Sp1 sites in the mouse \textit{aprt} gene promoter are required to prevent methylation of the CpG island

Donald Macleod, 1 Jillian Charlton, 1 John Mullins, 2 and Adrian P. Bird 1

1Institute of Cell and Molecular Biology, University of Edinburgh, Darwin Building, King’s Buildings, Edinburgh EH9 3JR, Scotland; 2Centre for Genome Research, Roger Land Building, King’s Buildings, Edinburgh EH9 3JR, Scotland

In an attempt to find the mechanism by which CpG islands remain free of methylation we have undertaken a detailed examination of the mouse adenine phosphoribosyltransferase (aprt) gene. This housekeeping gene has a CpG island that extends over the gene promoter and includes the first two exons. We show that the island is free of methylation at all CpGs, whereas the flanks are methyated. Detailed patterns of methylation beyond the boundaries of the CpG island vary between cells. In vivo footprinting across the island region shows that three GC boxes clustered at the 5' edge of the CpG island are occupied, most probably by Sp1. No other footprints are detected within the island region. Deletion or mutagenesis of the Sp1 sites causes de novo methylation of the CpG island in a transgenic mouse assay. Thus, the peripherally located Sp1 sites are necessary to keep the \textit{aprt} island methylation free.

[Key Words: CpG islands; aprt; DNA methylation; in vivo footprinting; chromatin; LMPCR]

Received July 11, 1994; revised version accepted August 10, 1994.
have identified regions that are bound to proteins in vivo. To test the significance of bound factors we have deleted and mutated the binding sites. Both treatments change nonmethylated cytosines into uracils while leaving methylated cytosines unchanged. Subsequent PCR amplification and sequencing reveals the methylated bases. Unlike the genomic sequencing method, which examines the combined pattern of methylation of a large number of molecules simultaneously, this method allows the sequence of individual molecules to be resolved. A large region of the aprt gene (1936 bp), including the whole CpG island, was analyzed in this way using liver DNA as starting material. These results show the position of methylated CpG sites in four independent clones, each corresponding to the sequence of the gene within an individual liver cell (Fig. 1B). In each case it can be seen that the central region, which also shows the highest frequency of CpGs, is completely nonmethylated. The analysis also shows which sites in the flanking DNA are methylated and allows the boundaries of the CpG island to be determined in terms of both CpG frequency and methylation. The methylation-free region extends from position 610 to 1425, and this coincides with the region of elevated CpG. In the flanking region it is notable that individual sites are methylated in some molecules and nonmethylated in others. Thus, although the overall pattern of methylation is inherited (Holliday and Pugh 1975; Riggs 1975) methylation of individual sites is not clonally derived.

Results

DNA methylation analysis of the aprt CpG island

To accurately position the CpG island relative to the structure of the mouse aprt gene, the distribution of the CpG dinucleotides was plotted using the published DNA sequence (Fig. 1A; Dush et al. 1985). The greatest concentration of CpGs extends from the promoter at the left edge over the first two exons into the second intron of the gene. To analyze the boundaries of the island with respect to methylation, the bisulfite-modification technique was used (Frommer et al. 1992). This procedure changes nonmethylated cytosines into uracils while leaving methylated cytosines unchanged. Subsequent PCR amplification and sequencing reveals the methylated bases. Unlike the genomic sequencing method, which examines the combined pattern of methylation of a large number of molecules simultaneously, this method allows the sequence of individual molecules to be resolved. A large region of the aprt gene (1936 bp), including the whole CpG island, was analyzed in this way using liver DNA as starting material. These results show the position of methylated CpG sites in four independent clones, each corresponding to the sequence of the gene within an individual liver cell (Fig. 1B). In each case it can be seen that the central region, which also shows the highest frequency of CpGs, is completely nonmethylated. The analysis also shows which sites in the flanking DNA are methylated and allows the boundaries of the CpG island to be determined in terms of both CpG frequency and methylation. The methylation-free region extends from position 610 to 1425, and this coincides with the region of elevated CpG. In the flanking region it is notable that individual sites are methylated in some molecules and nonmethylated in others. Thus, although the overall pattern of methylation is inherited (Holliday and Pugh 1975; Riggs 1975) methylation of individual sites is not clonally derived.

In vivo footprinting across the aprt CpG island

Having established the detailed methylation pattern, we next asked whether protein factors were bound within

---

**Figure 1.** (A) A CpG plot of the mouse aprt gene. Each vertical line indicates the position of each CpG dinucleotide in the DNA sequence (Dush et al. 1985). The positions of exons are indicated by open rectangles and transcription start sites by open arrowheads (Dush et al. 1988). The scale (top) is in base pairs. (B) An enlarged plot of a section of the above CpG plot, showing the part of the APRT sequence used in the methylation analysis. The series of vertical lines indicates the position of each CpG dinucleotide, and underneath this, the open circles indicate whether the cytosine in the dinucleotide is nonmethylated (○) or methylated (●). Each horizontal line of circles represents the methylation analysis from one of four independent clones.
the CpG island. To detect protein–DNA interactions in vivo we used the ligation-mediated polymerase chain reaction (LMPCR) technique [Mueller and Wold 1989] to study the aprt promoter in F9 embryonal carcinoma (EC) cells. We designed pairs of DNA primers for LMPCR that spanned the island sequence [see Fig. 2A]. By comparing the aprt “G ladder” produced by dimethylsulfate [DMS] piperidine cleavage of naked DNA with that from DMS-treated F9 cells, we could establish sites of protein–DNA interaction. Unlike DNase I footprinting in vitro, where protein binding is shown as large gaps in the DNase I ladder, footprints using DMS are more subtle and are identified by missing or hypersensitive guanosine bases in the DNA from DMS-treated cells. Three footprints were obtained [Fig. 2B], which coincided with the three GC boxes identified in the aprt promoter [Dush et al. 1985]. The GC box is the consensus binding site for the ubiquitous transcription factor Sp1 [Kadonaga et al. 1986]. In the mouse aprt gene these sites have been shown previously to bind purified Sp1 by DNase I footprinting in vitro and are sufficient for transcription of the gene [Dush et al. 1988]. Aside from the Sp1 sites, no

