Volume 16 Number 19 1988 Nucleic Acids Research

The left end of rat L1 (L1Rn, long interspersed repeated) DNA which is a CpG island can function as a promoter

Israel Nur+, Esterina Pascale and Anthony V.Furano*

Section on Genomic Structure and Function, Laboratory of Biochemical Pharmacology, National Institute of Diabetes, and Digestive and Kidney Diseases, NIH, Building 8, Room 203, Bethesda, MD 20892, USA

Received April 25, 1988; Revised and Accepted September 6, 1988

ABSTRACT

Here we report that the 600 bp promoter-like region at the left end of a newly isolated and characterized rat L1 DNA element can activate the prokaryotic chloramphenicol acyltransferase gene in a rat cell line. Activation only occurs when the promoter region is oriented to the transferase gene as it is to the L1 protein encoding sequences and is 75% inhibited by methylation of just 5 of the 22 CpGs present in the promoter. The G + C rich promoter contains enough CpGs to qualify it as a CpG island, but in contrast to other CpG islands, genomic L1 promoters are fully methylated in both somatic cell and sperm DNA as judged by restriction enzyme analysis. Partial demethylation of the genomic promoters by treatment with 5-azacytidine failed to produce discrete L1 transcripts. The relationship of methylation to the evolutionary history and fate of the rat L1 promoter is discussed.

INTRODUCTION

L1 DNA elements (long interspersed repeated DNA, LINE family) are a ubiquitous feature of mammalian genomes (1-3) and recent evidence suggests that they most likely are transposable elements (4-10). Although the mechanism of transposition is unknown, the fact that mammalian L1 elements share highly conserved protein encoding sequences suggests that L1 DNA encodes at least some of the functions necessary for this process (11-14). In addition, each of the mammalian L1 families so far examined contains a promoter-like sequence at the left end of the element that could regulate transcription of the protein encoding sequences (11-14) (see upper diagram in Fig. 1).

Understanding the regulation of L1 DNA transcription would greatly facilitate the activation or isolation of a transpositionally competent L1 element. A necessary first step for such studies is to establish that the promoter-like region can in fact function as a promoter. To do this we recently isolated and characterized two additional rat L1 elements that contain an apparently full length promoter region (14). This 600 bp region is > 50% G + C and contains several different transcriptional factor binding

sites and enough CpGs to quality it as a CpG island (14-17). No other region of rat L1 DNA has the characteristics of a CpG island (11, 14).

Here we show that the promoter-region of one of these elements can activate the transcription of the Escherichia coli chloramphenicol acyltransferase (CAT) gene in a rat tissue culture line. In vitro methylation of the promoter HpaII sites, which account for 5 of the 22 CpGs in the promoter region, partially inhibited promoter activity and the degree of inhibition was in part dependent on the sequence context. In contrast to other genomic and cloned CpG islands (17, 18), the rat L1 promoter HpaII sites are fully methylated in somatic cell and sperm DNA. These results are discussed in light of the fact that during evolution each mammalian L1 family has acquired a promoter region that has no sequence similarity to other L1 promoters but is nonetheless a CpG island.

MATERIALS AND METHODS

CAT Plasmids

The pS series of plasmids were constructed by standard techniques in our laboratory and the relevant parts of their structure are shown in Fig. 1. These plasmids are derivatives of pBR327 which is a deleted version of pBR322 (19) and contain the HindIII-AatII fragment of pSV2 which includes the prokaryotic CAT gene and all of the SV40 sequences that are 3' of it in pSV2 (20). Therefore, the CAT coding sequences and the 3' untranslated sequences are identical in the pS and pSV series of plasmids.

Cell Lines

The nontransformed rat fibroblast cell line, R2 (21), was obtained from Dr. K. Van Doran at Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. The nontransformed rat CREF cell line (22) and the MoMLV transformed CREF cell line (P113) were obtained from Dr. Robert H. Bassin, National Cancer Institute, National Institutes of Health. Cell lines were maintained in Dulbecco Modified Eagles medium containing 10% fetal calf serum, 2 mM glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 0.25 µg/ml fungizone. Only the P113 cell line was efficiently transfected by the $CaPO_4$ method (see below) and was used for all of the transfection experiments.

Cell Transfection and CAT Enzyme Assays

A different batch of freshly thawed P113 cells was used for each experiment and were passaged no more than once before transfection. The cells were plated at about 1 x 10^6 cells/100mm plate the day before an experiment and transfected when they were about 30% confluent. The media was changed 4 hours

before transfection using calcium phosphate-DNA precipitates prepared by a modification (23) of the procedure of Graham and van der Eb (24), except that the DNA and $CaCl_2$ was mixed thoroughly before the addition of the phosphate buffer. Each transfection was carried out in duplicate using independently prepared calcium phosphate-DNA precipitates and in the presence of 0.1 mM chloroquine (25). After 4 hours the medium was replaced, and after 48 hours the cell monolayers were washed twice with phosphate-buffered saline (PBS, 10 mM $NaPO_4$, pH 7.4, 0.15 M NaCl), scraped from the plate, washed once in 1 ml of 0.15 M NaCl, 10 mm Tris-HCl, pH 8.0, 5 mm sodium EDTA, and sonicated 5 seconds in 200 µl of cell lysis buffer (62.5 mM Tris-HCl, pH 7.6, 15% glycerol, and 5 mm DTT).

CAT activity was measured on the clarified lysate by one of three assays (20, 26, 27). All gave comparable results on a given set of extracts, but we found the assay described in Ref. 27 generally more convenient and used it routinely. For this assay, 20-50 µl of extract is incubated at 37°C for 1 hour, in a mixture (final volume, 200 µl) containing 0.156 M Tris-HCl, pH 7.6, 0.45 mg/ml of butyryl CoA, and 0.025 mM [^{14}C]chloramphenicol (0.23 µCi, about 400,000 cpm). The reaction mixture is extracted with 400 µl of Pristane:xylene (2:1, vol/vol), and the radioactivity of the organic phase (butyryl-chloramphenicol) is determined. Experiments with purified CAT enzyme showed that the reaction is linear up to the conversion of about 180,000 cpm (about 50%) of the [^{14}C]chloramphenicol to the butyrylated form. The no enzyme (or no extract) blank was 700-900 cpm, and was subtracted from all of the values shown. Assays with extracts prepared from nontransfected cells or cells transfected with salmon sperm DNA gave the same results as the no enzyme blank.

