ELSEVIER

Methods of Nutritional Biochemistry

A method to estimate the percent loss of cytosine methyl groups at defined CpG sites in liver DNA from methyl-deficient rats

Igor P. Pogribny and S. Jill James

Division of Biochemical Toxicology, National Center for Toxicological Research, Jefferson, AR 72079

Dietary methyl deficiency provides an ideal in vivo model system in which to study progressive alterations in DNA methylation patterns as they occur during multistage hepatocarcinogenesis. Weanling male F344 rats were given a semipurified diet deficient in the methyl-donors choline, methionine, and folic acid for a 36-week period with sampling intervals at 3, 9, 24, and 36 weeks. Using a genomic sequencing procedure based on the PCR amplification of bisulfite-modified DNA, the methylation status of individual CpG sites within exons 6 and 7 of the p53 gene in liver samples from control and deficient rats was assessed. Treatment of denatured nuclear DNA with sodium bisulfite converts unmethylated cytosine residues to uracil, which are then amplified as thymine in the PCR reaction. In contrast, methylated cytosines are resistant to bisulfite deamination under these reaction conditions and are amplified as cytosine. In this report, we describe a novel application of automated sequencing technology to estimate the proportion of methylated cytosines present at defined CpG sites within the total population of DNA molecules extracted. Using the bisulfite conversion-PCR genomic sequencing method, we demonstrate the validity of peak height analysis of co-eluting peaks in the autosequencer electrophoregram to estimate the percent methylation at a defined CpG site. The sensitivity of this method is demonstrated by the progressive loss of methyl groups at a defined CpG site in the methyl-deficient rats after 9, 24, and 36 weeks. The application of this sequence-specific technology will allow site-specific definition of the methylation status of each CpG site within a coding sequence or promoter region and should provide new insights into mechanisms and consequences of methylation dysregulation as a result of dietary deprivation of methyl donors. (J. Nutr. Biochem. 8:355–359, 1997) © *Elsevier Science Inc. 1997*

Keywords: 5-methyl cytosine; methylation; automated DNA sequencing

Introduction

Current understanding of the role of DNA methylation in normal and pathogenic processes is largely based on experimental data derived from the use of restriction endonucleases that cleave DNA depending on the methylation status of a defined CpG site. This approach is inherently limited in that only a small percentage of the total methylcytosines occur at these unique enzyme restriction sites and these may not be representative of the complete methylation profile of the gene. Recently, new methodology has been introduced that allows the determination of the methylation status of every cytosine residue within a defined sequence.[1] This method is based on bisulfite-mediated conversion of non-methylated cytosine to uracil followed by PCR amplification of the bisulfite-modified strands and DNA sequencing. Treatment of denatured nuclear DNA with sodium bisulfite

Financial support for this work was from the American Cancer Society Research Grant CN-73C, SJJ.
Address reprint requests to: Dr. S. Jill James at Division of Biochemical Toxicology, National Center for Toxicological Research, 3900 NCTR Rd., Jefferson, AR 72079 USA.
Received December 3, 1996; accepted February 21, 1997.

655 Avenue of the Americas, New York, NY 10010

0955-2863/97/$17.00
PII S0955-2863(97)00018-1

quantitatively converts all cytosine residues to uracil, which are then amplified as thymine in the PCR reaction. In contrast, 5-methylcytosine is resistant to bisulfite deamination under these reaction conditions and is amplified as cytosine. Manual or automated sequencing of bisulfite-modified DNA will then elucidate the methylation status of every cytosine residue within a defined gene sequence. With automated DNA sequencing, the bisulfite-modified and bisulfite unmodified cytosines at a given site will co-elute in the electrophoregram, as thymine and cytosine, respectively. Relative peak height analysis provides the potential to estimate proportion of cytosines that are methylated at each site within the total population of DNA molecules extracted. In this article, we describe the application of autosequencing with fluorescent dye terminator methodology to estimate the extent of partial methylation at each CpG site within exons 6-7 of the p53 gene. To validate the accuracy of peak height analysis for this purpose, the PCR product obtained from bisulfite-modified DNA was cloned into a pCRII vector. Twenty individual clones were sequenced and the proportion of clones with methylated cytosines at each CpG site was determined as reference control. The proportion of methylated cytosines obtained by cloning and sequencing was then compared with the calculated estimate obtained by peak height analysis. The progressive hypomethylation of cytosines within the hepatic p53 gene was quantified in liver DNA extracted from methyl-deficient rats to verify the sensitivity and utility of the peak height analysis.