Figure 2. In vivo footprinting of bound factors. (A) A schematic diagram showing regions that were analyzed by LMPCR. [Top] The position of the CpG dinucleotides as in Fig. 1. The position of overlapping primer pairs used for in vivo footprinting are indicated relative to the CpG plot above, by a dot and number at the end of the arrowed lines. The length of the arrowed line corresponds to the region analyzed with each primer set; arrows pointing to the left are primers complementary to the lower strand of the sequence, arrows to the right are primers complementary to the upper strand. The positions of the three GC boxes are indicated by three thick vertical lines under the CpG plot. The positions of methylation-sensitive restriction enzyme cleavage sites within the CpG island are indicated: (H) HpaII; (S) SmaI. The bracketed region under the CpG plot shows the position of the BsaBI–MaeII fragment [see text]. The scale at the top is in base pairs. [B] Results of in vivo footprinting reactions obtained using primer sets 4 [lanes 1–3] and X1, [lanes 4–6]. LMPCR reactions were run on standard sequencing gels, electroblotted, and hybridized to an aprt PCR probe [primers l–X1, see A]. Each lane shows the G ladder obtained from DMS-treated purified F9 DNA [lanes 1,4] or from F9 cells treated with DMS for 2 min [lanes 2,5] or 5 min [lanes 3,6]. Brackets show the positions of GC boxes (I, II, and III) that correspond to the DNA sequences below. Small circles show the position of weak or absent G bands (○) or hyper-reactive G bands (●). Arrows A and B show the position of transcription start sites relative to the GC boxes as mapped by Dush et al. [1988]. Footprints are detected at each of the three GC boxes, on both strands of the DNA, and were confirmed by repeating the reactions several times. (C) Results of footprinting reactions with two primer sets, 5 and 8. Lanes 1 and 4 are LMPCR amplifications of purified F9 DNA after DMS/piperidine cleavage. LMPCR amplifications of DMS-treated F9 cells for 2 min [lanes 2,5] or 5 min [lanes 3,6]. Although some variation in band intensity can be detected in these gels, it was not reproducible [repeated twice] and no footprints were identified using these primer sets.
other factor binding sites have been identified within the promoter [Dush et al. 1988] and there is no TATA box [Dush et al. 1985]. In our analysis, no other footprints could be identified using the other primer sets [e.g., Fig. 2C]. These results suggest that protein factors are not bound over the entire island but only within the promoter that is at its left-hand edge [see Fig. 2A].

**Nucleosome-like structures are positioned over the CpG island**

Transcription factors are not the only proteins expected to bind CpG island DNA, as nucleosomes are also likely to be present. To determine nucleosome position, F9 nuclei were incubated with micrococcal nuclease, which cleaves between nucleosomes. Sites cut by the enzyme in the *aprt* gene were detected by LMPCR using three primer sets spanning the CpG island region [primers 1, X1, and 7; Fig. 3B]. Micrococcal nuclease cleavage sites were clustered in the region that includes the three occupied GC boxes and transcription start sites [Fig. 3A]. The same region is also hypersensitive to cleavage by DNase I and endogenous nucleases [data not shown]. In addition to these cleavage sites, micrococcal nuclease cuts at ~200-bp intervals across the CpG island [primer 1, Fig. 3A]. At the boundaries of the island, cleavage sites

![Figure 3](https://example.com/figure3.png)

**Figure 3.** In vivo footprinting of nucleosomes. (A) Results of in vivo footprinting of nucleosomes using LMPCR amplification of DNA obtained from F9 nuclei restricted with micrococcal nuclease. The products were electrophoresed on 2% agarose gels, blotted to nylon membranes, and hybridized to an *aprt* probe [for primer sets 1 and X1, a PCR probe corresponding to position 45–883 of the *aprt* sequence was used and for primer set 7, a PCR probe between 1265 and 1502 was used]. Each pair of tracks shows LMPCR products from nuclei digested with 0.16 units of enzyme [lanes 1,3,5] or 0.32 units [lanes 2,4,6]. The primer sets used are indicated underneath. The open vertical rectangles represent protected regions and their position, relative to the CpG plot, is shown in B. The ladder of horizontal bands alongside the autoradiographs shows the position of 123-bp marker bands [BRL]. (HS) Nuclease hypersensitive region from our experiments. (B) The CpG plot of the *aprt* gene as described in Fig. 2. As in Figure 2A, the positions of the primers used to amplify micrococcal nuclease-cleaved DNA are indicated by a dot on the numbered, arrowed line under the CpG plot. The length of the arrowed line corresponds to the region analyzed with each primer set; arrows pointing to the left are primers complementary to the lower strand of the sequence, arrows to the right are primers complementary to the upper strand. The sites cleaved by the enzyme relative to the CpG plot are indicated by vertical arrows below. The arrow size is an approximate indication of the strength of signal obtained by hybridization to the probe as determined from A. The numbered circles represent the protected regions [which we suggest are nucleosomes]. [HS] The micrococcal nuclease/DNase I hypersensitive site.
Macleod et al.

are weak and irregularly spaced [primers X1 and 7, Fig. 3A]. The size of the regions that are protected from the nuclease implies that they correspond to nucleosomes that are uniquely positioned across the island (Fig. 3B).