Methylation of DNA In Vitro

Plasmid DNAs were incubated for 2 hours at 37°C with the prokaryotic HpaII methylase (New England Biolab) using the conditions suggested by the supplier in either the presence or absence of 0.08 mM S-adenosylmethionine. We refer to DNA subjected to the later reaction as mock-methylated DNA. The reactions were then heated at 65°C for 20 minutes to inactivate the methylase and then adjusted to 0.3 M sodium acetate. The DNA was precipitated with 2 volumes of ethanol, washed with 70% ethanol, dried, and dissolved in 10 mm Tris-HCl, pH 8.0, and 1 mm sodium EDTA.

Each experiment shown in Table 1 was carried out with a different preparation of methylated DNA, and in each case the electrophoretic profile of the methylated and mock-methylated plasmids were indistinguishable from the

untreated plasmid; i.e., > 90% of the DNA was negatively supercoiled form I (results not shown). In addition, each preparation was tested for sensitivity to digestion by HpaII restriction endonuclease; in each case the methylated plasmid was completely resistant to digestion, whereas the mock-methylated or untreated plasmids were completely sensitive to the enzyme. Finally, we could not detect any difference between the production of CAT enzyme by the mock-methylated plasmids and their untreated counterparts.

R2 cells were used for the demethylation experiments and were treated with 5-azacytidine (5-azaC) as follows: Cells at approximately 20% confluency were incubated for 18 hours with 4-5 μM 5-azaC. At this time the medium was replaced and the cells allowed to reach confluency. The cells were then subcultured, and when they reached approximately 20% confluency they were treated again with the drug. This procedure was repeated one more time, and at the end of the third treatment, DNA and RNA were extracted from the cells as described below. Little or no demethylation of L DNA HpaII sites was observed after one round of treatment, and maximal demethylation occurred after 3 such cycles.

Preparation of DNA and RNA

DNA was prepared from tissue culture cells and rat sperm (extruded from rat epididymus and judged to be at least 95% pure by microscopic examination) by a method that was provided to one of us (I.N.) by Dr. Howard Cedar and modified by us. After rinsing the cell monolayers twice in PBS, the cells were lysed directly in the tissue culture flasks by the addition of lysis buffer (0.05 M Tris-Cl, pH 8.0, 0.15 M NaCl, 0.05 M sodium EDTA, 1% NaSDS, 100 μg/ml proteinase K). We used 0.05 to 0.1 ml of lysis buffer/square cm of growing surface and the flask was incubated at 55°C for 30 minutes. Pancreatic RNAse to 100 μg/ml was added, and the incubation was continued for an additional 60 minutes. After phenol extraction, the DNA was wound out of the aqueous phase on a glass rod, washed successively in 70% and 95% ethanol, and treated again successively with pancreatic RNAse and proteinase K. After phenol extraction, the DNA was wound out of the aqueous phase, rinsed in 70% and then 95% ethanol, dried briefly, and dissolved in 0.01 M Tris-Cl, pH 7.5, 0.001 M sodium EDTA (TE). Rat liver DNA was prepared as described previously (28). RNA was prepared from tissue culture cells or liver tissues by cesium chloride centrifugation (29) of guanidinium thiocyanate extracts (30).

DNA and RNA Blot Hybridizations

DNA was digested with an excess of MspI or HpaII for 3 hours at 37°C using the conditions suggested by the supplier. The DNA was extracted with phenol,

precipitated with ethanol, and reincubated as above with the respective enzymes. In some cases ØX174 DNA was added as a control to monitor the completeness of the endonuclease digestion. Bidirectional transfer of the DNA to nitrocellulose was carried out as described in Ref. 31, and after the appropriate treatments, the nitrocellulose filters were hybridized with the radioactive nucleic acids indicated in the figures.

RNA was treated with glyoxal in the presence of 50% dimethylsulfoxide at 55°C (32) and resolved on a 1.5% agarose gel by electrophoresis at 120 V in a buffer containing 10 mm Tris-HCl, pH 7.3, 5 mM sodium acetate, and 0.5 mM sodium EDTA. Buffer circulation was maintained throughout the procedure. RNA was electrophoretically transferred to nylon membrane ("Zeta probe") using the "Transblot" cell (Bio-Rad) at 0.8 amp (about 60 V) for 2 hours in buffer of the same composition as that used for electrophoresis using the conditions suggested by the manufacturer. The RNA blots were incubated *in vacuo* at 80°C for 2 hours to release bound glyoxal before hybridization. Some of these blots were reused after first removing the previously hybridrized probe by heating the blot for 30 minutes at 90° in TE. Autoradiography of the reclaimed blots verified complete removal of the probe.

Methyl C Determination

The methyl C content of DNA was determined as follows. DNA was dried, and then hydrolyzed in sealed ampules at 180°C for 2 hours in concentrated (88%) formic acid. After evaporation under N_2, the bases were resolved in 0.2 M sodium phosphate buffer, pH 3.5, on a Partisil 10 SCX (Whatman Co.) HPLC column. The bases were detected as they eluted by their absorbance at 280 nm, and the molar fraction of each base was determined from the area of the elution peak and the molar absorbance of the base at 280 nm. The elution position of each base was determined by comparison to the appropriate standards.

RESULTS

The upper part of Fig. 1 shows a partial restriction map of the left end of L1Rn B6, hereafter referred to as L B6 (14), and the lower part of Fig. 1 shows the relevant portions of the plasmids constructed from L B6 as well as other CAT gene containing clones used here.