Methods and materials

Rats and diets

Weanling male F344 rats were housed two per cage in a temperature-controlled (24°C) room with a 12-hr light/dark cycle, and given ad libitum access to water and NIH-31 pelleted diet. At 50 g body weight (approximately 4 weeks of age), rats were randomly allocated to receive either a diet without added folic acid, low in methionine (0.18%), and lacking in choline (Dyets, Inc., Bethelehem, PA USA) or the same diet supplemented with 2 mg/kg folic acid, 0.4% methionine, and 0.3% choline as control. The omission of folic acid from the low methionine/choline-deficient diet was designed to enhance the severity of methyl group deficiency in this semipurified diet preparation. Four rats per diet group were killed by exsanguination under light ether anesthesia at 3, 9, 24, and 36 weeks after diet initiation. The livers were excised and frozen immediately in liquid nitrogen for subsequent DNA extraction as described by Ausebel.[2]

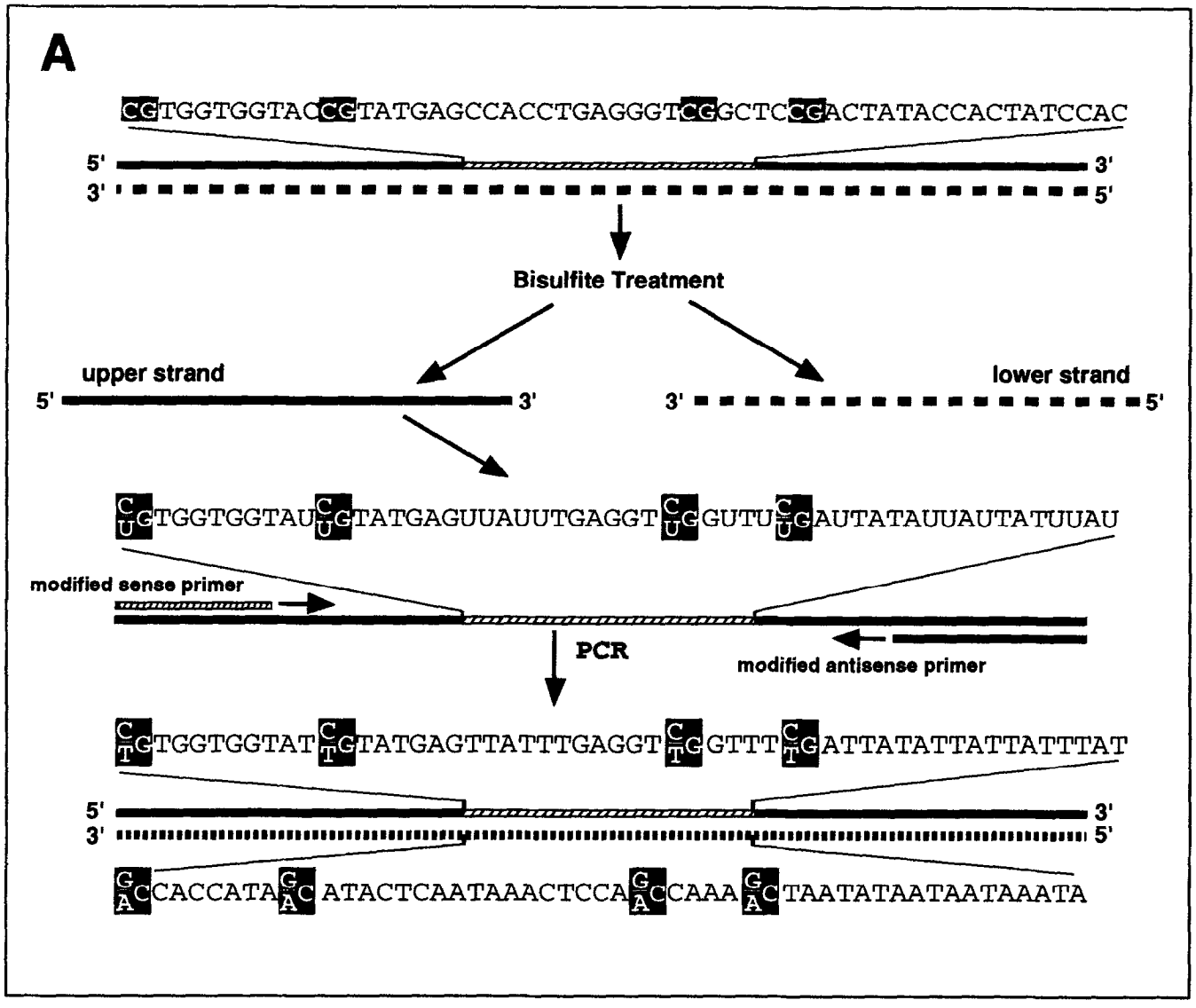
Bisulfite-genomic sequencing

The bisulfite-mediated cytosine deamination method as described by Clark et al. was used in the present study with a minor modifications.[1] The bisulfite-modified DNA template was amplified by PCR in a reaction mix containing 10 μL bisulfite-modified DNA, PCR buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl), 1.5 mM $MgCl_2$, 0.2 mM deoxynucleotide triphosphates, 100 pmol of each primer, and 5 U of AmpliTaq DNA polymerase (Perkin-Elmer Cetus, Norwalk, CT