CpG Island Promoter Region Methylation Patterns of the Inactive-X-Chromosome Hypoxanthine Phosphoribosyltransferase (*Hprt*) Gene

JONG-GWANG PARK AND VERNE M. CHAPMAN*

Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Buffalo, New York 14263

Received 10 June 1994/Returned for modification 11 July 1994/Accepted 6 September 1994

Inactive-X-chromosome genes in mammalian females have methylated CpG islands. We have questioned whether there are variable levels of cytosine methylation at different CpG sites within the island that might indicate the presence of primary sites of methylation which may be critical for the maintenance of gene repression and candidate sites for the initiation of inactivation. To address these questions, we have analyzed the methylation patterns of 32 CpG sites of the X-linked hypoxanthine phosphoribosyltransferase (*Hprt*) gene on the active and inactive X chromosomes of mouse tissues and cell lines, using genomic sequencing of bisulfite-treated genomic DNA. Cytosine is deaminated by bisulfite, but methylcytosine is not affected. Cell lines that were heterozygous for the *Hprt* deletion mutation (*Hprt*^{b-m3}) and a functional *Hprt* allele were selected with 6-thioguanine. The resulting cell populations uniformly carry the intact *Hprt* allele on the inactive X chromosome. The methylation of these CpG sites was determined either by the direct sequence analysis of bisulfite-treated and amplified DNA or by the sequence analysis of clones derived from the amplified DNA. No CpG methylation was detected on the active *Hprt* genes from either males or the active X chromosome of females. On average, 22 CpGs were methylated in the other 50% of female DNA, and the level of methylation at individual sites varied from 42 to 100%. Analysis of the inactive *Hprt* gene in two cell lines showed that averages of 14 and 18 CpGs were methylated and that the frequency of methylation at 32 individual sites ranged from 3 to 100%. The highest frequency of methylation in cell lines coincided with the sequences flanking transcription initiation sites. These results suggest that methylation patterns are heterogeneous within a tissue and even in clonal cell populations and that specific subsets of CpG sites sustain high methylation frequencies which may be critical for the maintenance of X-chromosome inactivation. The bisulfite method identified which CpG sites were methylated on the inactive X chromosome, and it provided a quantitative estimate of the frequency of methylation of these sites in genomic DNA.

Cytosine methylation in mammalian cells has been established as a regulatory feature in genomic imprinting (3, 27), fragile X syndrome (44), and the process of X-chromosome inactivation (16). This methylation is typically found in CpG island-associated promoters that are normally free of methylation when a gene is expressed (5). DNA methylation of X-chromosome genes is correlated with an inactive X chromosome (Xi) in female mammals (16, 33). One of the two X chromosomes in somatic cells of females becomes inactive during early embryonic development, and, once the inactive state of a chromosome is established, it becomes somatically heritable in a cell lineage. The onset of DNA methylation on the Xi may occur after the inactive condition is established, although the timing of this may vary with different sites and genes on the X chromosome.

Treatment of cultured cells with the DNA-demethylating agent 5-azacytidine results in reactivation of some but not all genes from the Xi (32, 34). Furthermore, genomic DNA from the Xi of mouse extraembryonic endoderm can transform hypoxanthine phosphoribosyltransferase (HPRT)-negative cell lines to HPRT positive, whereas comparable DNA from other embryonic lineages cannot (26). These results indicated that methylation is not an essential feature of the Xi and that methylation may be a mechanism for stabilizing inactivation and maintaining the Xi condition rather than a primary factor

in initiating the inactivation. In line with this idea, Lock et al. (28) found that DNA methylation of several restriction sites in *Hprt* did not occur until several days after X-inactivation had occurred in mouse embryos. In addition, methylation of *Pgk-1* and *Hprt* genes was not observed in retinoic acid-induced differentiated embryonal carcinoma cells that have undergone X-inactivation (4).

The results of several studies suggest the presence of site-specific methylation in the regulation of X-linked genes as well as in other genes: (i) the specific pattern of methylation of the Xi includes hypomethylation of some sites and hypermethylation of others (46); (ii) the stability of methylation is different among restriction sites of the human *HPRT* (48) and *PGKI* (18) genes, in which methylation of more stable sites correlates well with gene inactivation activity; (iii) one *HpaII* site (H7) of the mouse *Pgk-1* gene CpG island that is fully methylated on the Xi was shown to be methylated at 5.5 days postcoitum, closer to the time of X-inactivation in female embryos (42); (iv) in embryonic stem (ES) cells, the in vitro-methylated *Appt* transgene was demethylated in the 5' CpG island region, whereas the 3' end of the gene remained modified and was even methylated de novo at additional sites (14); (v) allele-specific methylation at the *Igf2r* intron and at one CpG in *Igf2* is established either in the gametes or shortly after fertilization, while the overall methylation pattern is established after the blastocyst stage during embryogenesis (10); and (vi) in the 5'-flanking region of the vitellogenin II gene, the hormone-specific demethylation of the upper strand is temporally parallel with the rate of mRNA accumulation

* Corresponding author. Mailing address: Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Elm and Carlton Sts., Buffalo, NY 14263. Phone: (716) 845-5840. Fax: (716) 845-8169.

(40). Therefore, it is possible that there are CpG sites that differ from others in the level and/or the timing of their methylation. It is also possible that there are primary CpG sites where methylation is critical for regulating gene expression and that the methylation of these sites on the X chromosome may contribute to the early steps in silencing genes in the inactivation process.

The assessment of cytosine methylation in genomic DNA has been accomplished previously by using Southern analysis with methylation-sensitive endonucleases, PCR amplification of genomic DNA following digestion with methylation-sensitive enzymes (42), and ligation-mediated PCR (37). In general, it has been difficult to establish quantitative determinations of the levels of methylation for each of the CpG sites within CpG islands by these methods. In order to analyze the occurrence of methylation at all CpG sites in the CpG island of mouse *Hprt*, we have adapted the method of PCR amplification of bisulfite-treated genomic DNA (15). Using this approach, we have examined the following features of *Hprt* methylation. First, we have established that there is no methylation present on the active-X-chromosome (Xa) gene of males. Second, we have determined the patterns of CpG methylation in female DNA. Third, we have examined the levels of methylation in cultured cell lines that uniformly have the intact *Hprt* gene on the Xi. Finally, we have examined the stability of *Hprt* methylation in clonal cell populations to establish the fidelity of cytosine methylation through DNA replication and cell division.

