be noted that the micrococcal nuclease treatment of nuclei and deproteinized DNA gave similar size distributions as determined by analytical gel electrophoresis.

external C may also influence the reaction. This enzyme is not only useful as an assay for undermethylation, but may provide the technology for analyzing the function of eukaryotic methylation, since in vitro methylated DNA can be used as templates in specific transcription or replication assays. In one series of experiments (15) methylated DNA was inserted into mouse L-cells by gene mediated gene transfer. The methyl modification was thus shown to be an inheritable property of the DNA.

ACKNOWLEDGMENTS

This research was supported by a grant (GM 20483) from the U.S. Public Health Service.

REFERENCES

1. Grippo, P., Iaccarino, M., Parisi, E. and Scarano, E. (1968) J. Mol. Biol. 36, 195-208.

2. Harbers, K., Harbers, B. and Spencer, J.H. (1975) Biochem. Biophys. Res. Commun. 66, 738-746.
3. Solage, A. and Cedar, H. (1978) Biochemistry 17, 2934-2938.
4. Miller, O.J., Schnedl, W., Allen, J. and Erlanger, B.F. (1974) Nature 251, 636-637.
5. Razin, A. and Cedar, H. (1977) Proc. Natl. Acad. Sci. USA 74, 2725-2728.
6. Mandel, J.L. and Chambon, P. (1979) Nucleic Acids Res. 7, 2081-2103.
7. Van der Ploeg, L.H.T. and Flavell, R.A. (1980) Cell 19, 947-958.
8. McGhee, J.O. and Ginder, G.D. (1979) Nature 280, 419-420.
9. Singer, J., Roberts-Ems, J. and Riggs, A.D. (1978) Science.
10. Wallowijk, C. and Flavell, R.A. (1978) Nucleic Acids Res. 5, 4631-4641.
11. Cedar, H., Solage, A., Glaser, G. and Razin, A. (1979) Nucleic Acids Res. 6, 2125-2132.
12. Sutter, D. and Doerfler, W. (1980) Proc. Natl. Acad. Sci. USA 77, 253-256.
13. Desrosiers, R.C., Mulder, C. and Fleckenstein, B. (1979) Proc. Natl. Acad. Sci. USA 76, 3839-3843.
14. Mann, M.B. and Smith, H.O. (1977) Nucleic Acids Res. 4, 4211-4221.
15. Pollack, Y., Stein, R., Razin, A. and Cedar, H. (1980) Proc. Natl. Acad. Sci. USA, in press.
16. Sutcliffe, J.G. (1978) Nucleic Acids Res. 5, 2721-2728.
17. Weinstock. R., Sweet, R., Weiss, M., Cedar, H. and Axel, R. (1978) Proc. Natl. Acad. Sci. USA 75, 1299-1303.