DNA Methylation

Molecular Biology and Biological Significance

Edited by J. P. Jost
H. P. Saluz

Birkhäuser Verlag
Basel · Boston · Berlin

Contents

DNA Methylation: Molecular Biology and Biological Significance
ed. by J. P. Jost & H. P. Saluz

Structure, function and regulation of mammalian DNA methyltransferase

Heinrich Leonhardt and Timothy H. Bestor

Department of Anatomy and Cellular Biology, Laboratory of Human Reproduction and Reproductive Biology, Harvard Medical School, 45 Shattuck St, Boston, Massachusetts 02115, USA

1 Introduction

The haploid mammalian genome contains $\sim 5 \times 10^7$ CpG dinucleotides (Schwartz et al., 1962), about 60% of which are methylated at the 5 position of the cytosine residue (Bestor et al., 1984). The unmethylated fraction of the genome is exposed to diffusible factors in nuclei (Antequera et al., 1989), perhaps due to the action of proteins which bind to methylated sequences and induce their condensation (Meehan et al., 1989). Methylation may therefore control the availability of regulatory sequences for interaction with the transcriptional apparatus. Activation of tissue-specific genes is often accompanied by the disappearance of methyl groups from promoter regions, and differentiated cell types display characteristic unique methylation patterns. It has been argued that the selective advantage of such a regulatory mechanism would be expected to be most pronounced for those organisms with large genomes, and in fact 5-mC is absent from the DNA of most organisms with genomes smaller than 5×10^8 base pairs but essentially universal among organisms having genomes above this size (Bestor, 1990).

Methylation patterns undergo sweeping reorganization during gametogenesis and early development (Monk, 1990). Measurements of bulk 5-mC levels (Monk et al., 1987) and studies of an imprinted transgene (Chaillet et al., 1991) have shown that the DNA of primordial germ cells has a very low 5-mC content, and that sperm DNA is relatively more methylated than oocyte DNA. Methylation levels actually decline significantly in the preimplantation embryo (with the paternal DNA being more affected) to reach a minimum at the blastocyst stage. Methylation levels increase in the postimplantation embryo and adult levels of 5-mC are attained only after completion of gastrulation (Monk, 1990). It must be pointed out that these findings are averaged over a very large number of CpG sites, and that the behavior of many

individual DNA sequences may be quite different than the genome-size average. It should also be noted that the number of methylated CpG sites exceeds the number of genes by a factor of about 50, and it is likely that the methylation status of a significant proportion of CpG dinucleotides is not subject to close regulation; this is consistent with the finding that many CpG sites show partial methylation in clonal cell populations. Strain-specific modifiers affect the methylation status of imprinted transgenes in mice and are likely to influence the methylation status of endogenous DNA sequences as well (Engler et al., 1991), and patterns of methylated CpG sites around certain genes undergo changes in aging mice (Uehara et al., 1989). The above findings make it clear that vertebrate methylation patterns are dynamic and subject to genetic and developmental control.

Several contributors to this volume discuss the role of methylation patterns in a variety of biological processes. Here we will be concerned with the mechanisms which establish and maintain patterns of methylated cytosine residues in the vertebrate genome. Because the only characterized component of the undoubtedly complex DNA methylating system is DNA methyltransferase itself, this enzyme will be the focus of attention.

2 Purification of mammalian DNA MTase

There is a long history of attempts to purify and characterize DNA (cytosine-5)-methyltransferase (DNA MTase) and numerous and often contradictory sizes and biochemical properties have been reported over the years (for a list of reported sizes, see Adams et al. (1990)). Recent purification and antibody studies have most frequently given an apparent M_r on SDS-polyacrylamide gel electrophoresis of around 190,000 for DNA MTase extracted from a number of proliferating human and murine cell types and tissues (Bestor and Ingram, 1985; Pfeifer and Drahovsky, 1986), including preimplantation mouse embryos (Howlett and Reik, 1991). DNA MTase is very sensitive to proteolysis, especially within the ~N-terminal 350 amino acids (Bestor, 1992), and smaller but enzymatically active cleavage products accumulate during purification. Proteolysis is presumably responsible for the smaller forms of DNA MTase that have been observed *in vivo* in non-dividing Friend murine erythroleukemia (MEL) cells, where a DNA MTase species of M_r 150,000 is found (Bestor and Ingram, 1985), and in full-term human placenta, where forms of DNA MTase of various smaller sizes have been identified (Pfeifer et al., 1985; Zucker et al., 1985). However, MEL cells have amplified the DNA MTase gene and express high levels of DNA MTase (Bestor et al., 1988), and full-term human placenta is an unusual non-proliferating tissue. At the present time it has not been