MATERIALS AND METHODS

Mice and cell lines. C57BL/6 male and female mice were used for the analysis of DNA methylation in adult tissues. Crosses between the ES cell-derived *Hprt^{b-m3}* animals (BM3) (20, 45) and a C57BL/6-*Hprt^a* *Pgk-1^a* congenic strain (AT29) (13) or *Mus spretus* (*Hprt^a*) (23) were established. Two 6-thioguanine-resistant (TG^r) cell lines, HOBMSL2 (shortened to BML-2) (unpublished data) and HOMSKI (11), were established from primary cultures of (BM3 × AT29)F₁ female liver and (BM3 × *M. spretus*)F₁ female kidney cells, respectively. Since the BM3 mice carry a deletion of the 5' genomic sequence, including the first two exons of *Hprt*, the resulting F₁ female cells have only one intact copy of the gene from AT29 or *M. spretus*. Therefore, TG^r cells carry an Xa of the BM3 origin and an Xi derived from AT29 or *M. spretus*. HSK-1, HSK-2, and HSK-3 are subclones isolated from single cells of the HOMSKI line by limiting dilution in 96-well dishes. All cell lines were routinely maintained by culture in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum.

DNA and plasmids. Mouse genomic DNA was isolated from liver, brain, spleen, and thymus tissues of adult C57BL/6 mice as described by Chapman et al. (12). DNA was isolated from cultured cells by a similar procedure except that instead of being homogenized, the cells were washed with ice-cold phosphate-buffered saline solution and lysed in buffer containing 50 mM Tris (pH 8.3), 50 mM NaCl, 25 mM EDTA, 0.5% sodium dodecyl sulfate, and 300 µg of proteinase K per ml at room temperature. A plasmid, pHPT2.0, containing the promoter, first exon, and 5' part of the first intron of the mouse *Hprt* gene was constructed by subcloning a 1.37-kb fragment (*EcoRI*-*SacI*) from the plasmid pHPT6 (31) and a PCR-amplified 0.6-kb fragment (*SacI*-*HincII*) into pUC19. For *in vitro* methylation, 1 µg of pHPT2.0 was methylated for 4 h by using M·*HpaII* methylase (New England Biolabs) under conditions recommended by the manufacturer. After phenol extraction,

DNA was subjected to *HpaII*-*MspI* digestion to confirm the extent of methylation.

***Xist* expression by reverse transcription (RT)-PCR.** Total RNA was extracted from tissues and cell lines by the RNazol (Biotech) method. After precipitation with isopropanol, samples were treated with RNase-free DNase (Promega) for 1 h at 37°C and reprecipitated after phenol-chloroform extraction. A 1-µg sample of RNA was reverse transcribed with 1 U of avian myeloblastosis virus reverse transcriptase (Promega) and 250 ng of random hexamer primers (Promega) in 20 µl for 20 min at 42°C and then subjected to 5 min of denaturation at 95°C. For each RNA sample, DNA contamination was determined by omitting the reverse transcriptase in parallel. The PCR was carried out with 5 µl of the cDNA in a 50-µl reaction mixture containing 10 pmol of *Xist* primers, 2 mM MgCl₂, and 0.2 mM deoxynucleoside triphosphates for 30 cycles of 95°C (1 min), 60°C (1 min), and 73°C (1 min) in a PTC-100 thermal controller (MJ Research). To determine the origin of the Xi, we digested the RT-PCR products of C57BL/6 and HOMSKI RNAs with *HindIII* and those of BML-2 with *HaeIII* on the basis of allele-specific sequence polymorphisms in the *Xist* cDNA. After digestion, samples were subjected to electrophoresis in a 2.5% agarose gel containing ethidium bromide. Primer pairs used for the study of *Xist* expression were 588-2F (7) and *Xist*1448 (5'-CACCAAATTACTCGGCCACT-3') and MX23b and MIX20 (25). All primers were synthesized by the DNA Synthesis Laboratory at the Roswell Park Cancer Institute.

Cytosine deamination (CD) by bisulfite and subsequent PCR (CD-PCR). Genomic DNA was linearized with *EcoRI*, ethanol precipitated, and quantitated. In order to analyze the methylation pattern by genomic sequencing, we used the sodium bisulfite method of Frommer et al. (15) with a few modifications. Alkaline-denatured DNA (2 to 5 µg) was incubated in 3.1 M (based on SO₂ assay) sodium bisulfite (J. T. Baker Co.), pH 5.0, at a concentration of 8 µg/ml for 40 h at 50°C. The bisulfite-treated DNA was recovered by the GeneClean (Bio 101) method at room temperature. The eluted DNA was alkaline denatured in 0.2 N NaOH, neutralized with ammonium acetate, precipitated in ethanol, resuspended in 10 mM Tris-HCl (pH 8.0), and stored at -20°C until use. To analyze the CpG island, including the promoter and first exon sequence of *Hprt*, a 371-bp fragment from the bisulfite-treated DNA was amplified by PCR. All PCRs were carried out in 50-µl volumes containing 10 mM Tris-HCl (pH 8.3 at 25°C), 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM deoxynucleoside triphosphates, 10 pmol of each primer, and 2 U of *Taq* polymerase (Boehringer Mannheim). The amplification consisted of 4 min of denaturation at 95°C, 25 to 30 cycles of 95°C (1 min), 58°C (1 min), and 73°C (1 min), and 5 min of extension at 73°C. A 5-µl aliquot was analyzed on 2% agarose gels containing ethidium bromide. Primers Hm-aI (5'-GGTCTAGAGGTTG CGGTATGGTTAGTATTA-3') and Hm-aII (5'-ACTCTA GAAAAATCCCCTTAACCTACCAC-3') were used in our study to produce a 371-bp fragment from the upper strand of *Hprt*. These primers were designed to have *XbaI* sites at the 5' ends for cloning of the PCR product, but the *XbaI* site cloning was not used in this study (see below for cloning method). One set of primers specific to the bisulfite-treated lower-strand DNA has failed to amplify a 151-bp fragment, so we present the results for the upper strand.

