Volume 9 Number 18 1981 Nucleic Acids Research

Distribution of 5-methylcytosine in the DNA of somatic and germline cells from bovine tissues

Karin S.Sturm and J.Herbert Taylor

Institute of Molecular Biophysics, and Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA

Received 23 June 1981

ABSTRACT

Genomic DNA of calf thymus contains 1.5 times as much 5-methylcytosine as similar sperm DNA, but the major EcoRI repeat fragment from satellite I of thymus contains ten times as much 5-methylcytosine as the corresponding fragment from sperm DNA. Restriction enzyme analyses of the total DNA and the satellite I fragment show that three HpaII sites in the fragment are completely unmethylated in sperm but fully methylated in thymus DNA. Undermethylation of many sites in the satellite DNAs can probably account for the lower level of methylation of sperm DNA rather than hemimethylation as previously suggested. These results are also discussed in relation to maintenance and de novo (initiation-type) methylases.

INTRODUCTION

The availability of isoschizomers of restriction enzymes, which distinguish recognition sites on the basis of methylation of cytosine or adenine, have revived interest in studying the pattern of methylation of DNA in relation to its functional significance (1). Since most reports to date have shown that in eucaryotes DNA methylation is restricted mainly to cytosines in CpG doublets (2,3), the most useful enzymes for analysis of this DNA are the pair, HpaII and MspI. Both enzymes recognize the sequence 5'-CCGG-3', but distinguish between the methylation of the cytosine residues (4). MspI cleaves the DNA at the site even if one or both of the inner cytosines are methylated, but does not cut if one or both 5' cytosines are methylated (5). HpaII, on the other hand, cuts the sequences -CCGG- and -^{m}CCGG-, but not -C^{m}CGG- (6). By using this pair of restriction enzymes and electrophoresis followed by Southern blotting and the use of probes for specific sequences, it has been possible to analyze DNAs for the methylation of specific CpG doublets. Results from these experiments suggest a correlation between methylation of these CpG sites and gene function in a particular cell type (7-9). The hypothesis being tested is that DNA methylation is correlated

with reduced transcription, but the evidence so far is insufficient to draw definite conclusions. Part of the problem is that only a limited number of all CpG sites present can be tested for methylation by restriction enzyme analysis.

For sequencing studies the cellular DNA is usually amplified by cloning in a bacterial plasmid, which results in the loss of the original methylation pattern. This leaves the repeated sequences as the only available material for direct analysis by sequencing.

Although the hypothesis concerning regulation of transcriptional activity may not be testable with most repeated sequences, the pattern of methylation in different tissues, the initiation and maintenance of these patterns, and the mutability of the 5-methylcytosine (m^5Cyt) containing base pair can be studied.

With these ideas in mind we have examined the DNA of bovine thymus and sperm to test a hypothesis that the reduced methylation of some mammalian sperm DNAs might be explained by strand asymmetric hemimethylation (10).

In conformity with a previous report (10) we found that thymus DNA contains about 1.5 times as much m^5Cyt as sperm DNA. However, at least for the MspI/HpaII sites tested, this difference is not due to hemimethylation.

Bovine satellite I, density 1.714 g/cm^3 (12), which is cut by EcoRI into segments of about 1400 base pairs (13-16), is very much undermethylated in sperm DNA compared to thymus DNA. The difference in methylation of the fragments from both tissues is large enough to suggest that reduced methylation of all highly repeated sequences could account for all or most of the difference in the level of methylation between total sperm and somatic cell DNA.

MATERIALS AND METHODS

Isolation of DNAs

Calf thymus DNA was isolated according to published procedures (17) with some variations. The frozen tissue was macerated in 250 mM sucrose, 50 mM Tris-HCl (pH 7.5), 25 mM KCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$ (buffer A) in a Waring Blender. Nuclei were prepared by Teflon-glass homogenization in buffer A containing 0.5% Triton X-100, and lysed in 100 mM $NaHCO_3$ (pH 10.8), 50 mM EDTA, 0.02% NaN_3, 1% Sarcosyl and 65 µg/ml proteinase I. The lysate was extracted with phenol, treated with 50 µg/ml RNAse A, reextracted and precipitated with isopropanol.

