Ligation-mediated PCR improves the sensitivity of methylation analysis by restriction enzymes and detection of specific DNA strand breaks

Sabine D.Steigerwald, Gerd P.Pfeifer and Arthur D.Riggs

Department of Biology, Beckman Research Institute of the City of Hope, Duarte, CA 91010, USA

Received December 15, 1989; Revised and Accepted February 11, 1990

ABSTRACT

DNA methylation at specific sites is most frequently studied by use of methylation-sensitive restriction endonucleases and Southern blotting. We report here that the sensitivity of this method can be increased several-hundred-fold by applying a ligation-mediated polymerase chain reaction (LM-PCR) procedure following enzyme treatment. DNA is cleaved simultaneously with two restriction enzymes, one sensitive and one insensitive to methylation. After cleavage, a gene-specific oligonucleotide primer is used for primer extension, followed by linker ligation and then conventional PCR. Using this technique, we demonstrate that DNA from 100 cells (about 0.6 ng) can be prepared and qualitatively analyzed for methylation at sites in an X-linked CpG island, and 50 ng of DNA can be analyzed quantitatively. A site 23 bp downstream of the major transcription start site of human phosphoglycerate kinase-1 (PGK-1) is 52 ± 7 percent methylated in DNA from female blood and greater than 98 percent unmethylated in DNA from male blood or HeLa cells. This method detects quantitatively specific breaks in either double stranded or single stranded DNA. Thus new assays for enzymes and DNA structure can be devised.

INTRODUCTION

DNA methylation is part of a transcription silencing system that operates in mammals and other vertebrates (1,2). For example, methylation seems to be an important part of X chromosome inactivation and genomic imprinting (1–5). These phenomena take place in the mammalian embryo at early stages where adequate amounts of tissue are difficult to obtain. Thus these studies, and, for example, studies on tissue samples obtained by microdissection should benefit from improved DNA methylation assays. Southern blot analysis following methylation-sensitive restriction enzyme digestion has been the most informative methylation assay, but it requires about 1 μ g of DNA (about 10^5 cells) per lane. Another assay for DNA methylation at specific sites is direct genomic sequencing (6). Genomic sequencing has the great advantage that methylation at all CpG sites can be assayed, but it has been a technically difficult procedure requiring about 50 μ g of DNA per lane and often giving poor quality data. Application of a ligation-mediated polymerase chain reaction

(LM-PCR) to genomic sequencing has greatly improved this technique (7,8). Nevertheless, even when improved by LM-PCR, genomic sequencing still requires about 1 μ g DNA per lane, in large part due to inescapable statistical limits. To visualize a sequence ladder, between 10 and 100 molecules are needed for each band, and this puts a firm theoretical limit on sensitivity (8). However, we reasoned that if only one specific site were analyzed, then the statistical limit would be lowered and the sensitivity of methylation analysis increased. We report here that this is the case; about 1 ng of DNA from 100 cells can be qualitatively assayed for methylation at specific sites, and 50 ng of DNA can be assayed quantitatively. In its essence, the procedure described here is a high resolution, sensitive, and quantitative assay for specific strand breaks and thus may have broad application.

MATERIALS AND METHODS

DNA samples and DNA isolation

Human-Chinese hamster hybrid cell lines containing either an inactive (X8-6T2) or an active (Y162-11C) human X-chromosome (9) were kindly provided by R.S. Hansen and S.M. Gartler (Seattle, WA). DNA from human blood cells was a gift of J. Singer-Sam (Duarte, CA). HeLa cells were strain S315 obtained from G. Attardi (Pasadena, CA). Nuclei were isolated by published procedures (10), and DNA from nuclei was purified by proteinase K treatment, phenol/chloroform extraction and ethanol precipitation.