proven that forms of DNA MTase smaller than M_r 190,000 are not the result of proteolysis either *in vivo* or during purification. It is most likely that the sole or predominant form of DNA MTase in normal somatic tissues and proliferating cell types has an apparent M_r of 190,000. The open reading frame in the cloned DNA MTase cDNA yields a calculated mass for the primary translation product of about 170,000, and expression of the cloned cDNA in COS cells yields a protein of about this apparent size (Czank et al., 1991). This observation suggests that DNA MTase normally undergoes a post-translational modification in mouse cells which retards its rate of migration on SDS-polyacrylamide gels. The nature of the modification is not yet known.

3 Sequence and structure of DNA MTase

The cDNA for DNA MTase from murine erythroleukemia cells was cloned by means of a degenerate synthetic oligonucleotide probe whose sequence was based on the amino acid sequence of a fragment of the purified enzyme (Bestor et al., 1988). The cDNA sequence revealed that DNA MTase consists of a 1,000 amino acid N-terminal domain linked to a C-terminal domain of about 500 amino acids that is closely related to bacterial type II DNA C5 methyltransferases. About 30 of the bacterial enzymes have been sequenced, and all contain 10 conserved motifs in invariant order (Lauster et al., 1989; Posfai et al., 1989). For reasons that are not clear none of the known DNA C5 methyltransferases have recognition sequences of 6 bp. All also contain a variable region between conserved motifs VIII and IX (Fig. 1) which has been shown by mutagenesis experiments to confer sequence specificity to the transmethylation reaction (Klimasauskas et al., 1991; Lange et al., 1991). Figure 1 shows the organization of conserved motifs in the C-terminal domain of DNA MTase compared to M. *Ddel* (the most closely related bacterial enzyme; Szynter et al., 1987; Bestor et al., 1988) and M. *Sssl*, a *Spiroplasma* methyltransferase whose recognition sequence is the dinucleotide CpG (Renbaum et al., 1990). Despite the fact that DNA MTase and M. *Sssl* recognize the same DNA sequence, the variable region of DNA MTase is dissimilar in amino acid sequence and more than twice as large as that of M. *Sssl*, and in fact is the longest of the monospecific C5 DNA methyltransferases. As discussed elsewhere it is likely that mammalian DNA MTase is the result of fusion between genes for a prokaryotic-like restriction methyltransferase and an unrelated DNA binding protein (Bestor, 1990).

The C-terminal methyltransferase domain of DNA MTase is joined to the N-terminal domain by a run of 13 alternating glycyl and lysyl residues. In the center of the N-terminal domain is a cluster of 8 cysteinyl residues which has been shown to bind Zinc ions (Bestor,

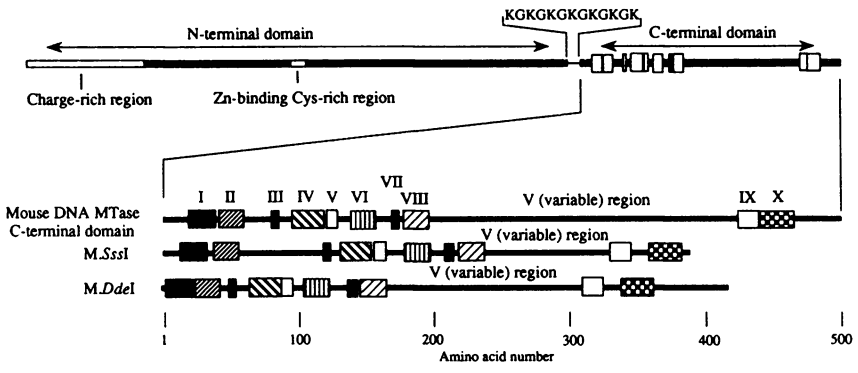

Figure 1. Sequence features and conserved motifs in mammalian DNA MTase. At top is a diagram of sequence features in DNA MTase; below is a depiction of elements in the C-terminal domain conserved between bacterial and mammalian DNA C5 methyltransferases. Boxes with common fill patterns indicate conserved motifs and are numbered I through X. Motif I is the putative S-adenosyl L-methionine binding site (Ingrosso et al., 1989), IV is the prolylcysteinyl active center (Wu and Santi, 1987; Chen et al., 1991), and the variable region is involved in sequence recognition (Lange et al., 1991; Klimasauskas et al., 1991). Note that the order of the conserved motifs is invariant and the variable region of DNA MTase is much longer than that of M. *SssI*, which recognizes the same sequence. M. *DdeI* methylates the cytosine residue in the sequence CTNAG.