USA) in a final volume of 50 μL. After bisulfite treatment, each strand of DNA must be amplified separately with two new sets of primers specific to each strand, because the conversion of cytosine to uracil creates new and noncomplimentary strands. The PCR primers of exons 6-7 of the rat p53 gene for the amplification of the bisulfite-modified DNA were designed according to the recommendations of Clark et al.[1] The primers for upper strand were sense primer 5′ GTTTTTGATTTATTTTTT-GTTTTTAGGTTTG 3′, antisense 5′ AACTAAAATCTTCCAA-CATAATAATAATAA 3′. The primers for lower strand were sense 5′ CTAACTCCTCCCCAACATCTTATCCAAATA 3′ and antisense 5′ ATTGGAGTTTTTTAGTGTGATGATGGTAAG 3′. The amplification procedure consisted of 2 min of denaturation at 95°C, 30 cycles of 1 min at 95°C, 1.5 min at 60°C, and 2 min at 72°C, and 10 min of extension at 72°C. To verify the quality of the amplification, a 5 μL aliquot of the PCR product was evaluated by electrophoresis in 4% NuSieve GTG agarose gel (FMC BioProducts, Rockland, ME USA) containing ethidium bromide. The PCR products amplified from coding and noncoding strand were sequenced directly using the ABI PRIZM Dye Terminator Cycle Sequencing Kit with AmpliTaq DNA Polymerase, FS (Perkin-Elmer-ABI, Foster City, CA USA) and an ABI PRISM 377 autosequencer. To clone PCR products, the pCRII vector provided with the TA cloning kit from Invitrogen (San Diego, CA USA) was used according to manufacturer's instructions and 20 individual clones were directly sequenced.