Cloning and sequencing. CD-PCR products were electrophoresed on 2% SeaPlaque-1% SeaKem agarose gels and purified by the GeneClean (Bio 101) procedure. Gel-purified DNA was either sequenced directly to analyze the average level of methylation in a mixed population of cells or cloned into a pCRII vector of the TA cloning system (Invitrogen) and

FIG. 1. Schematic representation of the mouse *Hprt* gene 5' region. The first exon (positions 1 to 143) (box) and the coding region of *Hprt* (shaded part) are shown. Three Sp1 motifs (arrowheads) are indicated. Primers Hm-aI and Hm-aII are specific to the upper strand and amplify a 371-bp fragment (positions -199 to 171).

then sequenced to determine the methylation pattern of individual molecules. One of the primers used for CD-PCR was end labeled with [32 P]ATP for DNA sequencing. The sequencing reaction was performed according to the fmol DNA sequencing system (Promega) and analyzed on denaturing polyacrylamide gels (6 to 8% acrylamide, 7 M urea). Vacuum-dried gels were exposed to Kodak X-Omat AR film without intensifying screens overnight at room temperature.

RESULTS

CpG methylation of the *Hprt* gene on the Xi. We assayed *Hprt* methylation in two TG⁺ cell lines, BML-2 and HOMSKI, established from (BM3 $\times$ AT29)F₁ and (BM3 $\times$ *M. spretus*)F₁ females, respectively. BM3 carries a deletion of the *Hprt* sequence including the first two exons and the 5' CpG island (20, 45). Somatic tissues of F₁ females (HPRTnull/HPRTA) are mosaic for HPRT expression as a result of random X-inactivation. Therefore, TG⁺ cells are HPRT negative and carry an intact *Hprt*^a gene on the Xi. Northern (RNA) analysis and HPRT activity assays confirmed the lack of *Hprt* expression in these cell lines (data not shown). We also identified the Xi in these cell lines by RT-PCR analysis for the allele-specific expression of the *cis*-acting *Xist* gene. Sequence polymorphisms of the *Xist* transcripts that distinguished C57BL/6-*Hprt*^a (a 302-bp *Hae*III fragment) and *M. spretus* (a 503-bp *Hind*III fragment) (25) alleles from the C57BL/6 allele (including that of BM3) were used in this study. We found that *Xist* expression in BML-2 and in HOMSKI cells was exclusively from the AT29 and the *M. spretus* alleles, respectively (data not shown). These results suggested that the *cis*-associated intact *Hprt* genes in these cells were present on the Xi. We also characterized three clonal cells derived from the HOMSKI line for allele-specific *Xist* expression by RT-PCR as described above. *Xist* expression from all three clones was exclusively from the *M. spretus* allele, as it was in the original HOMSKI cells (data not shown), suggesting that these clonal cells maintain an Xi that carries the *cis*-intact *Hprt*^a allele.

Analysis of methylated CpG sites in genomic DNA. We used a genomic sequencing method that is based on bisulfite modification of single-stranded DNA followed by PCR amplification for the analysis of DNA methylation patterns (15). Bisulfite treatment converts preferentially unmethylated cytosine (C) residues to uracil (U) so that C residues are amplified as thymine (T) and only methylated cytosine (m⁵C) residues appear as C in the PCR product. CD-PCR provides a means of identifying and quantitating DNA methylation in genomic DNA. We designed one set of PCR primers (Hm-aI and Hm-aII) that is specific for only the upper strand of bisulfite-reacted DNA. Figure 1 shows schematically the mouse *Hprt* 5' CpG island, including the sequence region we analyzed in this study. As a control for the specific conversion of only C to U, we performed CD-PCR sequencing of both the

FIG. 2. Control experiment showing preferential conversion of cytosine to thymine relative to m⁵C under CD-PCR conditions. Sequences of the two *Hpa*II sites at -18 and -70, relative to the major transcription start site, are illustrated. The sequence of unmethylated plasmid (Original DNA) is shown. Bisulfite-treated nonmethylated plasmid [*M.Hpa*II(-)] and plasmid that had been methylated in vitro with *Hpa*II methylase [*M.Hpa*II(+)] were PCR amplified and sequenced with primer Hm-aI. C residues of the CCGG *Hpa*II recognition sequences (arrows) are indicated. CC sequences of the control DNA have been converted to TT and TC in *M.Hpa*II(-) and *M.Hpa*II(+), respectively, because of bisulfite treatment and PCR amplification.

plasmid DNA containing the *Hprt* CpG island and the in vitro-methylated plasmid within the CCGG *Hpa*II sequences. Figure 2 shows representative sequences flanking two *Hpa*II sites, obtained by direct sequencing of CD-PCR product for control reactions. Under our conditions (see Materials and Methods for details), all Cs other than m⁵Cs of *Hpa*II sites were uniformly converted to uracil and read as T in the sequencing gel (Fig. 2).