1992). As described below, the N-terminal domain is involved in the discrimination of unmethylated and hemimethylated DNA, and the Zinc binding site is likely to be involved in this function. The first 200 amino acids of the N-terminal domain are very rich in charged and polar amino acids, and the first ~350 amino acids are very sensitive to proteolysis. Deletion of these sequences does not affect *in vitro* enzymatic activity or preference for hemimethylated sites (Bestor and Ingram, 1985).

4 *De novo* and maintenance methylation

Riggs (1975) and Holliday and Pugh (1975) predicted that vertebrate methylation patterns could be transmitted by clonal inheritance through the action of a DNA methyltransferase that was strongly stimulated by or dependent on hemimethylated DNA, which is the product of semiconservative DNA replication. This led to the expectation of two types of DNA methyltransferases: *de novo* enzymes, which would establish tissue-specific methylation patterns during gametogenesis and early development (in concert with a system that erased methylation patterns in the germline), and maintenance enzymes, which would ensure the clonal transmission of lineage-specific methylation patterns in somatic tissues. Razin and collaborators (Gruenbaum et al., 1982) showed that a DNA

MTase activity in extracts of somatic nuclei preferred hemimethylated substrates by a large factor, although *de novo* methylation was also observed. It was later shown that the *de novo* and maintenance activities reside in the same protein and that the preference for hemimethylated sites was 30–40 fold higher (Bestor and Ingram, 1983; Pfeifer et al., 1983; Bolden et al., 1984). Somatic cells do have the capacity to perform *de novo* methylation; methylation patterns are slowly restored after treatment with the demethylating drug 5-azacytidine (Flatau et al., 1984), and *de novo* methylation of the promoter regions of tissue-specific genes is observed in cells in long-term culture (Antequera et al., 1990). These findings confirm that *de novo* methylation is not confined to cells of the germline or early embryo, although *de novo* methylation of foreign DNA does appear to be much more efficient in embryonic cells (Jahner and Jaenisch, 1985). While the prediction of a distinct class of *de novo* DNA methyltransferases has not been confirmed, the existence of such enzymes cannot yet be excluded. It should soon be possible to answer the question definitively through use of a sensitive, versatile, and highly specific probe for DNA C5 methyltransferases recently introduced by Gregory Verdine's laboratory (Chen et al., 1991). Oligonucleotides containing the modified nucleoside 5-fluorodeoxycytidine (FdC) have been shown to trap a covalent transition state intermediate between DNA and DNA methyltransferases in a form that is stable to strong denaturing conditions, as predicted by Santi et al. (1983). If the FdC-containing oligonucleotide is radioactive, the covalent complexes with DNA methyltransferases can be visualized by autoradiography after electrophoresis on SDS-polyacrylamide gels. This mechanism-based probe and inhibitor should provide sub-femtomol sensitivity, and it will be possible to test lysates of cell populations in which *de novo* methylation are occurring (especially germ cells and cells of the preimplantation embryo) for species of DNA methyltransferase distinct from the known M_r 190,000 form. Immobilization of the FdC-containing oligonucleotides on a solid support should allow rapid purification of any new species, and amino acid sequencing of proteins purified in this way will allow cloning.