DNA from a small numbers of cells (10^4 or less) was isolated by a guanidine-HCl procedure (11) as modified for small samples by Singer-Sam et al (12). Pelleted cells were resuspended in 140 μ l of 6 M guanidinium hydrochloride, 10 μ l 7.5 M ammonium acetate, 10 μ l 20% sodium sarkosyl and 1 μ l proteinase K (20mg/ml), and 2 μ g herring sperm DNA was added as a carrier (bacterial DNA can also be used as a carrier). After incubation at 60°C for 1.5 h, 2.2 volumes of room temperature ethanol were added, and the mixture centrifuged at $14,000 \times g$ for 30 min at room temperature. The pellets were washed twice with 75% ethanol, dried in a Speedvac concentrator and dissolved in water overnight at room temperature.

Oligonucleotide primers and linker

The following gene-specific oligonucleotide primers were used for primer extension and PCR amplification: Primer F1,

5'-CGTCCAGCTTGTCCAGC; primer F2, 5'-TCCAGCGTC-AGCTTGTAGAAAGCG; primer C1, 5'-ATTCCAGGGG-TTGGGGT; primer C2, 5'-GGGTTGGGGTTGCGCCTT-TTCAA. Two oligonucleotides formed the linker (7), a 25-mer (5'-GCGGTGACCCGGGAGATCTGAATTC) and an 11-mer (5'-GAATTCAGATC). These were annealed in 250 mM Tris-Cl, pH 7.7, by heating to 95 °C for 3 min and gradually cooling to 4 °C over a time period of 3 h. The linker-primer is the longer oligonucleotide of the linker.

Restriction endonuclease digestion for methylation analysis

For analysis of a CCGG site at position +23 with primer set F, genomic DNA was simultaneously digested with HpaII and HaeIII (see Figure 1). For analysis of a GGCGCC site at position -343 with primer set C, DNA was simultaneously digested with NarI and NdeII. Restriction enzyme digestion was performed for 16 h in the appropriate reaction buffers as recommended by the supplier (Bethesda Research Laboratories) with at least 4-fold excess of enzyme over DNA. The completeness of the restriction digest was controlled by adding lambda DNA to an aliquot of the reaction mixture. After restriction digestion, the reaction was terminated by addition of EDTA to a final concentration of 20 mM, phenol extracted and ethanol precipitated. The DNA pellets were dissolved in water to a final concentration of 100 ng/ μ l. When DNA was isolated from small numbers of cells, the DNA was dissolved in 10 μ l water.

Primer extension, ligation, and amplification

DNA (0.7 ng to 2 μ g) was mixed in a final volume of 24 μ l with 0.6 pmoles of a gene-specific primer (C1 or F1) in SequenaseTM reaction mix (8). After denaturation at 95°C for 3 min, the primer was annealed at 45°C for 30 min; 4.5 units SequenaseTM (SequenaseTM 2.0, USB) per sample were added and primer extension was performed for 20 min at 45°C. The reaction was terminated by addition of 6 μ l of 310 mM Tris-HCl, pH 7.7, and heating to 67 °C for 15 min. 100 pmoles of the linker was added together with 3 units of T4 DNA ligase (Promega) and 45 μ l of ligation mix containing 13.3 mM MgCl₂, 30 mM dithiothreitol, 1.66 mM ATP, and 83 μ g/ml bovine serum albumin. The ligation was performed at 17 °C for 16 h. After heat denaturation of DNA ligase (10 min, 70 °C), 10 μ g *E. coli* tRNA was added and the DNA was ethanol precipitated. The fragments were amplified in 50 μ l 10 mM Tris-HCl, pH 8.9, 40 mM NaCl, 2 mM MgCl₂, 0.01% gelatin, 0.25 mM dNTPs, 10 pmoles of gene-specific primer (C2 or F2), 10 pmoles of linker primer, with 3 units *Taq* polymerase (AmplitaqTM, Cetus), according to the Cetus-Perkin Elmer protocol. 20 thermal cycles were done (1 min at 95 °C, 2 min at 67 °C and 3 min at 75 °C with an extension of 5 sec of the last segment in each cycle). When starting with 10³ or smaller numbers of cells or less than 10 ng DNA, respectively, 25 or 30 cycles were performed. A final cycle was done by adding dNTPs (final concentration 100 μ M), 5 pmoles of each primer and 1 unit *Taq* polymerase per sample and incubating at 95°C for 2 min, 67°C for 2.5 min and 75°C for 10 min. One third of each sample was analyzed by electrophoresis on 2.2% agarose gels and alkaline blot transfer to GeneScreen nylon membranes (13). For hybridization, labeled cDNA probes were reverse-transcribed from RNA-templates (14). Hybridization probes were made from fragments XmaIII-PvuII (150 bp; -69 to +81) or StyI-ApaI (142 bp; -408 to -267) for primer sets F and C, respectively. In general, double-stranded probes or labeled oligonucleotides should be adequate