Fig. 2 shows that extracts of cells transfected with pSB62 in which the L1 promoter is oriented to the CAT gene as it is to the L1 ORFs contain significant amounts of CAT activity relative to cells transfected with the pSS2 plasmid which contains the SV40 promoter. By contrast, cells transfected

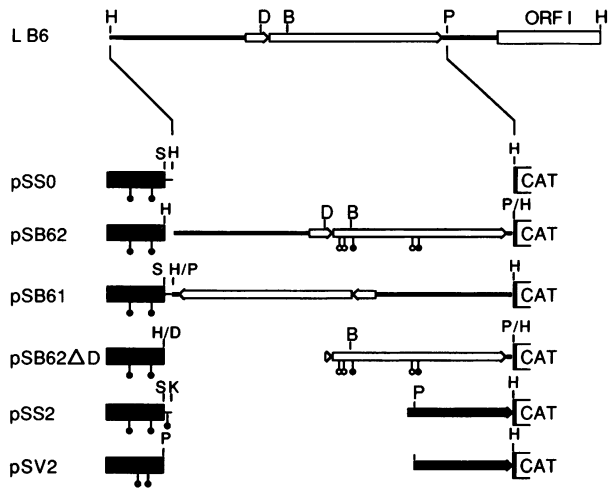

FIG. 1. Promoter fusions.
The upper diagram shows a partial restriction map of the left end and flanking DNA of L B6 (L1Rn B6) (14). The thin black bar indicates non-L1 rat DNA sequence and the large open arrow indicates the L1 promoter. The lower diagram shows the relevant region of the CAT fusion plasmids. The left side depicts vector DNA and for the pS plasmids (stippled box) the region shown is from position 851 to the SalI (S) site at position 651 of pBR322; for the pSV2 plasmid (stripped box) the region shown is from position 1866 to the PvuII (P) site at position 2066 of pBR322. H, D, B, and K represent the sites for HindIII, DraIII, BamHI and KpnI respectively and the thin line represents portions of the multicloning site from pUC19. During the construction of the pSB62 clone the vector SalI (S) site was converted to a HindIII (H) site. Two letters separated by a slash mark indicates a fusion between the indicated non-compatible restriction sites. ◔, HpaII site. The filled symbol represents highly conserved L1 HpaII sites. The filled heavy arrow depicts the SV40 early promoter region from the KpnI site (position 294) to the HindIII site (position 5171) (pSS2) or from the PvuII site (position 270) to the same HindIII site (pSV2).

with the pSS0 plasmid which contains no promoter or the pSB61 plasmid which contains the L1 promoter in the opposite orientation as that in the L B6 element contain no detectable CAT activity. In the experiment shown in Fig. 2, the activity with pSB62 was about 30% of that observed with the SV40 promoter. In other experiments this value ranged from 20 to 35% (results not shown and Table 1).

Table 1 shows that deleting all of the non-L1 rat DNA to the left of the L1 promoter in the pSB62 clone to produce plasmid pSB62ΔD (Fig. 1) has no significant effect on the amount of CAT activity produced in transfected cells (cf. lines 3 and 5). Therefore, L1 DNA is the only rat DNA sequence that is

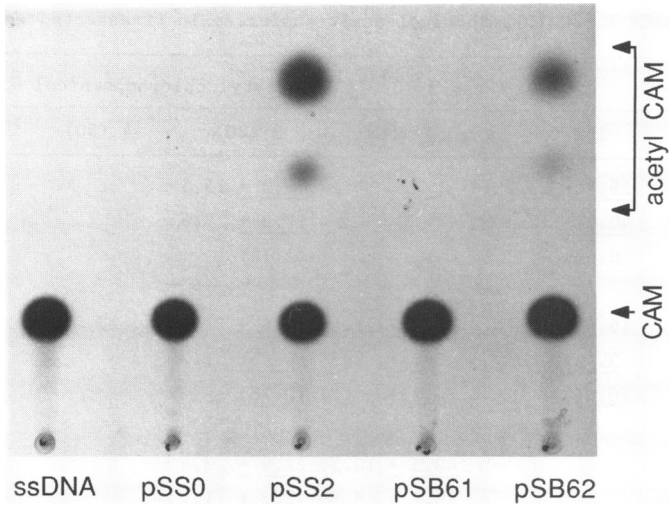

FIG. 2. Chloramphenicol acyltransferase activity in transfected cells. Extracts were prepared from P113 cells transfected with the indicated DNA as described in the Materials and Methods. SSDNA, salmon sperm DNA; CAM, chloramphenicol.

required for the production of CAT enzyme in P113 cells. When the orientation of the B62ΔD sequence is opposite to that shown in Fig. 1, no activation of the CAT gene occurs (data not shown). Therefore, as was implied from the results of Fig. 2, the rat L1 promoter is apparently unidirectional.

Experiments not presented here showed that the L1 promoter is also quite active in non-rat cell lines. For this reason we are carrying out structural analysis of the L1 promoter and the mapping of transcriptional start site(s) in both rat cells and several non-rat cell lines and will present the results of these studies in a future report.

In the remainder of this paper we present experiments that pertain to the fact that the rat L1 promoter is a CpG island (14-17); i.e., the L1Rn promoter is > 50% G + C and the content of CpG is generally that expected from the base composition and not only 0.25 of the expected value as is typical for mammalian DNA in general. Since decreased activity of a number of genomic promoters has been corrected with methylation of the C of promoter CpGs (reviewed in Refs. 17 and 33-36) and since methylation of certain promoters in vitro diminishes their activity when introduced into an appropriate cell line (37-40), we examined the effect of DNA methylation in vitro on the activity of the L1 promoter.

TABLE 1. Chloramphenicol acyltransferase in transfected cells

	cpm x 10^{-3} [^{14}C]butyryl chloramphenicol [a]				
Plasmid	1 (20)	2 (50)	3 (20)	4 (50)	5 (50)
pSS2	59.7 ± 22.2 [b]		83.0 ± 15.5		
pSS2 - Me	40.0 ± 12.3		92.8 ± 9.6		
	(0.67) [c]		(1.12)		
PSB62	16.6	34.2 ± 1.6	16.6 ± 1.3		
pSB62 - Me	3.4 ± 0.4	8.0 ± 2.3	4.3 ± 1.2		
	(0.21)	(0.23)	(0.26)		
pSB62ΔD		40.5 ± 10.5	21.9 ± 3.4		
pSB62ΔD - Me		26.7 ± 8.8	12.2 ± 1.1		
		(0.66)	(0.56)		
pSV2				41.3 ± 20.6	40.1 ± 0.14
pSV2 - Me				59.5 ± 12.7	55.2 ± 12.0
				(1.44)	(1.38)

[a] Cells were transfected with 20 µg of plasmid and either 20 or 50 µl of extract was assayed as given in the parenthesis beside the experiment number. Extracts (50 µl) from cells transfected with up to 40 µg of pSSO (no promoter, see Fig. 1) gave 300-800 cpm above the blank value (see Fig. 2).