5 Discrimination of hemimethylated and unmethylated CpG sites

Bacterial and mammalian DNA methyltransferases differ most markedly in that the type II bacterial enzymes do not discriminate between hemimethylated and unmethylated recognition sequences. Adams and colleagues (Adams et al., 1983) observed an increased rate of *de novo* methylation after treatment of a crude DNA MTase preparation with trypsin and concluded that the enzyme must contain a protease-sensitive domain that makes contacts with the C5 methyl group of hemimethyl-

ated sites. In double-stranded B form DNA the C5 positions of cytosine residues in CpG sites are separated by only a few Ångstroms in the major groove, and analysis of bacterial restriction methyltransferases have suggested that at least 3 regions of the protein must be very close to the target cytosine (Fig. 1); these are the S-adenosyl L-methionine binding site near the N-terminus (Ingrosso et al., 1989), the prolylcysteinyl dipeptide at the catalytic center (Wu and Santi, 1987), and a region near the C-terminus that mediates sequence-specific DNA binding (Lange et al., 1991; Klimasaukas et al., 1991). All these regions are within the C-terminal domain of mammalian DNA MTase. While contacts between the methyl group and any of these motifs might be expected to mediate discrimination of unmethylated and hemimethylated sites, the results of recent proteolysis experiments indicate that the discrimination is carried out by distant sequences in the N-terminal domain of DNA MTase. Protease V8 cleaves DNA MTase between the N- and C-terminal domains, as shown by microsequencing of fragments. Cleavage caused a large stimulation in the rate of *de novo* methylation without significant change in the rate of methylation of hemimethylated DNA; this demonstrates that the N-terminal domain inhibits the *de novo* activity of the C-terminal methyltransferase domain (Bestor, 1992). The finding was unexpected, as the close proximity of the methyl group in a hemimethylated CpG site to the C5 position of the target cytosine imposes severe steric constraints and it seems unlikely that an additional protein structural element could be accommodated near the target cytosine in the major groove. This and other lines of evidence (Bestor, 1992) lead to the conclusion that it is methylation-dependent structural alterations in DNA, rather than direct contact of the protein with major groove methyl groups, that is responsible for discrimination of unmethylated and hemimethylated CpG sites. This conclusion is not without precedent; DNase I preferentially cleaves methylated CpG sites (Fox, 1986), and yet this enzyme makes contacts only in the minor groove of B form DNA (Lahm and Suck, 1991). DNase I must therefore sense cytosine methylation indirectly through alterations of DNA structure rather than via direct major groove contacts. However, the physical separation between the catalytic and regulatory regions of DNA MTase suggests that the mechanism used by DNA MTase in the discrimination of unmethylated and hemimethylated CpG sites is fundamentally different than any known type of DNA:protein interaction.

Cleavage between the N- and C-terminal domains stimulates *de novo* methylation, and because most purification schemes measure *de novo* activity in assays, the purification method which gives the best apparent yield will be that which most favors proteolysis. The sensitivity of DNA MTase to proteolysis and the fact that most biochemical characterization of the enzyme has involved partially purified enzyme preparations

with unknown extents of proteolysis is likely to be part of the cause for the wide range of enzymatic properties ascribed to DNA MTase.

6 *De novo* sequence specificity

Little is known of how sequence-specific methylation patterns are established in the mammalian genome. The sequence specificity of purified DNA MTase does not extend much past the CpG dinucleotide (Gruenbaum et al., 1981; Simon et al., 1983; Hubrich et al., 1989; Bestor and Ingram, 1983), and cell types with different methylation patterns contain species of DNA MTase that are identical by all criteria, including *de novo* sequence specificity (Bestor et al., 1988).

There are several candidate mechanisms for sequence-specific methylation. First, as mentioned earlier it is possible that tissue- and sequence-specific *de novo* methyltransferases are expressed at specific stages of development and that the altered methylation patterns are maintained in somatic tissues through the maintenance activity of the known form of DNA MTase. While there is no evidence for a family of DNA methyltransferases, their existence remains a possibility. The sensitive and versatile FdC-oligonucleotide probes described earlier should provide an answer to the question of multiple species of DNA methyltransferases in mammals. Second, *de novo* methylation may be relatively indiscriminate during certain stages of development, and critical CpG sites might be protected from methylation by sequence-specific masking proteins. At such times the *de novo* activity of DNA MTase might be stimulated by proteolytic cleavage between the N- and C-terminal domains or interaction with a factor which counteracts the inhibitory effects of the N-terminal domain. The masking model cannot be looked on with much favor, as it is precisely the unmethylated CpG sites which are accessible to diffusible factors in nuclei (Antequera et al., 1989), and genomic sequencing has not shown a bias in the sequences flanking methylated and unmethylated CpG sites (Jost et al., 1990). Sequence-specific masking proteins would be expected to leave some evidence of a consensus sequence around unmethylated CpG sites. Third, a family of specificity factors, analogous to the specificity subunits of bacterial type I restriction-modification systems, might interact with DNA MTase to confer sequence specificity while enhancing *de novo* methylation activity. This possibility suffers the same problem as the masking proteins: there is no evidence of a consensus sequence around methylated or unmethylated CpG sites. Furthermore, proteins that interact strongly with DNA MTase have not been identified. Fourth, it is possible that *de novo* methylation is indiscriminate and that tissue-specific methylation patterns are established by sequence-specific demethylation. Sequence-specific demethylation, presumably through a