for hybridization, especially since some additional PCR cycles could be done if needed. However, the hybridization probes should not overlap with the primer sequences. After hybridization (8), nylon filters were washed at 70°C in 10 mM sodium phosphate (pH 7.2), 1 mM EDTA, 1% SDS and exposed for 0.5 to 15 h to Kodak XAR-5 film with intensifying screens at -70°C. A video densitometer (Biorad Model 620) together with programs for peak integration (The 1-D Analyst programs, Biorad) were used for densitometric analysis from exposures obtained without intensifying screens.

RESULTS

Description of the test system and methylation assay

Located at the 5' end of the human X-linked phosphoglycerate kinase-1 (PGK-1) gene is a CpG rich island that contains the promoter. Previous studies (9,15) have established that this island, in contrast to autosomal islands which are characteristically unmethylated, is heavily methylated on the inactive X chromosome but is unmethylated on the active X chromosome. To test the LM-PCR approach, we studied two sites recognized by methylation-sensitive restriction endonucleases. The first is a NarI site (Figure 1A) at nucleotide position -343 relative to the start of transcription (16); the second is a HpaII site at nucleotide position +23, the methylation of which has been shown to inversely correlate with gene activity (9,15). For the analysis of the NarI site, we cleaved DNA simultaneously with NarI and NdeII (nt -292) and used primer set C. To study the HpaII site, we digested DNA simultaneously with HpaII and HaeIII (nt -20) and used primer set F (Figure 1A). As illustrated in Figure 1A, if the site for the methylation-sensitive enzyme

Figure 1: Outline of the test system and method. (A) The 5' region of the X-linked human PGK-1 gene. The oligonucleotide primer sets C and F used for LM-PCR are indicated by arrows, and the analyzed restriction sites and PCR products are shown. The asterisk represents the transcription start site. (B) The ligation-mediated amplification method. The open bars represent single DNA strands. Oligonucleotide primers are shown as shaded or black boxes.

(HpaII or NarI) is methylated, then a full length fragment is produced after primer extension; if the internal methylation-sensitive site is completely unmethylated, then only a shorter fragment is seen. If methylation is partial, then the ratio of long fragment to short fragment is a measure of the percent methylation of the internal site.

Figure 1B further illustrates the procedure. Most restriction enzymes leave on each fragment a 5' phosphate, which is essential for successful ligation following primer extension (steps 2 and 3, Figure 1B). In more detail, primer 1 (primer C1 or F1) is annealed to the denatured, enzymatically cleaved DNA and extended using SequenaseTM. This creates blunt-ended molecules which then can be ligated to a linker. The introduction of a common linker sequence (7) allows the use of this sequence as a linker-primer together with a second gene-specific primer (primers C2 or F2) in a PCR amplification system (steps 4 and 5, Figure 1B). Primers 1 (e.g., C1 or F1) are usually 17- to 18mers with a calculated T_m of 50–54°C, primers 2 (e.g., C2 or F2) are 23 to 26 nucleotides long and have a calculated T_m of 74–80°C. PCR amplification was routinely performed for 20, 25 or 30 cycles depending on the amount of DNA used as starting material. The DNA fragments were subsequently separated on an agarose gel, alkaline blotted to a nylon membrane, and hybridized with a probe which does not overlap the primer sequences.