Methylation of the CpG island of *Hprt* in adult tissues. CD-PCR products of *Hprt* genes from male and female liver DNAs and from plasmid DNA of the *Hprt* genomic clone are shown in parallel with the original sequence for a 139-bp segment of the *Hprt* region analyzed (Fig. 3). The position of each m⁵C residue in 19 CpG sequences in the original DNA was identified by the presence of a band in the C lane in the sequencing reaction with the primer Hm-aI. Plasmid and male DNAs have all C residues converted to T with the exception of the second C of CpCpTpG (*Eco*RII site) in the plasmid DNA, probably because of *dcm* methylation in *Escherichia coli*, which indicates a lack of methylation in the active *Hprt* upstream sequences. However, a mixture of C and T was seen at each of the CpG sites in female *Hprt* sequences. Similar patterns of C and T were obtained throughout the entire 371-bp region containing 32 CpG sites. The identification of C and T in the sequence of female *Hprt* CD-PCR product was consistent with a mixture of nonreactive m⁵C and reactive C that were from the *Hprt* genes of Xi and Xa, respectively. We also examined the methylation patterns of the *Hprt* promoter region from female DNA extracted from brain, liver, spleen, and thymus tissues to determine whether there were tissue-specific differences in the extent of methylation. Sequence analysis using the primer Hm-aII showed a G band when opposite C on the complementary strand that was an m⁵C on the original upper strand. A mixture of A and G was seen at each CpG site, and the proportions of A and G did not differ between tissues (data not shown). This result is consistent with the housekeeping function of *Hprt* and suggests that the methylation of CpG sites was maintained in stable condition. These results suggest that this *Hprt* sequence region was not methylated on the Xa but that it was hypermethylated on the Xi.

FIG. 3. Direct sequencing of CD-PCR-amplified DNAs from plasmids, male and female mice, and a cell line, BML-2. DNA was treated with sodium bisulfite, amplified, and sequenced with the primer Hm-aI reading the upper strand of the original DNA from the 5'-to-3' direction. Thus, unmethylated C is changed to T while m^5C remains as C. Male DNAs have all C residues converted to T, but a mixture of C and T is seen at each of these CpG sites in the female and BML-2 *Hprt* sequences. Arrowheads and dashed lines, all C of CpG sites present in the original DNA before bisulfite reaction and C or T of the corresponding sites present after CD-PCR; open arrows, positions of Cs of CCTG sequences that were not converted to T in plasmid DNA; small arrows (p), polymorphic sequence differences between B/6 and BML-2; large arrow, position of +1, major *Hprt* transcription start site; B/6, C57BL/6.

Sequence analysis of cloned CD-PCR products was used to test whether the pattern of C/T or A/G observed in females represented a summing of methylated and nonmethylated sequences. We initially determined that there was a complete lack of nonreactive C at all 32 CpG sites in a total of 16 clones from CD-PCR of male genomic DNA (data not shown). The detailed methylation status of specific CpG sites was established by sequence analysis of 46 randomly collected clones derived from CD-PCR of female liver DNA. Typical sequencing gels from the analysis of female-derived clones are shown in Fig. 4. Clone 1 (FC1) has no C, while three other clones have Cs and Ts at different CpG sites. This result suggested (i) that there were unmethylated and methylated *Hprt* genes in female DNA presumed to be from the Xa and Xi, respectively, on the basis of the facts that there was no methylation in male DNA (data not shown) and that all of the inactive *Hprt* sequences were methylated in cell lines (see the next section), and (ii) that

FIG. 4. Sequences analysis of clones derived from female DNA after CD-PCR. Sequencing was performed as described for Fig. 3 with primer Hm-aI. All of the CpG sites in this region (arrows) are indicated. Clone 1 (FC1) has no C and is presumed to be from the Xa *Hprt* gene, while the three other clones have Cs at different sites and are presumed to be from the Xi *Hprt* gene.

there was heterogeneity of methylation at CpG sites. Two different groups of methylation patterns were observed in 46 clones from female DNA (Fig. 5). In one group of 22 clones, all of the CpG sequences were hypomethylated (*Hprt* on the Xa), while in the second group of 24 clones all of the 32 CpG sites were hypermethylated (*Hprt* on the Xi) (Fig. 5). In the 22 hypomethylated clones, 20 clones had all of the Cs converted to Ts by the bisulfite treatment, indicating that there was no methylation at any site (Xa). However, two clones showed either one or two nonreactive sites. The retention of C in these clones could have been due to a failure to deaminate by bisulfite or due to mistakes of *Taq* polymerase (Xa). In the second group of 24 clones from the Xi *Hprt* gene, all 32 CpG sites showed some amount of methylation but not all of the sites were methylated in any given cell. The number of CpG sites ranged from 16 to 28 of 32 sites, with an average ($\pm$ standard deviation) of 22.46 ± 0.66 sites ($\sim 70\%$) per individual DNA (Fig. 5). The levels of methylation at individual CpG sites were determined from the frequency of clones having C at

FIG. 5. Proportion of methylated sites in clones of *Hprt* DNA following CD-PCR. The methylation of CD-PCR product clones at 32 sites (33 in HOMSKI) was determined. Two different groups were derived from female DNA, 22 clones which are hypomethylated (Xi *Hprt*) and 24 clones which are hypermethylated (Xa *Hprt*) at 32 CpG sites. By contrast, the proportion of methylated CpG sites was lower in the two TG⁺ cell lines, BML-2 (34 clones) and HOMSKI (33 clones), that have the intact *Hprt* on the Xi. Methylated sites in the Xi (means $\pm$ standard errors): 22.46 ± 0.66 in female DNA, 14.06 ± 0.63 in BML-2, and 18.15 ± 0.52 in HOMSKI. B/6, C57BL/6.

FIG. 6. Frequency of DNA methylation at individual CpG sites in the region from -114 to +156 of *Hprt*. Data are from 24 clones of female liver DNA, 34 clones of BML-2 DNA, and 33 clones of HOMSKI DNA analyzed in Fig. 5. Most of these sites show a high frequency of methylation in female clones, but TG-selected cell lines maintain a high level of methylation in subsets of CpG sites. One of these subsets (or domains indicated by brackets) spans the transcription start (+1) region of the *Hprt*. Potential Sp1 motifs are underlined.

each site in the sequence analysis. The levels for the second group varied from 42 to 100% for 32 sites across the *Hprt* promoter region (Fig. 6).