To prepare sperm DNA, the sperm from one bull was pelleted from the semen, washed with 0.15 M NaCl, 0.015 M sodium citrate (SSC), followed by

ethanol and a 1:1 mixture of ethanol and ether to remove lipoproteins from the capsule (18). The washed cells were stirred in 0.25 M 2-mercaptoethanol to reduce disulfide bonds in the keratin-like outer membrane of the sperm head. The cells were lysed by the addition of trypsin to a final concentration of 0.5 mg/ml. DNA was extracted as described above.

Restriction Enzyme Digestion

DNA samples were digested for 6 hours with 1-5 units of enzyme per µg of DNA. A universal restriction buffer (20 mM Tris-HCl (pH 7.4), 5 mM $MgCl_2$, 1 mM dithiothreitol, ± 50 mM NaCl) was used for all reactions. Preparative digests of bull sperm and calf thymus DNA contained 1-5 mg DNA and 100-1000 units of EcoRI (Miles Laboratories). The restriction mixture was incubated for 24 hours at 37°C before separation of the products by gel electrophoresis.

Gel Electrophoresis

The digestion products were separated on 0.7% - 1% agarose gels or on 5% - 8% polyacrylamide gels. After electrophoresis, the gels were stained with ethidium bromide (0.5 µg/ml) in the running buffer (40 mM Tris-acetate (pH 7.5), 5 mM sodium acetate, 1 mM disodium EDTA) and photographed to show the bands. Restriction fragments from preparative digests were recovered from horizontal agarose gels by electrophoresis into hydroxyapatite, and subsequent elution of the fragment with 0.6 M sodium phosphate buffer (pH 7.5) over a Sephadex G-50 Column (19).

Base Analysis

Hydrolysis of multiple, small DNA samples to bases was performed as described earlier (20) with minor variations. The bases were separated on an Aminex-HPC Column (BioRad) at 3,500 psi in 20 mM ammonium carbonate buffer (pH 10.2). Typically 10-15 µg of DNA was hydrolyzed for each determination and the bases could be separated off the Column in 15 minutes. Bases were detected by their absorbance at 273 mm. The relative proportions of the individual bases were calculated from the area under the peaks of the absorbance trace and the molar extinction coefficients of the bases. The extinction coefficients of all bases in this buffer system had been determined separately as a basis for this calculation.

RESULTS

Digestion of Total DNA

Digestion of total calf thymus and sperm DNA with enzymes that did not contain a CpG in their recognition sequence, resulted in an indistinguishable restriction pattern for both DNAs (Fig. 1). When thymus and sperm DNAs were

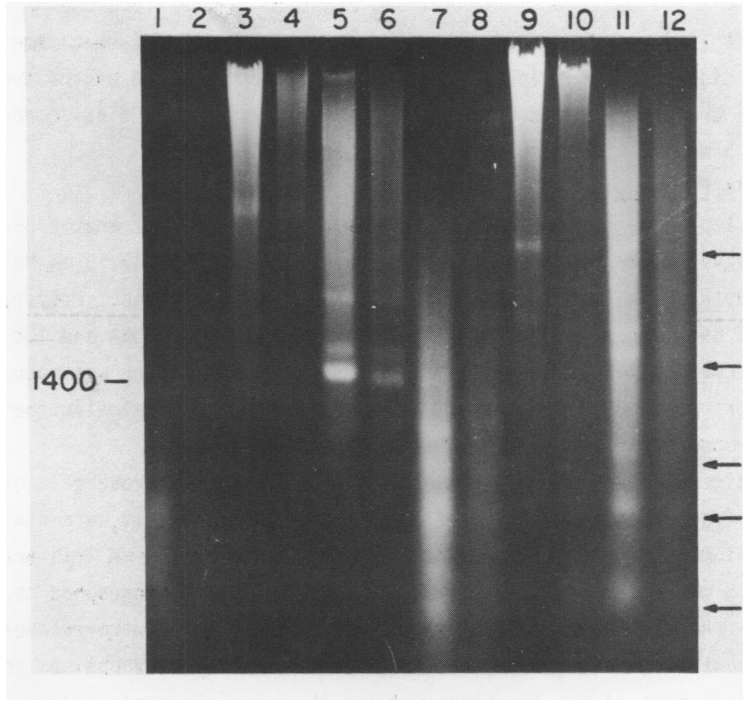

Fig. 1. Digestion pattern of total calf thymus DNA (lanes 1,3,5,7,9,11) and total sperm DNA (lanes 2,4,6,8,10,12) with various enzymes. (1,2) AluI; (3,4) BamHI; (5,6) EcoRI; (7,8) HaeIII; (9,10) HpaII; (11,12) MspI. The arrows indicate the differences between HpaII and MspI digests of both DNAs (see text).