DNA methylation analysis and quantification

Figure 2, lanes 12 and 13 show results from application of the method to 50 ng of human female blood DNA. Both long and short PCR products (168 bp and 126 bp) are seen, as expected for methylation only on the inactive X chromosome. Lanes 1–11 of Figure 2 provide data for a standard curve (see Figure 3) allowing determination of the percent methylation of the HpaII site at nucleotide position +23. HeLa cell DNA is unmethylated at site +23 (ref. 8 and Figure 4), and therefore can be used to construct a standard curve by cleaving separately with either HpaII or HaeIII, and mixing different amounts of the two reactions prior to ligation and amplification. An equimolar mixture (Figure 2, lane 6) shows a stronger band for the longer fragment. The difference in band intensities between the longer fragment and the shorter fragment was seen previously in genomic sequencing studies, is independent of carrier DNA, and is most likely due to varying ligation efficiencies (8). This difference is reproducible and a standard curve can easily be

established for the site to be analyzed. Densitometric measurements in the linear range of the films were performed for the autoradiographic data in Figure 2 and the ratio of the two bands is plotted in Figure 3. Using the standard curve, we find that methylation of female blood DNA (Lanes 12 and 13 of Figure 2) is 52 ± 7 percent, in excellent agreement with published data (15,17) and consistent with methylation of only the inactive X chromosome at this site. In Figure 4, male human blood DNA or HeLa cell DNA cut with both HaeIII and HpaII shows only the shorter fragment (lanes 1 to 4), indicating that this site is at least 98 percent unmethylated in HeLa cells and in male white blood cells.

Analysis of DNA from cell lines containing human active and inactive X-chromosomes

Methylation of sites in the human PGK-1 promoter was also determined in DNAs from human-Chinese hamster hybrid cells containing the human inactive (line X8-6T2) or active (line

Figure 3: Standard curve. Two experiments such as shown in in Figure 2 were done. Peak areas were determined by densitometry of the autoradiograms and the ratio of peak areas of the longer versus the shorter fragment is plotted against the percentage methylation.

Figure 2: Autoradiographic data for standard curve and methylation analysis of PGK-1 site +23 in female blood DNA. Two different samples of female blood DNA (50 ng) were treated with both HpaII and HaeIII, and analyzed by LM-PCR as described in the text (lanes 12 and 13). HeLa DNA is not methylated at site +23 (see Figure 4) and therefore could be used to generate a standard curve (Lanes 1–11). HeLa DNA was cut with either HaeIII only or HpaII only, and the cut DNA samples were mixed (50 ng total) simulating methylation levels of 0% (lane 1), 10% (lane 2), 20% (lane 3), 30% (lane 4), 40% (lane 5), 50% (lane 6), 60% (lane 7), 70% (lane 8), 80% (lane 9), 90% (lane 10) and 100% (lane 11).

Figure 4: Methylation analysis of PGK-1 site +23 in male blood DNA and HeLa DNA. Lanes 1 and 2: male human blood DNA cut with HpaII and HaeIII; lanes 3 and 4: HeLa DNA cut with HpaII and HaeIII; lanes 5 and 6: HeLa DNA cut with HaeIII only; lanes 7 and 8: HeLa DNA cut with HpaII only. In lanes 1,3,5, and 7, 100 ng DNA was used for the LM-PCR assay. In lanes 2,4,6, and 8, only 10 ng DNA was used.

Figure 5: Methylation analysis of human PGK-1 NarI site -343 in human-hamster hybrid cells. DNA from hybrid cells containing either an active (lanes 1 to 3) or an inactive human X-chromosome (lanes 4 to 6) was subjected to simultaneous cleavage with both NarI and NdeII (lanes 2,3,5,6) or with NdeII alone (lanes 1,4). Primer set C was used for amplification (20 cycles). Lanes 1,2,5: 100 ng DNA, lane 4: 50 ng DNA, lanes 3,6: 10 ng DNA.