Deletion of the aprt promoter leads to methylation of the CpG island

The occupied GC boxes are at the left edge of the CpG island (Fig. 2A). To determine whether Sp1 binding in this region alone could protect the whole island from methylation, we tested the consequences of deleting or mutating the promoter region. To assay methylation, transgenic mice were generated by the injection of DNA constructs into the male pronucleus of fertilized mouse eggs, which were then inserted into the uterus of pseudopregnant mice. The transgenic offspring were subsequently identified by analysis of tail DNA. The control construct for injection comprised a 3.4-kb BgiII–SphI fragment, pABS, containing the unaltered aprt island that had been subcloned from the genomic aprt clone pSAM 6.3 (Turker et al. 1991). A deletion (pAAB) was made by removing a 210-bp BsaBl–MaeII fragment that includes the transcriptional start sites and GC boxes from pABS (see Fig. 4). Each of the constructs was digested with EcoRI and SphI, which removed all prokaryotic DNA sequence, except for a short region of polylinker containing a KpnI site. These constructs were purified and used to produce transgenic mice. Analysis of tail DNA from founder animals was then carried out by Southern blotting to determine the methylation status of the transgenes. In each case, this was facilitated by the presence of the KpnI site within the polylinker at the 5’ end of the transgene (see Fig. 5A). Transgenes could be distinguished from the endogenous gene, as they gave a 1.7-kb restriction fragment (pABS) and a 1.5-kb fragment (pAAB) compared with the 5-kb fragment derived from the endogenous gene in KpnI digests. Methylation was assayed at several sites within the island using the methylation-sensitive restriction enzymes Smal and HpaII (see Fig. 2A). If the Smal site was nonmethylated, a KpnI–Smal double digest produced a transgene-specific 1.1-kb restriction fragment with pABS and pAAB. The 2.5-kb fragment produced from the endogenous gene served as an internal control for complete digestion. Further sites were tested by KpnI–HpaII digestion, which provided additional information on the level of methylation at other sites within the CpG island (Figs. 2A and 5A).

In an analysis of 10 founder mice with the wild-type construct pABS, (Fig. 5B) all but one were free of methylation at the Smal site, as the 1.7-kb KpnI restriction fragment was abolished completely by this enzyme. These transgenes were also nonmethylated at many, if not all, HpaII sites (Fig. 5B). Thus, the nonmethylated CpG island is reactivated at the aprt transgene in nearly all of the founder animals. Eleven pAAB founder mice were also tested in the same manner (Fig. 5C). Unlike the controls, the Smal site was heavily methylated in these transgenes, as the 1.5-kb restriction fragment is largely resistant to cleavage, and very little of the 1.1-kb transgene-specific fragment, which would be expected if this site were nonmethylated, is seen. The HpaII sites were also heavily methylated, as most of the transgene-specific 1.5-kb KpnI fragment is uncut by this enzyme. The

Figure 4. DNA sequence that was deleted (pAΔB) or mutagenized (pAZM2) from the wild-type aprt clone [pABS]. Part of the aprt sequence from the aprt promoter is shown (Dush et al. 1985). The arrows indicate the position of the transcription start sites. The positions of the BsaBl and MaeII sites are indicated. These enzymes were used to cut this fragment (210 bp) from the wild-type parent clone, pABS, to obtain pAAB. The deletion removes the three GC boxes and transcription start sites. The GC boxes are bracketed and labeled (I, II, and III). The base changes in each box [CGC → TTT] to create clone pAZM2 from the parent clone pABS are shown underneath. A single T → A change, which creates an XhoI site, was used for screening purposes (see Materials and methods).

Figure 5. Methylation analysis of transgenic founder mice using methylation-sensitive restriction enzymes. [A] Restriction enzyme map of the endogenous genomic aprt gene and transgenes. Each vertical line on the map shows the position of KpnI (K), Smal (S), BgiII (B), SphI (P), and HpaII (H) sites. One of the HpaII sites (outside the CpG island), indicated by an asterisk (‘), is partially methylated in the genome (see Turker et al. 1989). The solid horizontal bar represents the position of the probe used to hybridize to the Southern blots. The CpG island is denoted by brackets. [B] A Southern blot obtained from tail DNA of wild-type (+) or pABS transgenic mice (T1–T4) hybridized to an aprt-specific probe [see A]. DNA was restricted with KpnI [lanes 1,4,7,10,13], KpnI + Smal [lanes 2,5,8,11,14], or KpnI + HpaII [lanes 3,6,9,12]. [M] Radioactively labeled 1-kb marker (BRL); the sizes of some hybridizing restriction fragments are shown in the margin (in bp). The arrowed T shows the position of the transgene-specific restriction fragment in the KpnI digests. [C] As above, but with four pAΔB transgenic mice. DNA was restricted with KpnI [lanes 1,4,7,10], KpnI + Smal [lanes 2,5,8,11], or KpnI + HpaII [lanes 3,6,9,12]. [D] As above, but with four pAZM2 transgenic mice. DNA was restricted with KpnI [lanes 1,4,7,10], KpnI + Smal [lanes 2,5,8,11], or KpnI + HpaII [lanes 3,6,9,12].
degree of methylation varied between founder animals, but the transgenes were all shown to be highly methylated except in 1 of 11 examples, where the level of methylation was low (Fig. 5C).