A flow diagram of the bisulfite conversion-PCR sequencing method is presented in *Figure 1a*. The bisulfite method of mapping 5-methylcytosine exploits the fact that sodium bisulfite exposure at low pH reacts efficiently with nonmethylated cytosine residues in single-stranded DNA, but does not react with methylated cytosines. In the presence of H_2O, the bisulfite adduct of cytosine undergoes hydrolytic deamination yielding the sulfonated uracil. Desulfonation during alkali treatment results in the formation of uracil which is recognized as thymine by the Taq polymerase in the subsequent PCR reaction. The resultant PCR product sequence will read thymine at sites of unmethylated cytosines and cytosine only where 5-methylcytosine was present in the original DNA template. After bisulfite treatment, peak height analysis of co-eluting peaks of thymine and cytosine at the same sequencing position was used to estimate percent methylation at that site. The raw sequencing data from the bisulfite modified PCR product derived from the upper strand of exons 6-7 of the rat p53 gene is presented in *Figure 1b*. Cytosine residues located in non-CpG sites were used as internal control for the bisulfite conversion; 100% of these unmethylated cytosines were all converted to uracil and amplified as thymine. Only the methylated cytosines at CpG sites were read by Taq DNA polymerase as cytosines, although some co-eluting signal from thymine was always observed. A certain degree of heterogeneity in the degree of methylation at a given CpG site is to be expected, because constitutive processes such as DNA replication and repair will transiently alter methylation status.

Results and discussion

Theoretically, the co-eluting peaks from cytosine and thymine residues at the same sequencing position should reflect the population average of methylated and nonmethylated cytosines at this site. Using this assumption, the relative peak height of cytosine and thymine at the same position can be used to determine the percent methylation at that position. A possible confounder is the fact that the net rate of dideoxynucleotide incorporation by Taq DNA polymerase using dye-terminator chemistry is dependent on local sequence context and can result in sequence-dependent peak height variation.[3,4] To test whether peak height estimation of cytosine methylation is independent of sequence context and variations in peak height, complimentary strands were analyzed in opposite directions using sense and antisense