Figure 6: Methylation analysis of human PGK-1 HpaII site +23 in hamster-human hybrid cells. DNA from hybrid cells containing either an active (lanes 1 to 4 in A; lanes 1 to 3 in B) or an inactive human X-chromosome (lanes 5 to 8 in A; lanes 4 to 6 in B) was cleaved with HaeIII and HpaII (lanes 2 to 4 and 6 to 8 in A; all lanes in B) or with HaeIII alone (lanes 1 and 5 in A). Primer set F was used for amplification. A) Lanes 1,2,5 and 6: 100 ng DNA (20 cycles); lanes 3,7: 10 ng DNA (20 cycles); lanes 4,8: 1 ng DNA (30 cycles). B) The same experimental system was used starting with isolation of DNA from a small number of cells. Lanes 1 and 4: 10^4 cells (20 cycles), lanes 2 and 5: 10^3 cells (25 cycles), lanes 3 and 6: 10^2 cells (30 cycles).

Y162-11C) X-chromosome. In Figure 5 the methylation of the NarI site at nucleotide position -343 was studied by cutting the DNA with NarI and NdeII and using primer set C for amplification (see Figure 1A). The longer fragment is seen after cutting with NdeII alone (Figure 5, lanes 1 and 4). Only the shorter fragment is seen in DNA from line Y162-11C by cutting with NdeII and NarI simultaneously (Figure 5, lanes 2 and 3) whereas only the longer fragment appears in DNA from cell line X8-6T2 (Figure 5, lanes 5 and 6). This means that the NarI site is methylated on the inactive X chromosome and unmethylated on the active X chromosome in these hybrid cell lines. Figure 6A shows similar data for HpaII site +23 in the same DNA samples. Lanes 2 to 4 and lanes 6 to 8 (Figure 6A) show decreasing DNA amounts cut with HaeIII and HpaII and demonstrate that DNA amounts as low as 1 ng are sufficient to obtain methylation information.

Analysis of a small number of cells

Figure 6B shows LM-PCR data for DNA prepared from a small number of cells by a convenient, rapid procedure (11,12), which includes lysis and proteinase K treatment in guanidine-HCl, room temperature precipitation with carrier DNA, and no phenol extraction. The DNA was cut with HaeIII and HpaII simultaneously and primer set F was used for amplification. Lanes 1 to 3 of Figure 6B show results from 10^4 , 10^3 , and 10^2 hybrid cells containing the active human X-chromosome. Lanes 4 to 6 show data for hybrid cells containing the human inactive X-chromosome. As expected, the HpaII site at position +23 is unmethylated on the active X-chromosome and methylated on the inactive X-chromosome. The data show that 100 cells are sufficient to obtain information about DNA methylation at this site.

DISCUSSION

The LM-PCR procedure (7,8), as described here for methylation-sensitive restriction enzymes, permits the study of DNA methylation patterns in nanogram amounts of purified DNA, and we show that it can also be applied to samples prepared from a small number of cells. The procedure uses nested oligonucleotide primers on one side and a specific cut (or nick) on the other side to specify accurately the PCR product. In favorable cases, the extra level of specificity provided by the second, nested gene-specific primer may not be necessary, but its use assures low background. The method is almost 1000 times more sensitive than conventional Southern blotting, detecting specific cuts or nicks on DNA from about 100 cells. Because of the exponential PCR step, band detection does not limit sensitivity; rather sensitivity is limited by statistical sampling factors, as previously discussed (8). 100 diploid female cells should provide 200 molecules of a target sequence. However, only those molecules that successfully complete primer extension and linker ligation participate in the PCR reaction. Our results suggest about 5 to 10% overall efficiency of extension, ligation and amplification for the fragments studied here, since experiments using less than 1 ng DNA have shown occasional missing bands, consistent with sampling fluctuations expected for only a few molecules. Previous genomic sequencing studies showed that it is possible to include multiple primer sets in the reaction; therefore information about multiple methylation sites can be obtained in one experiment by rehybridization of the nylon membrane (8).