DNA methylation in cell lines. We assayed *Hprt* methylation in two TG^r cell lines, BML-2 and HOMSKI, that carried the *Hprt*^{b-m3} deletion mutation on the Xa and an intact *Hprt* allele on the Xi. Sequence analysis of *Hprt* CD-PCR products from BML-2 DNA is shown in Fig. 3. A patchy pattern of methylation was detected in *Hprt* sequences of BML-2, and extensive methylation was evident in the sequences near the major transcription start site (Fig. 3). A very similar pattern was seen in HOMSKI cells (data not shown). The methylation status at 32 CpGs (33 in HOMSKI cells) was analyzed in cloned CD-PCR products to determine the methylation of individual amplified sequences. All of the CpG sites were methylated, but the number of sites methylated in each clone was less than that observed in the 24 hypermethylated clones from female tissue DNA (Fig. 5). In addition, the distribution of methylation was not uniform across the total 32 CpG sites (Fig. 6). There were 14.06 ± 0.63 methylated sites in 34 BML-2 clones and 18.15 ± 0.52 methylated sites in 33 HOMSKI clones, which was significantly lower than the number of methylated sites observed in DNA from female tissues (22.46 ± 0.66) ($P < 0.01$). The frequency of clones showing methylation at individual CpG sites ranged from 12 to 100% in BML-2 and 3 to 97% in HOMSKI cells (Fig. 6). Three domains (a single CpG or a group of CpGs) that maintained a high frequency ($\geq 70\%$) of methylation were observed. These domains were ~ 100 bp away from each other, with the middle domain encompassing the major transcription start site. The first and the last domains

correspond to the regions of the two highest CpG densities within the sequence region we analyzed. These findings suggest that the overall pattern of CpG methylation may change markedly in the cell lines studied but that high levels of methylation were maintained at some sites. It is possible that these methylated CpGs were important for *Hprt* transcriptional regulation.

Analysis of methylation in clonal cell populations. Complex methylation patterns were observed in tissues and cultured cell lines. Cell-specific patterns of CpG methylation could be the result of differences in the original sites of methylation in the progenitor cells that are rigorously maintained by methyltransferases that recognize hemimethylated sites during cell replication. Alternatively, these differences could reflect a more dynamic process of methylation maintenance. To test this, we have studied methylation within the *Hprt* gene in three clonal cell populations (HSK-1, HSK-2, and HSK-3) isolated by single-cell cloning of the TG^r HOMSKI line. The RT-PCR analysis indicated that these clonal cells maintained the Xi, on which *M. spretus* *Xist* was expressed (data not shown), and were presumed to have *cis* intact *Hprt* on the Xi.

For the bisulfite reaction, DNA from single-cell-derived sublines passaged six times for approximately 20 generations without any selection was analyzed. Genomic sequencing was carried out by CD-PCR as described above, and representative sequence ladders from three clonal populations are shown in Fig. 7A. The presence of G-plus-A bands at most of the CpG sites demonstrated that cells were heterogeneously methylated at those sites despite their clonal origin. Quantitation of methylation by direct sequencing of the CD-PCR products is

FIG. 7. Analysis of methylation in HOMSKI clonal populations of cells. (A) DNA sequencing gel showing the analysis of the CD-PCR products derived from HOMSKI cells and three populations of cells cloned from HOMSKI. Primer Hm-aII, which reads the complement of the bisulfite-reacted coding strand of *Hprt*, was used for these analyses. This gel shows the methylation status of the CpG sites (25 to 30) (arrows between sequence positions 79 and 119 of the *Hprt* gene). A mixture of A and G after a C band indicates that there are heterogeneous methylation patterns in the clonal populations as well as in the parental HOMSKI cells. (B) Summary of CpG methylation at eight sites in HOMSKI cells and three clonal populations. The data were compiled from the individual clonal sequence analysis of HOMSKI CD-PCR products and the densitometry of direct sequencing shown in panel A.

difficult because of the lack of a reliable internal standard and the variable signal-to-noise ratio in a sequence lane. However, we attempted a semiquantitative estimation of methylation for eight CpG sites, using scanning densitometry of the direct sequencing autoradiogram shown in Fig. 7A, which is summarized in Fig. 7B. In this estimation, we assumed that the total of methylation levels at eight sites was equal in four cell lines and we utilized the quantitative HOMSKI data (Fig. 6) obtained from the sequence analysis of individual cloned DNAs as a reference. The general patterns of CpG methylation for the three clonal populations of cells are similar, and they parallel the cumulative pattern of methylation observed in the original HOMSKI cells. However, there are some sites, such as site 26 in HOMSKI-1 and site 31 in HOMSKI-3, that have increased methylation relative to that of the HOMSKI parent. Methylation at sites 29 and 30 is more extensive and invariant

in HOMSKI and its sublines. These results suggest that methylation at sites 29 and 30 was uniformly present in the clonal cells and that it was maintained through 20 generations, which is consistent with a methylation domain around CpG sites 29 and 30. Thus, specific sites of methylation were not strictly maintained in a lineage of cells growing in culture, but the levels of methylation were faithfully retained at some sites that may be important for the maintenance function of gene repression.

DISCUSSION

The methylation status of 32 CpG sites in the CpG-rich *Hprt* promoter region was analyzed by using the preferential deamination of cytosine residues by bisulfite in single-stranded DNA followed by PCR amplification of the bisulfite-treated DNA (CD-PCR). Nonreactive m^5C residues are identified in sequencing gels as cytosines (C), whereas unmethylated C residues are deaminated to uracil and amplified as thymine by PCR (15). The CD-PCR method is highly efficient and reliable. We observed nonreactive-DNA levels that were proportional to the input ratio of the methylated plasmid DNA in experimental mixtures (data not shown). Any incomplete bisulfite reaction was detected by identifying C at any non-CpG site in the DNA. Quantitative levels of methylation for specific CpG sites were established with this method through cloning and sequencing of the CD-PCR products from genomic DNA. We found, for example, that the level of methylation at five *HpaII* sites in the *Hprt* promoter region varied and that only 17% of the Xi in female DNA was methylated at all five sites. Comparable analyses of DNA methylation at these sites by Southern blotting of *HpaII*-digested genomic DNA would not be feasible.