cut with HpaII, the banding pattern after electrophoresis in agarose gels shows, that the highly repeated DNA in thymus is methylated at several sites which are not protected in sperm DNA. After MspI digestion the banding pattern for both DNAs is very similar, and comparable to the HpaII pattern of sperm DNA. This suggests that in the repeated sequences the CCGG sites are unmethylated in sperm and methylated at the inner cytosine in thymus DNA. The background smear in all lanes is produced by digestion of unique and moderately repeated sequences. In contrast to the highly repeated fraction, the unique and moderately repeated sequences show a similar size distribution after restriction of DNA from both tissues, but MspI cuts to a much smaller average size than HpaII. This indicates that most CCGG sequences are methylated to the same extent in both tissues, but minor variations in methylation at specific sites cannot be detected in this way.

Digestion of the Major EcoRI Fragment

In order to obtain more information about the distribution of methylated CCGG sites in the highly repetitive DNA, we isolated the major band from preparative agarose gels after EcoRI digestion of both sperm and thymus DNAs (Fig. 2). The fragment in this band has been shown to originate from satellite I (14). Typical yields with our isolation procedure were about 30 µg of fragment DNA per mg original DNA. The isolated fragment was then subjected to further restriction analysis.

Figure 3 shows the restriction pattern of the major EcoRI fragment from sperm and thymus DNA in the form of a scan of the polaroid negative of our ethidium bromide stained polyacrylamide gel. When the sperm fragment was cut with MspI, four bands were observed, corresponding to fragment sizes of 620, 330, 250, 185 base pairs, respectively. Peak I is due to contaminating DNA which is of similar size as the major EcoRI fragment and is therefore not separated from the fragment by preparative gel electrophoresis.

HpaII digestion of the sperm fragment yields a pattern similar to the MspI digest. Peak IV, is less distinct, however, and peak II has a shoulder of higher molecular weight indicating that one site is sometimes protected from HpaII cleavage. As already seen in the digest of total DNA, the MspI patterns of thymus and sperm fragments are very similar. However, as in the

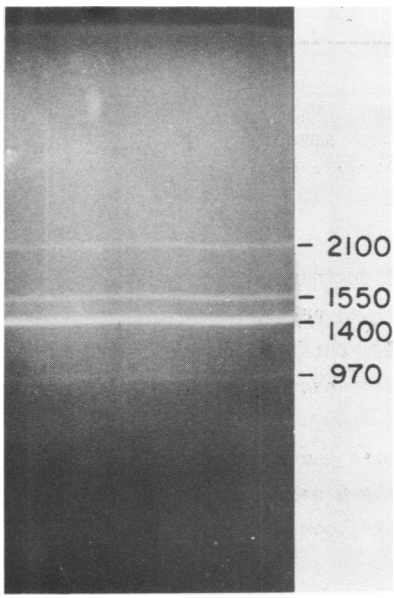

Fig. 2. Preparative gel (1% agarose) of an EcoRI digest of total calf thymus DNA.