Figure 5. (See facing page for legend.)
in methylation of the CpG island promoter, the transcription start sites, and also includes the Sp1 sites themselves were important. Specific mutations were introduced simultaneously into the three GC boxes by in vitro mutagenesis to produce the pAZM2 construct (see Fig. 4). To test whether mutagenesis of these sites occluded Sp1 binding, a 210-bp fragment containing the intact sites (pABS) or the mutated sites (pAZM2) was end-labeled and incubated with a crude protein extract from F9 cells. Protein–DNA interaction was shown by bandshift (Fig. 6). The pABS fragment formed a complex with the extract that was effectively competed out by a double-stranded oligomer containing a high affinity binding site for Sp1 (see Fig. 6B, lanes 2,3). This confirmed that Sp1 bound to this fragment. No other factors besides Sp1 interact with the probe to produce a complex. The pAZM2 fragment, in which the Sp1-binding sites have been mutated (see Fig. 4A), does not form a complex (Fig. 6B, lane 5). In a reciprocal experiment (Fig. 6C) a DNA duplex containing a high affinity binding site for Sp1 was used in competition with a double-stranded wild-type aprt oligomer (N1) or an oligomer carrying mutations in the GC boxes (M1) (see Fig. 6A). The results of the bandshifts clearly show that only the wild-type oligomer effectively competes with the high affinity site for Sp1 (Fig. 6C, lane 3). The mutant M1 oligonucleotide is unable to compete (Fig. 6C, lane 4).

The EcoRI–Spbl fragment was purified from the pAZM2 clone and used to produce transgenic mice. The results of the methylation analysis on four of the founder animals is shown in Figure 5D. As with pΔAB, the transgenes are highly methylated in four of five founder animals tested as the 1.7-kb KpnI fragment is largely resistant to digestion with Smal and HpaII (Fig. 5D). The transgene was nonmethylated in the fifth animal. This analysis shows that Sp1 sites are required to ensure that the aprt gene remains methylation free.

**Discussion**

We have shown that Sp1 elements play an important part in the maintenance of the aprt CpG island. Transgene constructs in which the GC boxes in the aprt promoter have been deleted or mutated no longer bind Sp1 and no longer exclude methylation from the CpG island [Table 1]. It has been suggested previously that factors may be required to protect CpG islands (Bird 1986) and also that Sp1 may be involved (Höller et al. 1988). Our data provide direct evidence for this view. Previous attempts to define the origin of methylation-free islands have not implicated factor binding, but with hindsight

**Mutagenesis of the aprt Sp1-binding sites results in methylation of the CpG island**

The above result showed that the 210-bp fragment from the aprt promoter was required to prevent methylation of the CpG island. As this region contains the entire promoter, the transcription start sites, and also includes part of the CpG island (see Fig. 2A), we tested whether the Sp1 sites themselves were important. Specific mutations were introduced simultaneously into the three GC boxes by in vitro mutagenesis to produce the pAZM2 construct (see Fig. 4). To test whether mutagenesis of these sites occluded Sp1 binding, a 210-bp fragment containing the intact sites (pABS) or the mutated sites (pAZM2) was end-labeled and incubated with a crude protein extract from F9 cells. Protein–DNA interaction was shown by bandshift (Fig. 6). The pABS fragment formed a complex with the extract that was effectively competed out by a double-stranded oligomer containing a high affinity binding site for Sp1 (see Fig. 6B, lanes 2,3). This confirmed that Sp1 bound to this fragment. No other factors besides Sp1 interact with the probe to produce a complex. The pAZM2 fragment, in which the Sp1-binding sites have been mutated (see Fig. 4A), does not form a complex (Fig. 6B, lane 5). In a reciprocal experiment (Fig. 6C) a DNA duplex containing a high affinity binding site for Sp1 was used in competition with a double-stranded wild-type aprt oligomer (N1) or an oligomer carrying mutations in the GC boxes (M1) (see Fig. 6A). The results of the bandshifts clearly show that only the wild-type oligomer effectively competes with the high affinity site for Sp1 (Fig. 6C, lane 3). The mutant M1 oligonucleotide is unable to compete (Fig. 6C, lane 4).