[b] Each value is the average and standard deviation of duplicate transfections using independently prepared calcium phosphate precipitates (see Materials and Methods).

[c] This value is the ratio of activity produced by the methylated plasmids to that produced by the mock-methylated plasmids. Only the values double-underlined are statistically significant.

In addition, since the presence of CpG islands in otherwise CpG-deficient genomes can be explained by the fact that, unlike most of the genome, CpG islands are unmethylated [(41) and references therein], we determined the methylation state of genomic L1 promoters.

CAT Enzyme Activity in Cells Transfected with Methylated Plasmids

The L B6 promoter contains 22 CpGs, and of these five are imbedded in the HpaII recognition sequence, CCGG (Fig. 1), and are therefore susceptible to methylation by the prokaryotic HpaII methylase. Table 1 shows that methylation of the pSB62 plasmid reduces CAT enzyme production by at least 75% as

compared to the nonmethylated control (lines 3 and 4). However, methylation of pSB62ΔD reduces CAT enzyme production by only about 40% as compared to the nonmethylated control (lines 5 and 6, Table 1). Since the deleted rat DNA does not contain HpaII sites [Fig 1, (14)], it appears that the effect of methylation of the L1 promoter is somewhat dependent on the sequence of the DNA that flanks the promoter.

Earlier studies by Kruczek and Doerfler (39) showed that production of CAT enzyme by a CAT gene fusion with the SV40 promoter which does not contain HpaII sites was not inhibited by HpaII methylation. They used the PSV2-CAT plasmid (20) in mouse L cells, and our results show that PSV2-CAT is also not inhibited by HpaII methylation in rat P113 cells (Table 1, experiments 4 and 5). Although the results of Table 1 indicate that PSS2 is also apparently not inhibited by methylation (lines 1 and 2), the results of experiment 1 and others not shown suggest that in some cases methylated PSS2 may produce somewhat less CAT enzyme than the mock methylated control. Since the only difference between PSS2-CAT and PSV2-CAT is in the DNA sequence 5' of the SV40 promoter (see Fig. 1), it is possible that methylation of different 5' flanking sequences could differentially affect the SV40 promoter. However, this putative differential effect is both variable and subtle. Therefore, its analysis will require a more sensitive and reliable assay than is possible with our present plasmids.

The Methylation State of Genomic L1 Promoters

By use of the isoschizomers HpaII, which only digests unmethylated HpaII sites (CCGG), and MspI, which digests both methylated and unmethylated sites, it is possible to determine the methylation state of a subset of CpG sites. The upper part of Fig. 3 shows the location of three classes of HpaII sites (a, b, and c) in the L1 promoter and two HpaII sites (d and e) near the right end of L1 DNA. Site e is also in a G + C rich region but this is not a CpG island (11).

The lower part of Fig. 3 shows the blot hybridizations of HpaII (H) or MspI (M) digestion of sperm (S), liver (L), or R2 rat cell (R) DNA with either oligonucleotide A1 or C2 which are specific for the left and right ends of L1 DNA, respectively (see diagram, Fig. 3). MspI digestion produces two sizes of short fragments that hybridize with A1: a major, somewhat diffuse class of about 0.23 kb which results from digestion of L1 HpaII sites at a and b, and a minor class of 0.4 kb. These larger fragments would be produced from L1 members that lack HpaII site(s) b but contain sites at a and c. Less than 20% of the genomic members are like this, since densitometric scans (not present-

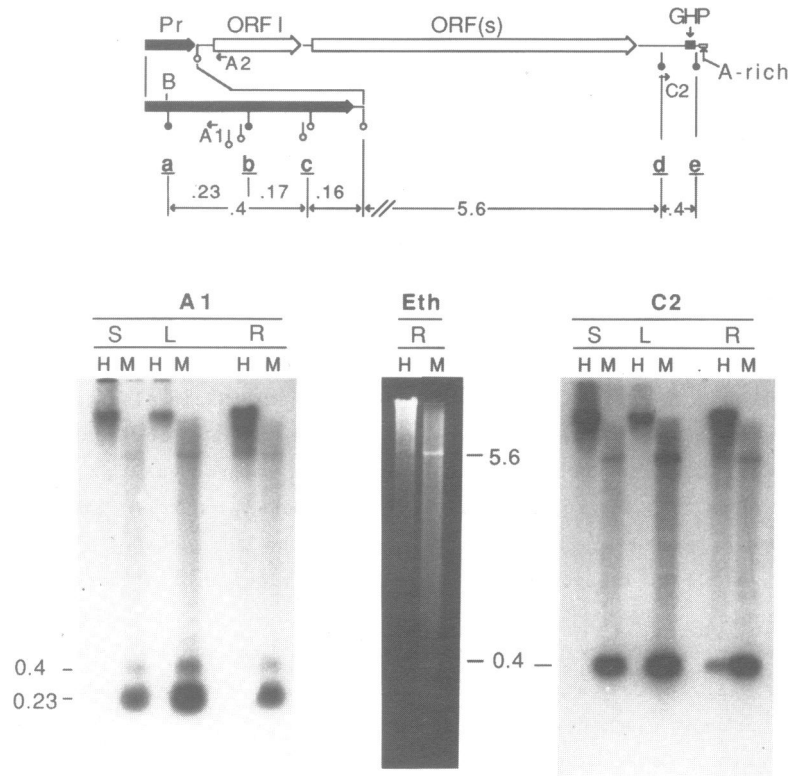

FIG. 3. Methylation state of genomic L1 promoters.
The upper diagram, depicts a generic rat L1 member. Pr, GHP and A-rich designate promoter, G-rich homopurine stretch, and A-rich right end respectively (11,14). B indicates BamHI site, and ● and ○ represent highly conserved and less highly conserved HpaII sites respectively. The numbers indicate size in kb. A1, A2, and C2 are oligonucleotides (30,30 and 40 mers respectively) and the arrow points in the 5' to 3' direction. The lower part of the figure shows the results of blot hybridzation of the indicated oligonucleotide to HpaII (H) or MspI (M) digests of sperm (S), liver (L), and R2 cell line (R) DNA. Eth indicates the ethidum bromide stained gel of the R2 cell line DNA digestion.

ed) showed that the MspI digestions of genomic DNA contained at least five fold more 0.23 kb fragments than 0.4 kb fragments.