The extent of methylation on the inactive Xi *Hprt* differed between somatic tissues and cell lines grown in culture. DNAs from cultured cells were methylated at 63 to 81% of the CpG sites that were methylated in female tissue (Fig. 5). Furthermore, a decreased frequency of methylation for individual CpG sites was detected in cell lines, despite the fact that the overall pattern of methylation of the Xi CpG was retained (three domains preserved their methylation level of $\geq 70\%$ in Fig. 6). Our results for *Hprt* CpG methylation differ from previous studies of the promoter region of the X-linked human phosphoglycerate kinase 1 (*PGKI*) gene, in which the inactive *PGKI* gene was fully methylated at 60 of 61 CpGs in cultured cells (36, 38). Nevertheless, mosaic patterns of CpG methylation and extensive methylation around the transcription start point were seen in the cells that were cloned from hamster-human hybrid cells after 5-azacytidine treatment (38). Heterogeneity of methylation was also found by using methylation-sensitive enzymes (41, 47, 48). The use of sodium bisulfite modification more accurately identifies the status of CpG methylation, and it may address the function of methylation in regulating gene expression.

Very little is known about the basis for de novo methylation of CpG islands of Xi genes. However, it is likely that the hemimethylated substrate of newly synthesized DNA and the processive functioning of the enzyme work in concert to maintain these levels (17). Where enzyme and substrate are abundant, it is possible that there is a gradual increase in the number of CpG sites that become methylated in an Xi gene. By contrast, if either enzyme or substrate is limited, it is possible that the frequencies of CpG methylation within an island will gradually decline in cells. Under these circumstances, the sites that remain methylated may reflect either intrinsic differences in the probability of methyltransferases binding or the highest

density of CpG methylation. Alternatively, it is also possible that the binding of secondary proteins could help to target these enzymes to DNA sequences or to chromatin regions (8). The observed mosaic patterns of methylation in HOMSKI clonal cell populations differ from each other, suggesting clonal inheritance, and may reflect de novo methylation or demethylation or both during subsequent cell divisions. Pfeifer et al. (36) estimated that the efficiency of maintenance methylation in the human-hamster hybrid system is ~99.9% for all sites of the human *PGK1* 5' region. Our results for the methylation at eight CpG sites of *Hprt* are rough estimates, but partial methylation in clonal populations of cells is clearly inconsistent with such a high efficiency. This result establishes that the original pattern is generally retained and that methylation at some sites is more strictly maintained within a CpG island.

Sites or domains that are fully methylated through successive cell generations may be involved in the mechanisms for *Hprt* repression. The *Hprt* promoter lacks TATA and CAAT consensus sequences, like most CpG island-associated genes, but it does have two copies of potential Sp1 binding sites (31). The frequencies of methylation at these two Sp1 sites were 54 to 63% in female DNA and decreased three- to ninefold in cell lines in this study. Indeed, Sp1 can bind and function equally well at methylated and unmethylated sites, regardless of methylation (19). Thus, methylation of CpG sites, especially those in the promoter region, could indirectly inhibit transcription. Genomic sequencing and footprinting analyses of the CpG island of the human *PGK1* gene have shown that the Xi did not contain any transcription factors but that it was wrapped around two positioned nucleosomes or nucleosome-like particles (35). Recent studies have shown that the methyl-CpG-binding protein (MeCP-1) (30) is an indirect inhibitor of transcription from many methylated promoters, including the *PGK1* promoter (8, 9). However, previous in vivo footprinting analysis of the 5' region of either *Hprt* (21) or *Pgk-1* (35, 38) genes could not detect any specific DNA-binding protein on the inactive alleles. The mechanisms of how these proteins are involved in gene silencing remain to be resolved (6, 39).

Recently, Ackerman et al. (1) have reported that an imperfect dyad symmetry is necessary and sufficient for basal promoter function of the murine adenosine deaminase (*Ada*) gene. This model postulates the formation of stem-loop structures around the major transcription start sites in CpG-rich promoters. The promoter of *Hprt* has TATA-less and CpG-rich sequence structures that are similar to those of *Ada*, which suggests that there could be similar mechanisms for the expression of *Hprt*. Figure 8a shows a secondary structure deduced from the *Hprt* sequence region that is coincident with the second domain of methylation (Fig. 6) flanking the major transcription initiation site (31). There are three reasons for our belief that this kind of structural motif is a candidate element for promoter function of *Hprt*. First, both *Hprt* and *Pgk-1* genes have no well-known sequence motif in the promoter region, with the exception of potential Sp1 binding sites. Second, we can deduce the potential for stem-loop structures from the primary sequence of both mouse and human *Hprt* and *Pgk-1* genes despite their poor overall sequence homology, as shown in Fig. 8 (the H-7 site of the mouse *Pgk-1* is downstream of the structure). Third, this study and other studies of the human *PGK1* (18, 36) and human *HPRT* (22) genes have shown that there is a consistent correlation between gene activity and the methylation status of the region around the major transcription start sites and downstream sequences. It is noteworthy that the predicted structures of these genes include the so-called HIP 1 binding site, which is known to be impor-

FIG. 8. Potential secondary structures deduced from the sequences of the mouse *Hprt* (31) (a), human *HPRT* (18) (b), mouse *Pgk-1* (2) (c), and human *PGK1* (36) gene transcription regions. The stem-loop structures were predicted by the Fold program (version 7; Genetics Computer Group, Inc., Madison, Wis.) (49) on a VAX computer, plotted with the Genetics Computer Group program Squiggles and drawn with the Canvas program (version 3; Deneba Systems, Inc.). Major transcription sites within the structures (arrows), sequences homologous to the HIP 1 binding site (29) (asterisks), and cytosine residues of CpG sites (boldface) are indicated.

tant in the transcription of the dihydrofolate reductase gene (29). Thus, DNA methylation of the conserved region in the inactive *Hprt* and *Pgk-1* genes may disturb the formation of open structures either directly or indirectly via methyl-binding proteins. Under these circumstances, a small number of m⁵CpG sites may be sufficient to maintain the inactive state of these genes. In the same line, Kass et al. (24) showed that regional methylation of CpG-rich sequences could drive plasmid minichromosomes to form heterochromatin. Whether the predicted structures occur in vivo remains to be determined.