**Table 1. Summary of the methylation results from transgenic mice generated with the three DNA constructs pABS, pΔAB, and pAZM2**

|         | Methylated | Nonmethylated |
|---------|------------|---------------|
| pABS    | [1]        | 9             |
| pΔAB    | 10         | 1             |
| pAZM2   | 4          | 1             |

The numbers refer to the number of founder animals tested. Parentheses indicate weakly methylated.
the results that were obtained are compatible with the conclusions drawn here. Szyf et al. (1990) identified a 214-bp fragment from the mouse Thy-1 gene promoter that contained a "portable signal," which protected the Thy-1 gene (and plasmid vector) from methylation after transfection into embryonic stem (ES) cells. It was shown later that this DNA fragment contains binding sites for transcription factors, including Sp1 (Spanopoulos et al. 1991). Mummaneni et al. (1993) proposed that the boundaries of CpG islands have sequences that block the spread of methylation from nearby methylation centers. The existence of methylation centers was inferred from the finding that a CpG island fragment from the aprt gene became methylated when placed upstream of its normal position following transfection into EC cells. The interpretation was that the island was unable to resist the influence of a methylation center when placed close to it. We note, however, that the CpG island fragment that was translocated omits two of the three GC boxes. Based on the work presented here, it seems likely that the loss of these sites may have been sufficient to lead to de novo methylation of the island wherever it was placed. Thus, these results are consistent with the proposal that binding sites for a group of factors, which may or may not always include Sp1, are required for protection of CpG islands from methylation. It has been noted previously that GC boxes are frequently found within CpG islands (Gardiner-Garden and Frommer 1987).

It is surprising that occupied Sp1 sites at one edge of the aprt CpG island are capable of preventing methylation of CpG sites some distance downstream. That these are the only occupied sites for sequence-specific binding proteins in the island is suggested by the absence of additional footprints in vivo and the presence of typically spaced nucleosomes across the island. The DNase I/micrococcal nuclease hypersensitive site is also located at the island periphery, coincident with the Sp1 sites and the transcription start sites. In the great majority of cases, CpG islands extend downstream from the promoter of a gene into its transcription unit. Thus, it may be generally true that bound transcription factors mark the 5' boundary of an island, the methylation-free domain extending several hundred base pairs downstream from this point. How factor binding prevents methylation and why the nonmethylated domain is relatively constant in length are unanswered questions at present.

Their asymmetrical location makes it unlikely that the Sp1 sites exclude methylation by sterically preventing access by the methyltransferase. If this were the case, the methylation-free region would be expected to straddle the bound factors symmetrically. As shown previously (Tazi and Bird 1990), CpG islands are generally highly accessible to proteins [nucleases] and have features of "open chromatin" such as hyperacetylated histones and depletion of histone H1. It is difficult to sustain the argument that such an open structure should be much less accessible to the methyltransferase than typical nucleosomal chromatin. It is more likely that the binding of factors to the island excludes DNA methylation by a mechanism other than steric hindrance.

It may be significant that the presence of Sp1 is also required to drive transcription of the aprt promoter (Dush et al. 1988). Thus, changes that abolish promoter activity also abolish the ability of the island to remain free of methylation. The relationship between transcription and lack of CpG island methylation, however, remains a puzzle. Many genes with CpG islands, such as the human a-globin gene, are highly tissue specific in expression yet are nonmethylated in expressing and nonexpressing tissues alike (Bird et al. 1987). A possible explanation is that genes of this kind are poised for transcription but do not, for some reason, make stable RNA.

We have investigated this possibility for the human a-globin gene and the mouse myoD1 gene by assaying for run-on transcription in isolated nuclei of nonexpressing cells and have failed to detect any inappropriate transcription [F. Antequera and D. Macleod, unpubl.]. The results confirm similar findings for the tissue-specific Thy-1 gene, which also has a CpG island (Kolsto et al. 1986). In addition, the results agree with evidence that neither the myoD1 gene (Mueller and Wold 1989) nor the a-globin gene (Yagi and Groudine 1986) display open chromatin or DNase I hypersensitive sites in nonexpressing tissues. Thus, CpG islands can be maintained methylation free in the apparent absence of either transcription or an open chromatin configuration. An alternative possibility is that there is a stage in early embryos where all island-associated genes, including tissue-specific genes, are transcribed or transcriptionally competent. Lack of methylation could be established at this time and be replicated because of maintenance methylation in somatic cells, where transcription no longer occurs. Future experiments are designed to test this and other possibilities.

Materials and methods

Bisulfite modification and sequencing of genomic DNA

This procedure was essentially as described by Frommer et al. (1992). Mouse liver DNA (10 μg) was first cleaved with EcoRI and then denatured in 0.2 M NaOH for 10 min at 27°C in a volume of 110 μl. The reaction was neutralized with 44 μl of 5 M ammonium acetate and precipitated with 620 μl of ethanol. The pellet was washed with 70% ethanol, air-dried, and resuspended in 100 μl of cold sterile distilled water on ice, and 1.04 ml of fresh 3.8 M sodium bisulfite and 60 μl of 10 mM Quinol were added. This mixture was overlayed with 200 μl of mineral oil and incubated at 50°C in the dark for 16 hr. The DNA was desalted and concentrated using GeneClean (Biol01), ethanol precipitated, and resuspended in sterile water. Aliquots were stored at −20°C. An aliquot of DNA was amplified using the modified primers CGAAATTCCTCCCATCCTCTTATTTAAACCAT and TAGAATTCACCCCAATTACTATAACTAAA, which are derived from the aprt sequence at positions 70–91 and 2390–2414 (Dush et al. 1985) and then reamplified using the internal set, TTGAATTCTTTGGTTGTTGAATTTTTGATT and TTGAATTCTTGGGGTATGGAATTTTAGGTATA at positions 191–210 and 2101–2128. PCR reactions were carried out in Thesit buffer (Ponce and Micol 1992). The resulting PCR
fragments, containing EcoRI ends, were cloned into pBluescribe vector (Stratagene) and sequenced using Sequenase [U.S. Biological].