Although the LM-PCR approach demonstrated here requires two additional steps, it is much more sensitive than Southern blotting and automatically provides an internal standard, making it suitable for quantitation of, for example, methylation at specific sites. Quantitation should be rather insensitive to absolute efficiencies, since it depends on relative efficiencies between two bands. Genomic sequencing (8) and footprinting (7) studies have firmly established that the relative intensity of bands generated by LM-PCR is very reproducible and an intrinsic property of the sequence. With cloned DNA as source of unmethylated DNA or by using a methylation-insensitive restriction enzyme (e.g. *MspI*), a standard curve easily can be obtained for any chosen site. The method provides a positive signal for both the methylated and unmethylated state of the analyzed site. This allows a quantification of either the methylated or unmethylated (uncut or cut) state of a site. For example, we have been able to determine that DNA from HeLa and a human-hamster cell line with an active human X chromosome is greater than 98 percent unmethylated (no long fragment) at PGK-1 site +23, whereas DNA from a human-hamster cell line with an inactive human X chromosome is at least 98 percent methylated (no short fragment) at this same site. Another PCR based methylation assay (17), which uses two primers which flank the recognition site of a methylsensitive restriction enzyme, has recently been improved by use of an internal standard (12). This method is more sensitive and can provide information on absolute concentrations. However, the standard must be specially made for each site to be assayed and added to the reaction, and an unmethylated (cut) fragment does not give a positive signal.

The LM-PCR procedure described here is a very sensitive method for detecting and quantifying specific single strand breaks or nicks. Therefore, several other applications are obvious. These should include studies on high-resolution mapping of hypersensitive sites and other techniques employing exogenous or endogenous DNases. It should be possible to study enzymes or chemical agents that do not leave 5-phosphates by adding a kination step. Thus quantitative assays with automatic internal standards can be devised for numerous enzymes, including aglycosylases, DNA methyltransferases and demethylases.

ACKNOWLEDGEMENTS

We give special thanks to Judith Singer-Sam for advice and for communicating to us prior to publication the procedure for isolating DNA from a small number of cells. This work was supported by a National Institute of Aging grant (AG08196) to A.D.R. and a fellowship from the Deutsche Forschungsgemeinschaft (Pf212/1-1) to G.P.P.

REFERENCES

1. Riggs, A.D. and Jones, P.A. (1983) *Adv. Cancer Res.* 40, 1–30.
2. Grünwald, S. and Pfeifer, G.P. (1989) *Prog. Clin. Biochem. Med.* 9, 61–103.
3. Monk, M. (1988) *Genes Dev.* 2, 921–925.
4. Solter, D. (1988) *Annu. Rev. Genet.* 22, 127–146.
5. Reik, W. and Surani, M.A. (1989) *Nature* 338, 112–113.
6. Church, G.M. and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* 81, 1991–1995.
7. Mueller, P.R. and Wold, B. (1989) *Science* 246, 780–786.
8. Pfeifer, G.P., Steigerwald, S.D., Mueller, P.R., Wold, B. and Riggs, A.D. (1989) *Science* 246, 810–813.
9. Hansen, R.S., Ellis, N.A. and Gartler, S.M. (1988) *Mol. Cell. Biol.* 8, 4692–4699.
10. Wijnholds, J., Philipsen, J.N.J. and AB, G. (1988) *EMBO J.* 7, 2757–2763.

11. Jeanpierre, M. (1987) *Nucl. Acids Res.* 15, 9611.
12. Singer-Sam, J., LeBon, J.M., Tanguay, R.L. and Riggs, A.D. (1989), *Nucleic Acids Res.*, in press.
13. Reed, K.C. and Mann, D.A. (1985) *Nucl. Acids Res.* 13, 7207–7221.
14. Weih, F., Stewart, A.F. and Schütz, G. (1988) *Nucl. Acids Res.* 16, 1628.
15. Keith, D.H., Singer-Sam, J. and Riggs, A.D. (1986) *Mol. Cell. Biol.* 6, 4122–4125.
16. Singer-Sam, J., Keith, D.H., Tani, K., Simmer, R.L., Shively, L., Lindsay, S., Yoshida, A. and Riggs, A.D. (1984) *Gene* 32, 409–417.
17. Singer-Sam, J., Yang, T.P., Mori, N., Tanguay, R.L., LeBon, J.M., Flores, J.C. and Riggs, A.D. (1989) in Clawson, G., Willis, D., Weissbach, A., Jones, P.A. (eds.) *UCLA Symposia on Molecular and Cellular Biology*, 'Nucleic Acid Methylation', Alan Liss, New York, New Series, Vol. 128, in press.