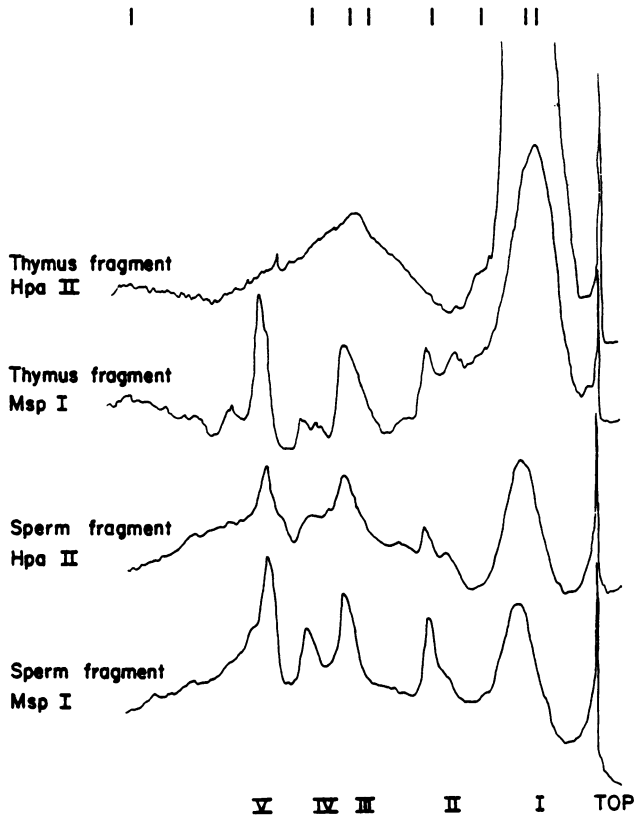

Fig. 3. Densitometer tracing of the negative of a polaroid picture of an 8% polyacrylamide gel stained with ethidium bromide. The molecular weight markers are an SV40/Sau3A digest separated on the same gel. The sizes of the marker fragments are 1330, 1264, 945, 610, 396, 384, 237 and 60 base pairs, respectively.

HpaII digest of the sperm DNA, peak IV is smaller than in the MspI digest and a shoulder on peak II is observed. In addition, a peak of slightly lower molecular weight than peak V is observed. HpaII cuts the thymus fragment only partially and the broad peak around 300 base pairs suggests cutting at more than one of the sites.

Since ethidium bromide fluorescence is not a quantitative measure of DNA concentration, these data cannot be used to determine relative efficiences of cutting at individual sites. In all experiments, control digests of SV40 DNA indicated complete digestion.

Table 1. Base composition of bovine DNAs as determined by HPLC chromatography after formic acid hydrolysis of the DNA to bases.

(means with standard deviations)

Origin of DNA	Number of Determinations	Base Composition (mole % of total bases) A	T	G	C	5-meC	$\frac{5\text{-meC}}{C + 5\text{-meC}} \times 100$	$\frac{A + T}{G + C}$
Calf thymus (total DNA)	9	29.04 ± 2.26	28.31 ± 1.18	20.76 ± 1.08	20.43 ± 1.00	1.46 ± 0.44	6.67	1.35
Bull sperm (total DNA)	9	28.80 ± 0.68	29.09 ± 1.26	20.72 ± 0.49	20.43 ± 1.02	0.95 ± 0.26	4.44	1.38
Calf thymus Satellite I (Eco RI fragment)	4	24.02 ± 3.83	25.47 ± 0.29	22.34 ± 2.25	25.33 ± 4.32	2.84 ± 0.66	10.08	0.98
Bull sperm Satellite I (Eco RI fragment)	4	26.08 ± 0.92	24.85 ± 1.16	22.20 ± 1.08	26.63 ± 1.63	0.25 ± 0.20	0.93	1.04

Analysis of Base Composition

Table 1 summarizes the data from all base analyses. Total calf thymus DNA contains about 1.5 times as much m^5Cyt as total sperm DNA. The EcoRI fragment from thymus contains roughly 10 times as much m^5Cyt as the corresponding sperm fragment. Since the m^5Cyt content is expressed as the fraction of cytosine residues that are methylated, the higher G + C content of satellite DNA is taken into account.

Assuming that the satellite I comprises about 5% of the total genomic DNA, the high level of m^5Cyt can account for about 25% of the over methylation of thymus compared to sperm. Possibly all the extra methylation could be due to satellite methylation if the same results should be found in the other satellites which comprise about 23% of total DNA in bovine cells.

DISCUSSION

We have confirmed a previous report that bovine sperm DNA contains much less 5-methylcytosine than DNA from thymus cells (10). Vanyushin et al. (11) have also reported much less m^5Cyt in bovine sperm than in liver, spleen, lung or kidney, but they did not examine thymus DNA. Our results obtained from restriction enzyme analysis of the major EcoRI repeat of satellite I, have revealed three HpaII sites, which are methylated at the inner cytosine residue in thymus DNA, whereas these same sites are not modified in sperm DNA. At least in this satellite fraction, bovine sperm DNA is not hemimethylated as proposed by Sneider (10). Undermethylation of the various satellite DNAs in sperm compared to somatic cells could largely account for the difference in total m^5Cyt content between the two tissues. Sequence studies now in progress in our laboratory should reveal the distribution pattern of the many other methylcytosines in this EcoRI repeat, which cannot be detected by HpaII and MspI digestion. Base composition and restriction analysis of the other satellites will be used to test the above hypothesis. Several recent studies of developmentally regulated genes have indicated that there are CG doublets in eucaryotic DNA which are methylated in the sperm even when the corresponding sequences in certain somatic cells are not modified. The methylation of the bovine satellite is a clear case of the opposite condition.