In vivo footprinting using DMS

DMS treatment of F9 DNA or whole cells and subsequent cleavage by piperidine, ligation of linkers, and LMPCR reactions were as described [Pfeifer et al. 1989] except for the following modifications. PCR reactions were in Thesit buffer [Ponce and Micil 1992], 30 mM tricine at pH 8.4, 2 mM MgCl₂, 50 mM β-mercaptoethanol, 0.1% gelatin, 0.1% Thesit with the addition of 10% DMSO. In some reactions, 1 μg of single stranded-binding protein (Promega) was also included in the Sequenase extension step, which improved results from G + C-rich regions. The sequence of each internal primer, of each nested set, is as follows: (1) CCAATTGGACCTCCCCACACC (position 52-72); (2) GGTTCACCCAAAAGGAGG (253-272); (3) GCCCTTGCTATGCGGCAG (453-472); (4) GCTTGTTTATTTGCAGCTGAAG (653-674); (5) CCGCTGCTCTACACGCAG (847-867); (6) CCTGCTTCCTAGGATATC (1071-1089); (7) GTCCTGCCACCCACCATCC (1266-1284); (8) GGGGACACAAATGCCGCG (1471-1489); and (X1) GGTTCCACCCCAAAACGAGG (253-272).

In vivo footprinting using micrococcal nuclease

Nuclei were prepared from 80% confluent F9 cells using the method of Shimada et al. (1986). Cells were harvested and resuspended in R buffer [10 mM Tris at pH 7.5, 10 mM NaCl, 3 mM MgCl₂, 0.1 mM PMSF, and 0.25 mM sucrose]. Cells were lysed by the addition of RNase-free DNase to a final concentration of 0.25% and gentle homogenization with a Dounce. Nuclei were pelleted through a sucrose cushion (1.1 M sucrose in buffer R) twice at 5000 rpm for 5 min and finally resuspended in 50 mM Tris, 5 mM MgCl₂, 0.1 mM EDTA. Glycerol was added to 40%, and the nuclei were frozen at −70°C in aliquots equivalent to 300–500 μg DNA/ml. Nuclei (1 ml aliquot) were pelleted for 5 min at 5000 rpm in an Eppendorf centrifuge and resuspended in 600 μl of buffer M [Shimada et al. 1986; 50 mM Tris at pH 7.4, 60 mM KCl, 3 mM CaCl₂, and 0.34 mM sucrose] and 100 μl aliquots were digested with increasing amounts of micrococcal nuclease [Worthington, resuspended at 17 units/μl in 10 mM Tris at pH 7.5, 0.5 mM EDTA, 0.5 mM DTT, and 50% glycerol]. Reactions were incubated at 37°C for 4 min and terminated with an equal volume of stop mix (1% SDS, 0.6 M NaCl, 20 mM EDTA, and 20 mM Tris-HCl at pH 7.5). Proteinase K was added to a final concentration of 100 μg/ml, and tubes were incubated at 55°C for 1 hr before extracting the DNA using established procedures. Digestion parameters sufficient to produce a nucleosomal ladder were determined empirically. DNA (3 μg) was subsequently used for LMPCR amplification [see above], and products were resolved on 2% agarose gels in TAE buffer (40 mM Tris-acetate, 1 mM EDTA at pH 8). DNA was blotted onto nylon membranes (Hybond N+, Amersham) and hybridized to labeled probe.

In vitro mutagenesis

An 84-bp DNA oligomer, M1 [Fig. 6A] was used to generate three base pair changes simultaneously in each of the three GC boxes using the method of Kunkel et al. [1987] using a mutagenesis kit [Muta-gene, Bio-Rad]. Clones were selected that contained an introduced XhoI site and then sequenced using an automatic DNA sequencer (ABI) to check that all of the desired base changes had been made.

Analysis of protein–DNA interactions by agarose-gel bandshift

A crude protein (NUN) extract was prepared from F9 nuclei [see above] using the method of Lavery and Schibler [1993]. Nuclei were pelleted, as described above, from frozen stocks, and nine volumes of NUN buffer (final concentration 0.3 M NaCl, 1 M urea, 1% NP-40, 25 mM HEPES at pH 7.6, 1 mM DTT, and 0.1 mM PMSF) were added. The mixture was vortexed for 5 sec before placing on ice for 15 min. The chromatin precipitate was removed by centrifugation in a microcentrifuge for 10 min at 10,000 rpm at 4°C, and glycerol was added to a final concentration of 10%. Aliquots were frozen at −70°C, and protein concentration was determined using a Bio-Rad protein assay reagent. Binding reactions were carried out as described by Somma et al. [1991]. In a 20 μl reaction mixture, 5 μg of F9 protein extract was preincubated on ice with 1 μg of nonspecific poly[dI-Cl] competitor [in some reactions, 0.1 μg of specific competitor was also added] in 25 mM HEPES, 60 mM KCl, 1 mM DTT, 0.5 mM EDTA, and 8.7% glycerol for 10 min. The radioactive probe was then added and incubation was continued for an additional 30 min on ice. Restriction fragments from cloned DNA or PCR products were end-labeled with Klenow and 32P-labeled nucleotides. Generally, 20–50 pg of probe was used in the binding reaction. The reactions were electrophoresed in 1.5% agarose gels in 0.5 × TBE buffer, subsequently dried under vacuum onto DE 81 paper, and exposed to Kodak XAR5 autorigradiographic film.