The fact that neither the 0.23 nor the 0.4 kb A1 specific fragments were produced by HpaII digestion of genomic DNA means that the HpaII sites at a, b, and c are fully methylated in sperm and somatic cell DNA. This result contrasts with the finding that mammalian CpG islands, when examined in either total genomic DNA or in the case of several different specific CpG islands

(15,18), are unmethylated in both somatic cell and sperm DNA. Although HpaII sites d and e are also fully methylated in sperm and liver DNA, there is some demethylation of these sites in the established rat cell line, R2. The tendency for the DNA of permanent rodent cell lines to become somewhat demethylated has been noted previously (see Ref. 33).

Fig. 3 also shows that the MspI fragments derived from the left end of genomic L1 elements are recovered in about the same yield as those from the right end. This could only occur if most of the rat L1 elements are full length, a point which we have made earlier (11,14). Since these results are from bidirectional blots of the same gel, the relative hybridization by the 2 probes, which were of the same specific activity, is directly comparable.

We next tried to demethylate genomic members by exposing cells to repeated rounds of treatment to 5-azaC (42) as described in Materials and Methods. 5-azaC is incorporated into DNA where it inhibits the cytosine methylase (43). In preliminary experiments we determined that the dose used, 4-5 μM, while somewhat inhibitory to the growth of R2 cells, did not kill them.

Fig. 4 shows the results of three such experiments: In experiments 1 and 2, the cells were exposed to 3 or 4 cycles, respectively, of treatment with 5-azaC which in both cases produced near complete demethylation of HpaII sites d and e, and significant but partial demethylation of the HpaII sites in or near the promoter. The HpaII and MspI digestions contains about the same amount of the 0.4 kb fragment that hybridizes with C2. By contrast, the yield of the 0.23 kb fragments that hybridize with A1 (sites a and b) is significantly less in the HpaII digestion than in the MspI digestion. In addition, the A1 probe detected a series of partial HpaII digestion products including substantial amounts of an approximately 6 kb fragment. Densitometric scans (not presented) of the original autoradiograms from experiments 1 and 2 showed that the total amount of hybridization to the A1 probe was about the same in the HpaII and MspI digestions and that 35-40% of the radioactivity in the HpaII digestion was present in the discrete partial products. For example, about 25% of the radioactivity was present in the 6 kb fragment alone. This fragmentation pattern can only be explained if genomic members are more often digested by HpaII at site a than at the other HpaII sites in or near the promoter. Furthermore, since the A1 specific 0.4 kb HpaII band was recovered in a higher yield than expected from the number of L1 members that lack site b (but contain sites a and c, see above), then a significant number of members must be demethylated at sites a and c but remain methylated at site b. Taken

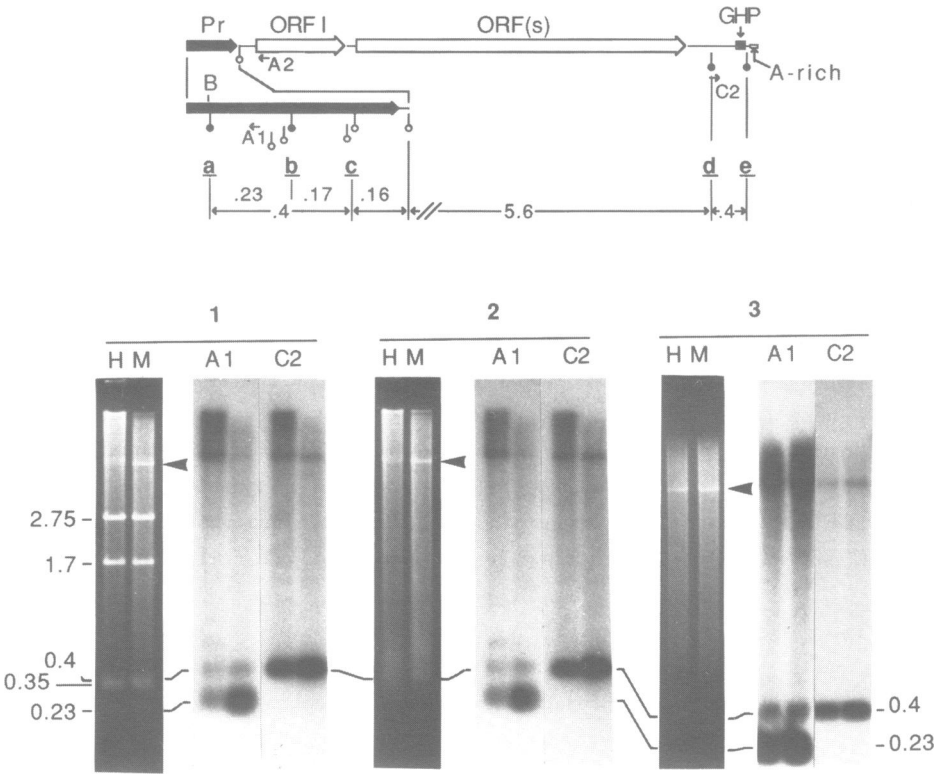

FIG. 4. The effect of 5-azacytidine treatment on R2 cell line DNA.
The upper diagram is the same as that in Fig. 3. The lower part of the figure shows the hybridization of the indicated oligonucleotide to HpaII (H) or MspI (M) digests of DNA from R2 cells exposed to 3, 4, or 3 rounds of treatment (experiments 1, 2, and 3, respectively) with 5-azaC as described in the Materials and Methods. The bands at 2.75, 1.7, and 0.35 kb in the ethidium bromide stained gel of experiment 1 are HpaII digestion products of ØX174 which was added to this digestion to monitor the completeness of the HpaII digestion. In experiment 3 the specific activity of the A1 oligonucleotide was higher then that of the C2 oligonucleotide. In experiments 1 and 2 they had the same specific activity.

together these results indicate that the promoter L1 HpaII sites differ in their susceptibility to demethylation.