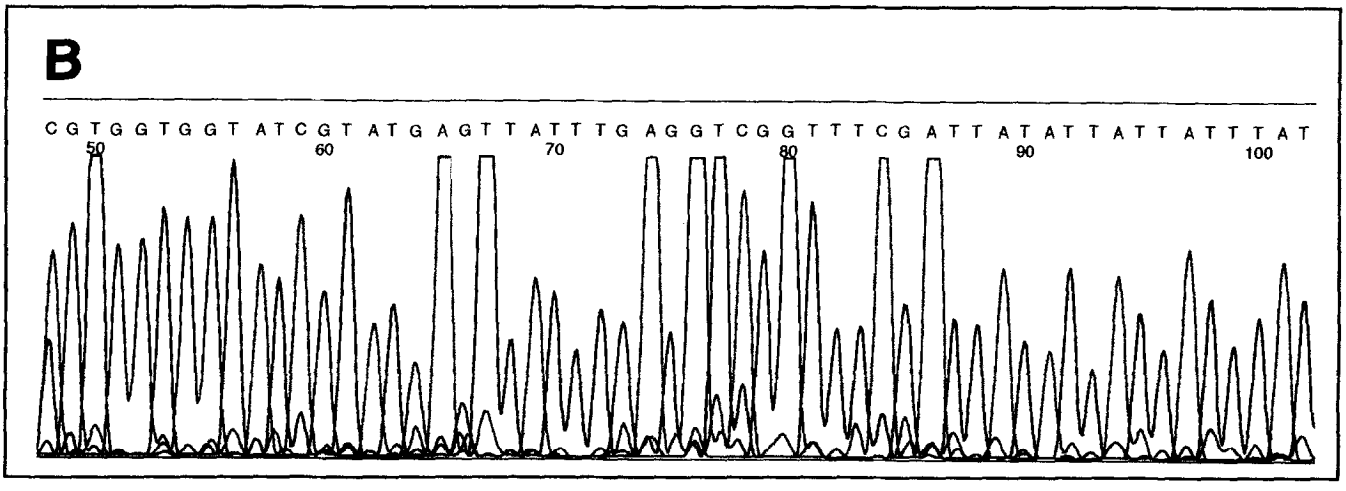


Figure 1 Sequence alterations with bisulfite treatment of genomic DNA. *A*, Flow diagram of the bisulfite-induced alterations at CpG sites within exons 6-7 of the rat hepatic p53 gene. A second set of primers for each strand is required and must be designed to accommodate the bisulfite-mediated conversion of cytosine to uracil (note that the bottom strand created is unique). *B*, An example of the autosequencer raw data showing the bisulfite-modified PCR product derived from the coding strand of exons 6-7 of the p53 gene.

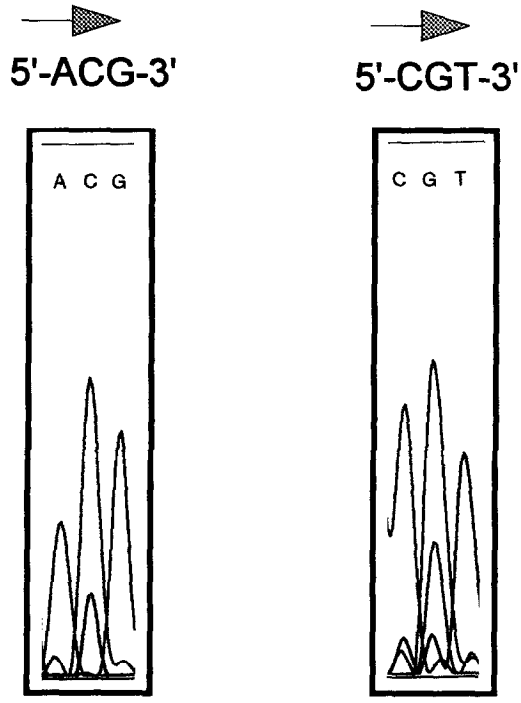


Figure 2 Coeluting peaks at CpG sites using sense and antisense primers. An example of C/T co-elution using the sense primer and an example of G/A co-elution on the complimentary strand using the antisense primer.

primers for the sequencing reaction. Hypomethylated hepatic DNA obtained from rats fed the methyl-deficient diet was used as DNA template.[5] An example of C/T co-elution using sense primer and G/A co-elution on the opposite strand using antisense primer is shown in *Figure 2*. In *Table 1*, the proportion of methylated cytosines at each of the nine CpG sites on complimentary strands of exons 6-7 of the p53 gene obtained by peak height analysis after bisulfite treatment is presented. The peak height analysis estimate is

Table 1 Determination of the proportion of methylated sites within exons 6-7 of the hepatic p53 gene by cloning analysis and by direct sequencing analysis of PCR product[a]

	% of 5-methylcytosine residues		
		Peak height analysis[c]	
Sequence	Cloning analysis[b]	C (C/T)	C (G/A)
CCG	15	12	10
CGA	35	38	24
CGG	45	40	33
CGT	35	43	29
CGG	40	47	30
CGA	35	33	22
CGC	50	39	29
CGG	95	90	75

[a] The hypomethylated DNA template was analysed from rat liver obtained after 36 weeks of folate/methyl deficiency.[5] In control rat liver, 80–95% of cytosines at CpG sites of p53 are methylated.
[b] % of methylated cytosines within exons 6-7 of p53 gene calculated from sequencing of 20 individual clones.
[c] Percent of methylated cytosines = C/(C + T) × 100%.