Production of transgenic mice

Transgenic mice were generated by microinjection of the respective transgene fragment, at a concentration of 1 μg/ml in 10 mM Tris-HCl, 0.1 mM EDTA (pH 7.5), using standard procedures [Hogan et al. 1986]. Transgenes were injected into single-cell embryos isolated from a cross between C57Bl/6 × CBA/Ca F1 mice. Offspring were weaned at 3 weeks of age, and tail biopsies performed for preparation of DNA. This was carried out by incubating the tail tips in proteinase K buffer (50 mM Tris at pH 8, 100 mM NaCl, 100 mM EDTA, 1% SDS, and 60 μg/ml of proteinase K) overnight at 55°C, and then extraction of nucleic acids was performed with phenol.

Acknowledgments

We thank Genevieve Stapleton for her advise on using the LMPCR method, Gillian Brooker, Aileen Greig, and Joan Davidson for their technical support, and Frank Johnson and Graham Brown for their photographic skills. We also thank Sally Cross and Richard Meehan for their helpful comments on the manuscript. This work was supported by The Wellcome Trust, The Imperial Cancer Research Fund, and The Howard Hughes Medical Institute. J.M. was supported by an Agricultural and Food Research Council [AFRC] grant [A550/40].

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

Note added in proof

Brandeis et al. [1994] [Nature, in press] have independently shown the importance of Sp1 sites in maintaining the methylation-free status of a CpG island at the hamster aprt gene.
References

Antequera, F. and A. Bird. 1993. Number of CpG islands and genes in human and mouse. Proc. Natl. Acad. Sci. 90: 11985–11990.

Antequera, F., J. Boyes, and A. Bird. 1990. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines. Cell 62: 503–514.

Bird, A.P. 1986. CpG-rich islands and the function of DNA methylation. Nature 321: 209–213.

——. 1987. CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3: 342–347.

Bird, A., M. Taggart, M. Frommer, O.J. Miller, and D. Macleod. 1987. CpG islands in non-methylated, CpG-rich DNA. Cell 40: 91–99.

Bird, A.P., M.H. Taggart, R.D. Nicholls, and D.R. Higgs. 1987. Non-methylated CpG-rich islands at the human alpha-globin locus: Implications for evolution of the alpha-globin pseudogene. EMBO J. 6: 999–1004.

Brandeis, M., T. Kafri, M. Ariel, J.R. Chaillet, J. McCarrey, A. Antequera, F., J. Boyes, and A. Bird. 1990. High levels of de novo methylation and altered chromatin structure at CpG islands in cell lines of non-methylated, CpG-rich DNA. Cell 62: 503–514.

Bird, A.P. 1986. CpG-rich islands and the function of DNA methylation. Nature 321: 209–213.

——. 1987. CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3: 342–347.

Bird, A., M. Taggart, M. Frommer, O.J. Miller, and D. Macleod. 1987. CpG islands in non-methylated, CpG-rich DNA. Cell 40: 91–99.

Bird, A.P., M.H. Taggart, R.D. Nicholls, and D.R. Higgs. 1987. Non-methylated CpG-rich islands at the human alpha-globin locus: Implications for evolution of the alpha-globin pseudogene. EMBO J. 6: 999–1004.

Carotti, D., F. Palitti, P. Lavia, and R. Strom. 1989. In vitro methylation of CpG-rich islands. Nucleic Acids Res. 17: 9219–9229.

Choi, Y.-C. and C.-B. Chae. 1991. DNA hypomethylation and germ cell-specific expression of testis-specific H2B histone gene. J. Biol. Chem. 266: 20504–20511.

——. 1993. Demethylation of somatic-specific histone H2A and H2B genes in F9 embryonal carcinoma cells. Mol. Cell. Biol. 13: 5538–5548.

Cooper, D.N., M.H. Taggart, and A.P. Bird. 1983. Unmethylated domains in vertebrate DNA. Nucleic Acids Res. 11: 647–658.

Driscoll, D.J. and B.R. Migeon. 1990. Sex difference in methylation of single-copy genes in human meiotic germ cells: Implications for X chromosome inactivation, parental imprinting, and origin of CpG mutations. Somatic Cell Mol. Genet. 16: 267–282.

Dush, M.K., J.M. Sikela, S.A. Khan, J.A. Tischfield, and P.J. Stambrook. 1985. Nucleotide sequence and organization of the mouse adenine phosphoribosyltransferase gene: Presence of a coding region common to animal and bacterial phosphoribosyltransferases that has a variable intron/exon arrangement. Proc. Natl. Acad. Sci. 82: 2731–2735.

Dush, M.K., M.R. Briggs, M.E. Royce, D.A. Schaff, S.A. Khan, J.A. Tischfield, and P.J. Stambrook. 1988. Identification of DNA sequences required for mouse APRT gene expression. Nucleic Acids Res. 16: 8509–8525.

Ferguson-Smith, A.C., H. Sasaki, B.M. Cattanach, and M.A. Surani. 1993. Parental-origin-specific epigenetic modification of the mouse H19 gene. Nature 362: 751–755.

Frank, D., I. Kesher, M. Shani, A. Levine, A. Razin, and H. Cedar. 1991. Demethylation of CpG islands in embryonic cells. Nature 351: 239–241.

Frommer, M., L.E. McDonald, D.S. Millar, C.M. Collis, J. Watt, G.W. Grigg, P.L. Molloy, and C.L. Paul. 1992. A genomic sequencing and methylation analysis by PCR-mediated cloning method. Nature 351: 239–241.