In the third experiment shown in Fig. 4, three cycles of 5-azaC treatment produced an apparent complete demethylation of all of the L1 HpaII sites; i.e. the HpaII and MspI digestions produced the same pattern of L1 DNA fragments. As the ethidium bromide stain gels show, most of the non-L1 genomic HpaII sites were also demethylated; again the HpaII and MspI digests are indistin-

TABLE 2. 5 Methyl C content of DNA

Source of DNA	moles % of 5 methyl C / moles of DNA base
Rat liver	1.19 ± 0.15 [a]
Rat muscle	1.29
R2 cells (not treated)	0.79 ± 0.11 [b]
R2 cells (treated) [c]	0.14
R2 cells (after passage) [d]	0.20

[a] Three different preparations of liver DNA.
[b] Five different preparations of DNA from R2 cells not treated with 5-azaC.
[c] Experiment 3, Fig. 5
[d] Cells from experiment 3, Fig. 5, after 3 passages (about 8 doublings).

guishable. Furthermore, blot hybridization with a rat satellite I probe showed that the HpaII sites in rat satellite I (44,45) were also completely demethylated in experiment 3 (results not shown).

It is thought that the methylation state of the CpG in HpaII sites is representative of CpG sites in general (41). However, HpaII digestion monitors only a small percentage of the total CpGs in the genome. Since we detected a different susceptibility to demethylation among just 5 of the L1 CpGs (i.e., those in HpaII sites a, b, c, d, and e), we thought it possible that digestion with HpaII and MspI might not accurately reflect the demethylation state of the genome. Therefore, we examined the total methyl C content of the DNA from experiment 3 directly by HPLC as described in Materials and Methods. The results in Table 2 show about 20% of the CpG sites remain methylated in the DNA from this experiment. The values that we obtained for rat liver or muscle DNA (Table 2) are similar to those for various normal mammalian DNAs [0.9 to 1.4 depending on the genus (46)]. The value obtained with R2 cells that were not exposed to 5-azaC is consistent with earlier observations that established rodent cell lines are partially demethylated when compared to normal rodent somatic cells (also see Fig. 3).

Although demethylation of CpG was not complete, it was quite extensive. We therefore examined the pattern of L1 transcripts in the cells from experiment 3. To do so, total RNA from these cells as well as from untreated R2 cells and normal liver and kidney cells was electrophoresed, transferred to nylon membranes, and hybridized to the probes indicated in the legend to Fig. 5. The results with both the total L1 DNA probe, the A2 oligonucleotide, which would identify sense transcripts beginning at ORF I, and the C2 oligonucleotide which would detect antisense L1 transcripts produced identical

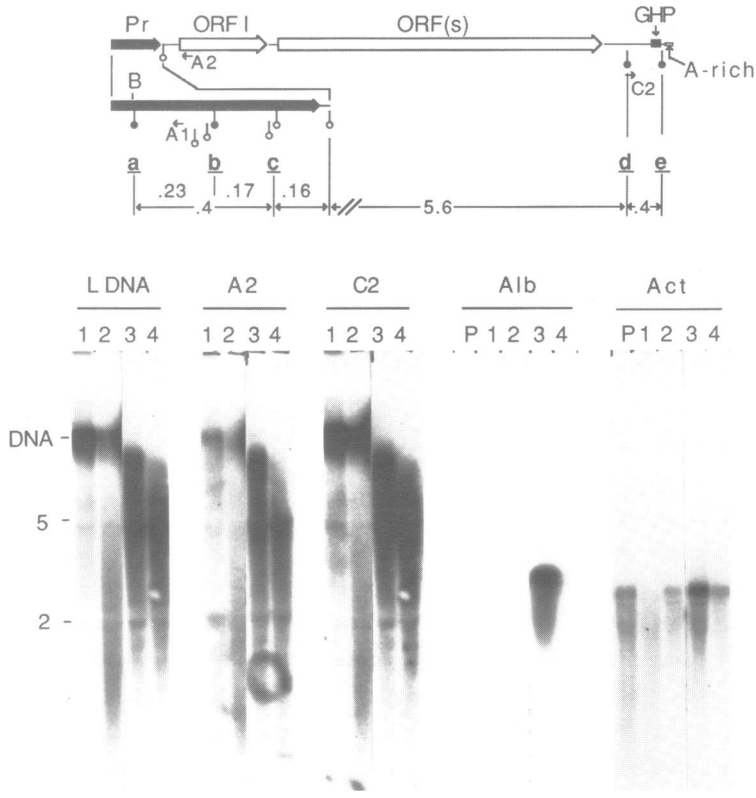

FIG. 5. L1 transcripts in 5-azacytidine treated cells.
The upper diagram is the same as that in Fig. 3. The lower part of the figure shows hybridzation with the indicated probe of total cellular RNA from untreated R2 cells, lane 1; the 5-azaC treated cells of experiment 3 (Fig. 4), lane 2; normal rat kidney and liver, lanes 3 and 4 respectively. Alb and act are probes that contain portions of the coding sequence of the mouse albumin and rat β actin genes respectively and in these panels normal liver RNA is in lane 3 and kidney RNA in lane 4. Lane P contains total RNA from the P113 cell line. The L1 probe was the essentially full length rat L1 element, L 3 (L1Rn 3), that was described previously (11), 5 and 2 refer to the approximate sizes in kb of the 28S and 18S ribosomal RNAs, respectively. The blot used for C2 probe was previously used for the L1 DNA probe which was completely removed as described in the Materials and Methods.

patterns. The demethylated cells, lane 2, while containing an increase in the amount of heterodisperse transcripts that migrate faster than 28S ribosomal RNA (approximately 5 kb), did not contain any novel, discrete sense L1 transcripts. Although we can not rule out the possibility that some of the heterodisperse, sense transcripts detected with the A2 probe may originate

from L1 promoters, we conclude that most of the L1 transcripts in normal rat cells or untreated R2 cells are most likely "read-through" transcripts from non-L1 promoters and that demethylation increases this. Consistent with this conclusion is our detection of satellite I DNA transcripts in the demethylated R2 cells of experiment 3 but not in untreated R2 cells nor in liver or kidney cells (results not shown). The pattern of L1 transcripts in the demethylated cells (Fig. 5, lane 2) reverts to that of the untreated cells (lane 1) when the R2 cell genome becomes remethylated by continued growth in the absence of 5-azaC (results not shown).

We did not detect transcripts of three tissue-specific genes in the demethylated cells: insulin, α fetal protein, or serum albumin. The hybridization with the serum albumin probe is shown in Fig. 5.