The results of our study demonstrate that there are characteristic differences in the levels of methylation at different CpG sites within the *Hprt* promoter region, particularly in the TG^r cell lines. These results suggest that the potential role of DNA methylation in the initiation of X-inactivation should be reassessed by determining when these primary sites become methylated during development. In light of the recently described patterns of methylation of imprinted genes during embryogenesis (43), it is possible that de novo methylation of genes does play a major role in establishing inactivation for individual genes. Whether de novo methylation of all genes on the Xi occurs uniformly or whether there are differential sites of methylation on chromosomes remains to be addressed.

ACKNOWLEDGMENTS

We thank William A. Held for critical reading of the manuscript, Kwang-Ho Lee for construction of BML-2 and HOMSKI cell lines, and Andrew R. Zinn at Whitehead Institute for providing *Xist* primers. The valuable assistance of Nancy Holdsworth and Greg Kohler in preparing the manuscript is gratefully acknowledged. We also thank D. Swiatek and R. Benz for assistance in the care of the mice.

This work was supported in part by grants GM33160 and HG00277 from the National Institutes of Health to V.M.C.

REFERENCES

- Ackerman, S. L., A. G. Minden, and C.-Y. Yeung. 1993. The minimal self-sufficient element in a murine G+C-rich promoter is a large element with imperfect dyad symmetry. *Proc. Natl. Acad. Sci. USA* **90**:11865–11869.
- Adra, C. N., P. H. Boer, and M. W. McBurney. 1987. Cloning and expression of the mouse pgk-1 gene and the nucleotide sequence of its promoter. *Gene* **60**:65–74.
- Barlow, D. P. 1993. Methylation and imprinting: from host defense to gene regulation? *Science* **260**:309–310.
- Bartlett, M. H., C. N. Adra, J. Park, V. M. Chapman, and M. W. McBurney. 1991. DNA methylation of two X chromosome genes in female somatic and embryonal carcinoma cells. *Somatic Cell Mol. Genet.* **17**:35–47.
- Bird, A. P. 1986. CpG-rich islands and the function of DNA methylation. *Nature (London)* **321**:209–213.
- Bird, A. P. 1992. The essentials of DNA methylation. *Cell* **70**:5–9.
- Borsani, G., R. Tonlorenzi, M. C. Simmler, L. Dandolo, D. Arnaud, V. Capra, M. Grompe, A. Pizzuti, D. Muzny, C. Lawrence, H. F. Willard, P. Avner, and A. Ballabio. 1991. Characterization of a murine gene expressed from the inactive X chromosome. *Nature (London)* **351**:325–329.
- Boyes, J., and A. P. Bird. 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64**:1123–1134.
- Boyes, J., and A. P. Bird. 1992. Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. *EMBO J.* **11**:327–333.
- Brandeis, M., T. Kafri, M. Ariel, J. R. Chaillet, J. MacCarrey, A. Razin, and H. Cedar. 1993. The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. *EMBO J.* **12**:3669–3677.
- Bressler, S. L., K.-H. Lee, D. A. Adler, V. M. Chapman, and C. M. Disteché. 1993. Maintenance of X inactivation of the *Rps4*, *Zfx*, and *Ube1* genes in a mouse in vitro system. *Somatic Cell. Mol. Genet.* **19**:29–37.
- Chapman, V. M., P. G. Kratzer, L. D. Siracusa, B. A. Quarantillo, R. Evans, and R. M. Liskay. 1982. Evidence for DNA modification in the maintenance of X-chromosome inactivation of adult mouse tissues. *Proc. Natl. Acad. Sci. USA* **79**:5357–5361.
- Chapman, V. M., D. R. Miller, D. Armstrong, and C. T. Caskey. 1989. Recovery of induced mutations for X chromosome-linked muscular dystrophy in mice. *Proc. Natl. Acad. Sci. USA* **86**:1292–1296.
- Frank, D., I. Keshet, M. Shani, A. Levine, A. Razin, and H. Cedar. 1991. Demethylation of CpG islands in embryonic cells. *Nature (London)* **351**:239–241.
- Frommer, M., L. E. McDonald, D. S. Miller, C. M. Collis, F. Watt, G. W. Grigg, P. L. Molloy, and C. L. Paul. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA* **89**:1827–1831.
- Grant, S. G., and V. M. Chapman. 1988. Mechanisms of X-chromosome regulation. *Annu. Rev. Genet.* **22**:199–233.
- Gruenbaum, Y., H. Cedar, and A. Razin. 1982. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature (London)* **295**:620–622.
- Hansen, R. S., N. A. Ellis, and S. M. Gartler. 1988. Demethylation of specific sites in the 5' region of the inactive X-linked human phosphoglycerate kinase gene correlates with the appearance of nuclease sensitivity and gene expression. *Mol. Cell. Biol.* **8**:4692–4699.
- Holler, M., G. Weistin, J. Jiricny, and W. Schaffner. 1988. Sp1 transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. *Genes Dev.* **2**:1127–1135.
- Hooper, M., K. Hardy, A. Handyside, S. Hunter, and M. Monk. 1987. HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature (London)* **326**:292–295.
- Hornstra, I. K., and T. P. Yang. 1992. Multiple in vivo footprints are specific to the active allele of the X-linked human hypoxanthine phosphoribosyltransferase gene 5' region: implication for X-chromosome inactivation. *Mol. Cell. Biol.* **12**:5345–5354.
- Hornstra, I. K., and T. P. Yang. 1994. High-resolution methylation analysis of the human hypoxanthine phosphoribosyltransferase gene 5' region on the active and inactive X chromosomes: correlation with binding sites for transcription factors. *Mol. Cell. Biol.* **14**:1419–1430.
- Johnson, G. G., T. A. Larsen, P. Blakely, and V. M. Chapman. 1985. Elevated levels of erythrocyte hypoxanthine phosphoribosyltransferase associated with allelic variation of murine *Hprt*. *Biochemistry* **24**:5083–5089.
- Kass, S. U., J. P. Goddard, and R. L. P. Adams. 1993. Inactive chromatin spreads from a focus of methylation. *Mol. Cell. Biol.* **13**:7372–7379.
- Kay, G. F., G. D. Penny, D. Patel, A. Ashworth, N. Brockdorff, and S. Rastan. 1993. Expression of *Xist* during mouse development suggests a role in the initiation of X chromosome inactivation. *Cell* **72**:171–182.
- Kratzer, P. G., V. M. Chapman, H. Lambert, R. Evans, and R. M. Liskay. 1983. Differences in the DNA of the inactive X-chromosomes of fetal and extraembryonic tissues of mice. *Cell* **33**:37–42.
- Li, E., C. Beard, and R. Jaenisch. 1993. Role of DNA methylation in genomic imprinting. *Nature (London)* **366**:362–365.
- Lock, L. F., N. Takagi, and G. R. Martin. 1987. Methylation of the *Hprt* gene on the inactive X occurs after chromosome inactivation. *Cell* **48**:39–46.
- Means, A. L., and P. J. Farnham. 1990. Transcription initiation from the dihydrofolate reductase promoter is positioned by HIP binding at the initiation site. *Mol. Cell. Biol.* **10**:653–661.
- Meehan, R. R., J. D. Lewis, S. McKay, E. L. Kleiner, and A. P. Bird. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* **58**:499–507.
- Melton, D. W., D. S. Konecki, J. Brennand, and C. T. Caskey. 1984. Structure, expression, and mutation of the hypoxanthine phosphoribosyltransferase gene. *Proc. Natl. Acad. Sci. USA* **81**:2147–2151.
- Mohandas, T., R. S. Sparkes, and L. J. Shapiro. 1981. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* **211**:393–396.
- Monk, M., and M. Grant. 1990. Preferential X-chromosome inactivation, DNA methylation and imprinting, p. 55–62. *In* M. Monk and A. Surani (ed.), *Genomic imprinting (development 1990 supplement)*. The Company of Biologists Ltd., Cambridge.
- Paterno, G. D., C. N. Adra, and M. W. McBurney. 1985. X-chromosome reactivation in mouse embryonal carcinoma cells. *Mol. Cell. Biol.* **5**:2705–2712.