Gardiner-Garden, M. and M. Frommer. 1987. CpG islands in vertebrate genomes. J. Mol. Biol. 196: 261–282.

Gunderson, C., A-B. KoloSt, and H. Prydz. 1991. Differential methylation of a CpG-island concatamer in hemi- and homozygous transgenic mice. FEBS Lett. 295: 214–218.

Hogan, B., F. Constantini, and E. Lacy. 1986. Manipulating the Mouse Embryo. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Höller, M., G. Westin, J. Jiricny, and W. Shaffner. 1988. SpI transcription factor binds DNA and activates transcription even when the binding site is CpG methylated. Genes & Dev. 2: 1127–1135.

Holliday, R. and J.E. Pugh. 1975. DNA modification mechanisms and gene activity during development. Science 186: 226–232.

Jones, P.A., M.J. Wolkowitz, W.M.I. Rideout, F.A. Gonzales, C.M. Marziasz, G.A. Coetzee, and S.J. Tapsccott. 1990. De novo methylation of the MyoD1 CpG island during the establishment of immortal cell lines. Proc. Natl. Acad. Sci. 87: 6117–6121.

Kadonaga, J., K. Jones, and R. Tjian. 1986. Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. Sci. 11: 20–23.

Kolsto, A.-B., G. Kollias, V. Giguere, K.-I. Isobe, H. Prydz, and F. Grosved. 1986. The maintenance of methylation-free islands in transgenic mice. Nucleic Acids Res. 14: 9679–9678.

Kunkel, T.A., J.D. Roberts, and R.A. Zalkor. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. 154: 367–382.

Lavery, D.J. and U. Schibler. 1993. Circadian transcription of the cholesterol 7a hydroxylase gene may involve the liver-enriched bZIP protein DBP. Genes & Dev. 7: 1871–1884.

Mehtali, M., M. LeMeur, and R. Lathc. 1990. The methylation-free status of a housekeeping transgene is lost at high copy number. Gene 91: 179–184.

Mueller, P.J. and B. Wold. 1989. In vivo footprinting of a muscle specific enhancer by ligation mediated PCR. Science 246: 780–786.

Mummmaneni, P., P.L. Bishop, and M.S. Turker. 1993. A cis-acting element accounts for a conserved methylation pattern upstream of the mouse adenosine phosphoribosyltransferase gene. J. Biol. Chem. 268: 552–558.

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.

Pfeifer, G.P., S.D. Steigerwald, F.R. Mueller, B. Wold, and A.D. Riggs. 1989. Genomic sequencing and methylation analysis by ligation-mediated PCR. Science 246: 810–813.

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.

Ponce, M.R. and J.L. Micol. 1992. PCR amplification of long DNA fragments. Nucleic Acids Res. 20: 623.

Riggs, A.D. 1975. X-inactivation, differentiation and DNA methylation. Cytogenet. Cell. Genet. 14: 9–25.

Riggs, A.D. and G.P. Pfeifer. 1992. X-chromosome inactivation and cell memory. Trends Genet. 8: 169–174.

Shimada, T., K. Inokuchi, and A.W. Nienhuis. 1986. Chromatin structure of the human dihydrofolate reductase gene promoter. J. Biol. Chem. 261: 1445–1452.

Somma, M.P., I. Gambino, and P. Lavia. 1991. Transcription factors binding to the mouse HTF9 housekeeping promoter differ between cell types. Nucleic Acids Res. 19: 4451–4458.

Spanopoulou, E., V. Giguere, and F. Grosved. 1991. The functional domains of the Thy-1 gene promoter. Mol. Cell. Biol. 11: 2216–2228.

Stapleton, G., M.P. Somma, and P. Lavia. 1993. Cell type-specific interactions of transcription factors with a housekeep-
Macleod et al. ing promoter in vivo. *Nucleic Acids Res.* 21: 2465–2471.

Stöger, 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.

Szyf, M., G. Tanigawa, and P.L.J. McCarthy. 1990. A DNA signal from the thy-1 gene defines de novo methylation patterns in embryonic stem cells. *Mol. Cell. Biol.* 10: 4396–4400.

Tazi, J. and A. Bird. 1990. Alternative chromatin structure at CpG islands. 1990. *Cell* 60: 909–920.

Tykocinski, M.L. and E.C. Max. 1984. CG clusters in MHC genes and 5′ demethylated genes. *Nucleic Acids Res.* 12: 4385–4396.

Turker, M.S., K. Swisshelm, A.C. Smith, and G.M. Martin. 1989. A partial methylation profile for a CpG site is stably maintained in mammalian tissues and cultured cell lines. *J. Biol. Chem.* 264: 11632–11636.

Turker, M.S., P. Mummaneni, and P.L. Bishop. 1991. Region- and cell type-specific de novo DNA methylation in cultured mammalian cells. *Somatic Cell Mol. Genet.* 17: 151–157.

Yagi, M. and M. Groudine. 1986. Chromatin structure and developmental expression of the human alpha-globin cluster. *Mol. Cell. Biol.* 6: 1108–1116.
Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island.

D Macleod, J Charlton, J Mullins, et al.

*Genes Dev.* 1994, 8:
Access the most recent version at doi:10.1101/gad.8.19.2282

---

**References**

This article cites 46 articles, 18 of which can be accessed free at:
http://genesdev.cshlp.org/content/8/19/2282.full.html#ref-list-1

**License**

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here.](#)