DISCUSSION

We have demonstrated here that the 600 bp promoter-like region at the left end of a newly isolated rat L1 element, L B6 (L1Rn B6) (14), can function as a promoter in vivo. Our evidence is that plasmids containing a fusion of this DNA sequence to the bacterial CAT gene produce CAT enzyme in the P113 rat cell line. Enzyme is produced only when the L1 promoter sequence is oriented to the CAT gene in the same orientation as it is to the L1 ORFs (Figs. 1, 2 and results not shown). Furthermore, enzyme production depends only on L1 DNA and not on the non-L1 rat DNA that flanks the left end of L B6 (Table 1, lines 3 and 5).

Cells containing the L1 promoter fusion plasmids produced at least 20% as much CAT enzyme as cells with the SV40 promoter fusion plasmids (Table 1). However, since CAT enzyme production is a function of the steady-state level of CAT gene mRNA, these values only partly reflect the relative strength of these promoters. Experiments not reported here showed that the level of CAT gene transcripts is very low in P113 cells transfected with either the L1 promoter or SV40 promoter plasmids. The instability and low concentration of CAT gene transcripts in certain mammalian cells is known (47), and for these reasons, among others, we are now extending our analysis of L1 promoter fusions to other mammalion cell lines.

The CpG content of the G + C rich rat L1 promoter is generally about that expected from the nucleotide composition (14). Therefore, the rat L1 promoter is a CpG "island", and we examined two aspects of this property of the promoter.

First, methylation of CpG to 5MeCpG has been correlated with the repres-

sion of a number of mammalian promoters [reviewed in (17,33-36)], and we found that methylation of just 5 of the 22 CpGs in the L B6 promoter reduced CAT enzyme production by at least 75%. This inhibition is somewhat dependent on sequence context since deletion of the non-L1 DNA which flanks the promoter and which contains no HpaII sites reduces the inhibitory affect of methylation (Table 1, Lines 4 and 6). Although we have not been able to accurately measure transcript levels in P113 cells (see above), we presume by analogy with the proposals of others (39,40) that the decrease in CAT enzyme production is due to a decrease in promoter activity.

The alternative explanation is that the structure and therefore some property (e.g., stability, translatability, etc.) of the CAT transcripts synthesized from the methylated and nonmethylated plasmids differ. Since both the L1 promoter and SV40 promoter CAT plasmids contain the identical CAT gene coding and 3' non-translated sequence (see Materials and Methods and Fig. 1), and since methylation does not affect CAT enzyme production from the SV40 promoter plasmids (Table 1), then transcripts with an altered structure could only be produced if methylation of the L1 promoter altered the transcriptional start site. Therefore, this explanation would also mean that methylation altered the properties of the L1 promoter.

We also examined the methylation state of L1 promoters in the rat genome and found that the HpaII sites of these promoters are completely methylated in rat liver and sperm DNA and even in a rat tissue culture cell line in which partial demethylation of L1 HpaII sites outside of the promoter region occurred (Fig. 3).

By use of 5-azaC we could demethylate L1 sequences, and in one experiment (experiment 3, Fig. 4) this treatment produced apparent complete demethylation of all of the L1 promoter HpaII sites and most of the genomic HpaII sites as well. The major transcriptional effect of this extensive demethylation was to increase the amount of apparent read-through transcription from non-L1 sequences (Fig. 5). The increased amount of heterodisperse L1 transcripts in the drug-treated cells could have obscured a small amount of discrete sense L1 transcripts, but our failure to detect such transcripts could also mean that the demethylation of the L1 promoters by 5-azaC was not sufficient to activate them. Direct chemical analysis showed that the DNA from experiment 3 still contained about 20% as much 5 MeC as untreated cells, and since genomic L1 promoters contains 3-6 HpaII sites (Fig. 3) but about 20 CpGs, then the L1 promoters could still be partially methylated and therefore repressed. The fact that the methylation in vitro of only about 25% of the CpG sites in the

L B6 promoter reduces CAT enzyme production by about 75% (Table 1) indicates that the L1 promoters are sensitive to partial methylation.

CpG islands are selectively associated with the regulatory regions of genes and are unmethylated in both somatic cell and sperm DNA (15-18, 34). In fact, their unmethylated state is thought to account for the evolutionary persistence of CpG islands in otherwise CpG-deficient genomes since 5 MeC is subject to deamination to T which in time leads to the loss of CpG and the compensatory gain of TpG and CpA [(41) and references therein]. Since the rat L1 promoter is a CpG island but is methylated in somatic cell and sperm DNA, we speculate that the rat L1 promoter only recently became subjected to methylation. Since the L1 promoter is completely distinct from either of the mouse L1 promoter regions (14), then the rat L1 family must have acquired its promoter some time subsequent to the divergence of rats and mice from their common ancestor some 10 million years ago (48). Therefore, it is possible that at the time of its acquisition the present day rat L1 promoter was a nonmethylated CpG island.

We presume that the methylation of the L1 promoter (whether it is a cause or result of promoter repression) is indicative of the fact that present day rats have acquired the ability to repress what previously might have been a constitutively active promoter sequence. Since uncontrolled transposition is deleterious, this ability would be essential to survival of the organism. However, acquisition of a novel constitutively active promoter by an L1 element competent for transposition would allow amplification of a "new" generation of L1 elements with probably dire consequences for most of the extant population of the genera that harbors it. Repopulation would only occur from those animals (descendants) that acquired the ability to repress the newly acquired promoter. Although this scenario is quite speculative, it is consistent with the fact that during mammalian evolution L1 elements have repeatedly acquired promoters that share no sequence homology but are nonetheless CpG islands.

ACKNOWLEDGEMENT

We thank Dr. Shirley Tilghman (Princeton University) for cDNA clones containing mouse α fetal protein and serum albumin sequences and Dr. Bruce Patterson (National Cancer Institute) for the cDNA clone containing rat β actin sequences. We also thank Dr. Michael Cashel (National Institute of Child Health and Development) for useful discussions.