Since somatic cells inherit their DNA from sperm and egg, our results indicate the action of an initiation type methylase, which modifies unmethylated DNA de novo during development. This initiation type methylase (21) has to act in addition to the maintenance type, which is the most abundant methylase in higher cells and uses preferentially half-methylated DNA. Pre-

sumably the maintenance enzyme scans the newly replicated helix behind the fork and methylates CG sites only if a modified cytosine is present on the other chain in a symmetrical position. There is now increasing evidence for this type of mechanism, since DNA introduced into cells by transformation retains most of the introduced methylated sites (1,22). The only case known, where a mammalian methylase appears to use completely unmethylated DNA as substrate in vivo, is during certain viral infections. The acquisition of a new methylation pattern seems to be correlated with integration of the viral DNA into the host genome, but the mechanism is not known (23).

Because only repeated sequences can be isolated with an intact methylation pattern in sufficient quantities for sequence determination, further studies on these satellite sequences may be justified even though the satellites are not known to be transcribed. Hopefully we will learn not only more about the stability of methylation patterns, but also about the mechanisms of their origin, i.e. the operation of initiation type methylases in higher cells.

ACKNOWLEDGMENTS

This research was supported by grants from the U.S. Department of Energy (EY-78-S-05-5854) and from the National Institute on Aging (1 R01 AG01807-01).

REFERENCES

1 Razin, A. and Riggs, A.D. (1980) Science 210, 604-610.
2 Doskocil, J. and Sorm, F. (1962) Biochim. Biophys. Acta 55, 953-959.
3 Bird, A.P. and Taggart, M.H. (1980) Nucleic Acids Res. 8, 1485-1497.
4 Waalwijk, C. and Flavell, R.A. (1978) Nucleic Acids Res. 5, 4631-4641.
5 Sneider, T.W. (1980) Nucleic Acids Res. 8, 3829-3840.
6 Mann, M. and Smith, H.O. (1978) Nucleic Acids Res. 4, 4211-4221.
7 Mandel, J.L. and Chambon, P. (1979) Nucleic Acids Res. 7, 2081-2103.
8 McGhee, J.D. and Ginder, G.D. (1979) Nature 280, 419-420.
9 van der Ploeg, L.H.T. and Flavell, R.A. (1980) Cell 19, 947-958.
10 Kaput, J. and Sneider, T.W. (1979) Nucleic Acids Res. 7, 2303-2322.
11 Vanyushin, B.F., Trachera, S.G. and Belozersky, A.N. (1970) Nature 225, 948-949.
12 Macaya, G., Cortadas, J. and Bernardi, G. (1978) Eur. J. Biochem. 84, 179-188.
13 Roizes, G. (1976) Nucleic Acids Res. 3, 2677-2696.
14 Gautier, F., Mayer, H. and Goebel, W. (1976) Mol. Gen. Genet. 149, 23-31.
15 Kopecka, H., Macaya, G., Cortadas, J., Thiéry, J.P. and Bernardi, G. (1978) Eur. J. Biochem. 84, 189-195.
16 Philippsen, P., Streeck, R.E., Zachau, G. and Müller, W. (1975) Eur. J. Biochem. 57, 55-68.
17 Rill, R.L., Hilliard, P.R., Bailey, J.T. and Levy, G.C. (1980) J. Am. Chem. Soc. 102, 418-420.
18 Borenfreund, E., Fitt, E. and Bendich, A. (1961) Nature 101, 1375-1376.
19 Tabak, J.F. and Flavell, R.A. (1978) Nucleic Acids Res. 5, 2321-2332.

20 Pollock, J.M., Swihart, M. and Taylor, J.H. (1978) Nucleic Acids Res. 5, 4855-4863.
21 Taylor, J.H. (1979) in Molecular Genetics III, Chromosome Structure, Taylor, J.H., Ed., pp. 89-115, Academic Press, New York.
22 Wigler, M.H., Levy, D. and Perucho, M. (1981) Cell 24, 33-40.
23 Sutter, D. and Doerfler, W. (1980) Proc. Natl. Acad. Sci. USA 77, 253-256.