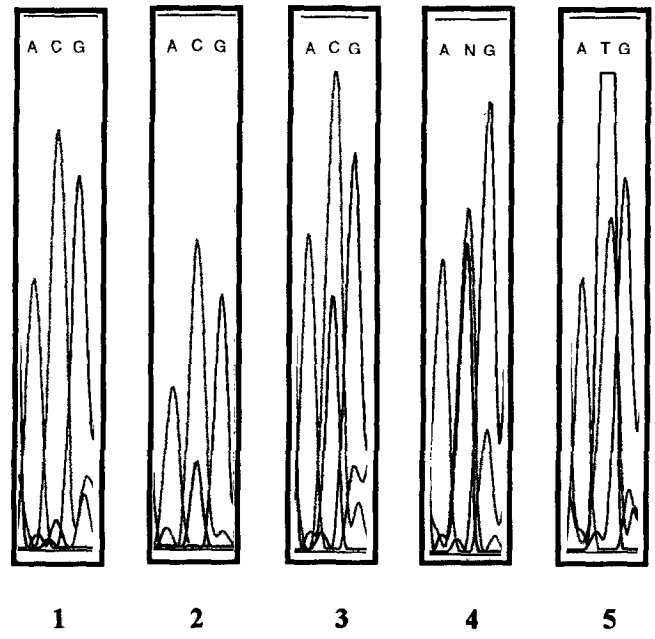


Figure 3 Progressive increase in cytosine hypomethylation at a CpG dinucleotide with dietary intervention. The progressive increase in the coelution of thymine with cytosine at an ACG site within the hepatic p53 gene from rats fed a methyl-deficient diet. *Panel 1* is from a control rat; panels 2–5 are from rats fed a folate/methyl-deficient diet for 3, 9, 24, and 36 weeks, respectively.[5] The ability to detect progressive hypomethylation (increased proportion of thymine) with increased time on the methyl deficient diet confirms the sensitivity of peak height analysis for the estimation of site-specific methylation status.

compared with the percent methylation at the same sites determined by sequence analysis of 20 independent clones containing the same PCR product amplified after bisulfite amplification. Using the proportion obtained from the cloned products as reference control, the calculated estimate derived from the sense primer was found to be more accurate than that obtained from the antisense primer. The mean error in the peak height estimate (averaged over the nine CpG sites) was 5.67% using the sense primer and 13.1% using the antisense primer. Thus, variation in sequence context and total peak height obtained by bidirectional sequencing of opposite strands of the same PCR product did not result in major variation in the estimate of cytosine methylation status by peak height analysis.

In *Figure 3*, using the sense primer, a progressive increase in the proportion of thymine to cytosine at a representative ACG site in DNA from rats given the methyl deficient diet was quantified. The detection of progressive hypomethylation (increased proportion of thymine) with increased time on the methyl deficient diet underscores the sensitivity of peak height analysis for the estimation of site-specific methylation status. The ability to quantify progressive alterations in site-specific cytosine methylation status with this new methodology should significantly advance the understanding of the pivotal role of DNA methylation during genomic imprinting and embryonic development as well as the pathologic dysregulation of methylation during cancer progression.

References

1 Clark, S.J., Harrison, J., Paul, C.L., and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22,** 2990–2997

2 Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., and Smith, J.A. (1989). Preparation of genomic DNA from mammalian tissue, 2.2.1–2.2.3. In: *Current Protocols in Molecular Biology*. Wiley-Interscience New York, NY USA

3 Parker, L.T., Deng, Q., Zakeri, H., Carlson, C., Nickerson, D.A., and Kwok, P.Y. (1995). Peak height variations in automated sequencing of PCR products using Taq-dye terminator chemistry. *Biotechniques* **19,** 116–121

4 Paul, C.L. and Clark, S.J. (1996). Cytosine methylation: quantitation by automated genomic sequencing and GENESCAN™ analysis. *Biotechniques* **21,** 126–133

5 Pogribny, I.P., Basnakian, A.G., Miller, B.J., Lopatina, A.G., Poirier, L.A., and James, S.J. (1995). Breaks in genomic DNA and within the p53 gene are associated with hypomethylation in livers of folate/methyl deficient rats. *Cancer Res,* **55,** 1894–1901