+Present address: Orgenics, Ltd, PO Box 360, Yavne 70650, Israel

*To whom correspondence should be addressed

REFERENCES

1. Burton, F. H., Loeb, D. D., Voliva, C. F., Martin, S. L., Edgell, M. H., and Hutchison, C. A. III (1986) J. Mol. Biol. 187, 291-304.
2. Fanning, T., and Singer, M. F. (1987) Nucleic Acids Res. 15, 2251-2260.
3. Singer, M. F., and Skowronski, J. (1985) Trends Biochem. Sci. 10, 119-122.
4. Economou-Pachnis, A., Lohse, M. A., Furano, A. V., and Tsichlis, P. N. (1985) Proc. Natl. Acad. Sci. USA 82, 2857-2861.
5. Lakshmikumaran, M. S., D'Ambrosio, E., Laimins, L. A., Lin, D. T., and Furano, A. V. (1985) Mol. Cell. Biol. 5, 2197-2203.
6. Katzir, N., Rechavi, G., Cohen, J. B., Unger, T., Simoni, F., Segal, S., Cohen, D., and Givol, D. (1985) Proc. Natl. Acad. Sci. USA 82, 1054-1058.
7. Burton, F. H., Loeb, D. D., Chao, S. F., Hutchison, C. A. III, and Edgell, M. H. (1985) Nucleic Acids Res. 13, 5071-5084.
8. Shyman, S., and Weaver, S. (1985) Nucleic Acids Res. 13, 5085-5093.
9. Soares, M. B., Schon, E., and Efstratiadis, A. (1985) J. Mol. Evol. 22, 117-133.
10. Fujita, A., Hattori, M., Takenaka, O., and Sakaki, Y. (1987) Nucleic Acids Res. 15, 4007-4020.
11. D'Ambrosio, E., Waitzkin, S. D., Witney, F. R., Salemme, A., and Furano, A. V. (1986) Mol. Cell. Biol. 6, 411-424.
12. Loeb, D. D., Padgett, R. W., Hardies, S. C., Shehee, W. R., Comer, M. B., Edgell, M. H., and Hutchinson, C. A. III (1986) Mol. Cell. Biol. 6, 168-182.
13. Hattori, M., Hidaka, S., and Sakaki, Y. (1985) Nucleic Acids Res. 13, 7813-7827.
14. Furano, A. V., Robb, S., and Robb, F. T. (1988) Nucleic Acids Res. 16, accompanying paper
15. Cooper, D. N., Taggart, M. H., and Bird, A. P. (1983) Nucleic Acids Res. 11, 647-658.
16. Bird, A., Taggart, M. H., Frommer, M., Miller, O. J., and Macleod, D. (1985) Cell 40, 91-99.
17. Bird, A. (1986) Nature (London) 321, 209-213.
18. Lindsay, S., and Bird, A. (1987) Nature (London) 327, 336-338.
19. Soberon, X., Covarrubias, L., and Bolivar, F. (1980) Gene 9, 287-305.
20. Gorman, C. M., Moffat, L. F., and Howard, B. H. (1982) Mol. Cell. Biol. 2, 1044-1051.
21. Topp, W.C. (1981) Virology 113, 408-411.
22. Fisher, P. B., Babiss, L. E., Weinstein, I. B., and Ginsberg, H. S. (1982) Proc. Natl. Acad. Sci. USA 79, 3527-3531.
23. Davis, L. G., Dibner, M. D., and Battey, J. F. (1986) Basic Methods in Molecular Biology, Elsevier Science Publishing Co., Inc., New York.
24. Graham, F. L., and Van Der Eb, A. J. (1973) Virology 52, 456-467.
25. Luthman, H., and Magnusson, G. (1983) Nucelic Acids Res. 11, 1295-1308.
26. Neumann, J. R., Morency, C. A., and Russian, K. O. (1987) BioTechniques 5, 444-447.
27. Seed, B., and Sheen, J. Y. (1988) Gene 67, 271-277.
28. Witney, F., and Furano, A. V. (1984) J. Biol. Chem. 259, 10481-10492.
29. Glisin, V., Crkvenjakov, R., and Byus, C. (1974) Biochemistry 13, 2633-2637.
30. Chirgwin, J. M., Przybyla, A. E, MacDonald, R. J., and Rutter, W. J. (1979) Biochemistry 18, 5294-5299.

31. Smith, G. E., and Summers, M. D. (1980) Anal. Biochem. 109, 123-129.
32. McMaster, G. K., and Carmichael, G. G. (1977) Proc. Natl. Acad. Sci. USA 74, 4835-4838.
33. Holliday, R. (1987) Science 238, 163-170.
34. Razin, A., and Cedar, H. (1984) Int. Rev. Cytol. 92, 159-185.
35. Doerfler, W. (1983) Annu. Rev. Biochem. 52, 93-124.
36. Riggs, A. D., and Jones, P. A. (1983) Adv. Cancer Res. 40, 1-30.
37. Stein, R., Razin, A., and Cedar, H. (1982) Proc. Natl. Acad. Sci. USA 79, 3418-3422.
38. Busslinger, M., Hurst, J., and Flavell, R. A. (1983) Cell 34, 197-206.
39. Kruczek, I., and Doerfler, W. (1983) Proc. Natl. Acad. Sci. USA 80, 7586-7590.
40. Langer, K-D., Weyer, U., and Doerfler, W. (1986) Proc. Natl. Acad. Sci USA 83, 1598-1602.
41. Bird, A. P. (1980) Nucleic Acids Res. 8, 1499-1504.
42. Flatau, E. Gonzales, F. A., Michalowsky, L. A., and Jones, P. A. (1984) Mol. Cell. Biol. 4, 2098-2102.
43. Taylor, S. M., and Jones, P. A. (1982) J. Mol. Biol. 162, 679-692.
44. Pech, M., Igo-Kemenes, T., and Zachau, H. G. (1979) Nucleic Acids Res. 7, 417-432.
45. Epstein, D. A., Witney, F. R., and Furano, A. V. (1984) Nucleic Acids Res. 12, 973-988.
46. Fasman, G. D. (1975) Handbook of Biochemistry and Molecular Biology: Nucleic Acids, Vol. II, CRC Press, Cleveland, Ohio.
47. Cullen, B. R. (1987) In Berger, S. L., and Kimmel, A. R. (eds) Methods in Enzymology, Academic Press, New York, Vol. 152, pp. 684-704.
48. Benveniste, R.E, Callahan, R., Sherr, C.J., Chapman, V., and Todaro, G. J. Virol. 21, 849-862.