35. **Pfeifer, G. P., and A. D. Riggs.** 1991. Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR. *Genes Dev.* **5**:1102–1113.
36. **Pfeifer, G. P., S. D. Steigerwald, R. S. Hansen, S. M. Gartler, and A. D. Riggs.** 1990. Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. *Proc. Natl. Acad. Sci. USA* **87**:8252–8256.
37. **Pfeifer, G. P., S. D. Steigerwald, P. R. Mueller, B. Wold, and A. D. Riggs.** 1989. Genomic sequencing and methylation analysis by ligation mediated PCR. *Science* **246**:810–813.
38. **Pfeifer, G. P., R. L. Tanguay, S. D. Steigerwald, and A. D. Riggs.** 1990. In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev.* **4**:1277–1287.
39. **Riggs, A. D., and G. P. Pfeifer.** 1992. X-chromosome inactivation and cell memory. *Trends Genet.* **8**:169–174.
40. **Saluz, H. P., J. Jiricny, and J. P. Jost.** 1986. Genomic sequencing reveals a positive correlation between the kinetics of strand-specific DNA demethylation of the overlapping estradiol/glucocorticoid-receptor binding sites and the rate of avian vitellogenin mRNA synthesis. *Proc. Natl. Acad. Sci. USA* **83**:7167–7171.
41. **Silva, A. J., K. Ward, and R. White.** 1993. Mosaic methylation in clonal tissue. *Dev. Biol.* **156**:391–398.
42. **Singer-Sam, J., M. Grant, J. M. LeBon, K. Okuyama, V. M. Chapman, M. Monk, and A. D. Riggs.** 1990. Use of an *HpaII*-polymerase chain reaction assay to study DNA methylation in the P_{gk}-1 CpG island of mouse embryos at the time of X-chromosome inactivation. *Mol. Cell. Biol.* **10**:4987–4989.
43. **Stoeger, R., P. Kubicka, C.-G. Liu, T. Kafri, A. Razin, H. Cedar, and D. P. Barlow.** 1993. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell* **73**:61–71.
44. **Sutcliffe, J. S., D. L. Nelson, F. Zhang, M. Pierretti, S. T. Caskey, D. Saxe, and S. T. Warren.** 1992. DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum. Mol. Genet.* **1**:397–400.
45. **Thompson, S., A. R. Clarke, A. M. Pow, M. L. Hooper, and D. W. Melton.** 1989. Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells. *Cell* **56**:313–321.
46. **Viegas-Pequignot, E., B. Dutrillaux, and G. Thomas.** 1988. Inactive X chromosome has the highest concentration of unmethylated HhaI sites. *Proc. Natl. Acad. Sci. USA* **85**:7657–7660.
47. **Wolf, S. F., and B. R. Migeon.** 1982. Studies of X chromosome DNA methylation in normal human cells. *Nature (London)* **295**:667–671.
48. **Yen, P. H., T. Mohandas, and L. J. Shapiro.** 1986. Stability of DNA methylation of the human hypoxanthine phosphoribosyl-transferase gene. *Somatic Cell. Mol. Genet.* **12**:153–161.
49. **Zuker, M., and P. Stiegler.** 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. *Nucleic Acids Res.* **9**:133–148.