mechanism related to excision repair, has been documented in the case of the chicken vitellogenin gene (Jost et al., 1990) and could be widespread. It is sobering to recognize that at the present time it is not known whether tissue-specific methylation patterns are established by sequence-specific *de novo* methylation, by indiscriminate *de novo* methylation and sequence-specific demethylation, or by some combination of the two.

7 Targeted disruption of the DNA MTase gene in mice and in mouse cells

The regulatory role of DNA methylation remains controversial, in large part because reversible, tissue-specific methylation patterns are restricted to large-genome organisms such as vertebrates and vascular plants in which genetic approaches are limited. It has recently become possible to introduce predetermined mutations in any mouse gene for which cloned probes are available by gene targeting in embryonic stem (ES) cells (Mansour et al., 1988). This approach has been used to disrupt both alleles of the DNA MTase gene in ES cells with a construct which introduces a short deletion-replacement at the translational start site (Li et al., 1992). The mutation is a partial loss of function allele which produces trace amounts of a slightly smaller protein, as established by gel electrophoresis and immunoblotting. Net enzyme activity *in vitro* assays is about 5% of wildtype. This is limiting, and the homozygous mutant ES cells and embryos have about one-third of the wildtype level of 5-mC in their DNA. The homozygous mutant ES cells show no discernible phenotype even after prolonged passage *in vitro*. The mutation has also been established in the germline of mice. Homozygous mutant embryos complete gastrulation and the early stages of organogenesis but are stunted, delayed in developmental stage, and fail to develop past the 20 somite stage. Histological analysis shows that many cells in the mutant embryos contain fragmented, pycnotic nuclei which are typical of apoptosis rather than necrosis. It was interesting to find that reduced 5-mC levels are lethal at the stage where normal embryos attain adult levels of 5-mC in their DNA (Monk, 1990). In addition to the partial loss of function mutation, a second independent mutation was constructed by means of a targeted insertion mutation in sequences downstream of the region targeted by the first construct. This presumptive severe loss of function mutation causes homozygous embryos to die at earlier stages and to have less 5-mC in their DNA than does the partial loss of function mutation, and embryos with one copy each of the partial and severe loss of function mutation die at intermediate stages and have intermediate levels of 5-mC in their DNA.

Embryos homozygous for the partial loss of function mutation retain $\sim 1 \times 10^7$ methylated CpG sites per genome, one-third of the wildtype

level. This finding shows that even fairly modest reductions in 5-mC content which have no apparent effect on the phenotype of cultured ES cells completely prevent normal development past midgestation. The cause of the developmental block is not known, but an attractive and testable hypothesis is inappropriate gene expression as a result of the activation of genes that are normally repressed by methylation.

Embryos homozygous for the partial loss of function mutation complete gastrulation and the early stages of organogenesis. They will therefore serve as a robust test system for hypotheses regarding the importance of DNA modification in developmental gene control, X inactivation, genomic imprinting, virus latency, and other biological phenomena in which DNA methylation has been proposed to play a role.

Acknowledgments. Supported by the National Institutes of Health and the March of Dimes Birth Defects Foundation.

Adams, R. L., Burdon, R. H., McKinnon, K., and Rinaldi, A. (1983) Stimulation of *de novo* methylation following limited proteolysis of mouse ascites DNA methylase. FEBS Lett. *163*, 194–198.

Adams, R. L. P., Bryans, M., Rinaldi, A., Smart, A., and Yesufu, H. M. I. (1990) Eukaryotic DNA methylases and their use for *in vitro* methylation. Phil. Trans. R. Soc. Lond. B *326*, 11–21.

Antequera, F., Boyes, J., and Bird, A. (1990) High levels of *de novo* methylation and altered chromatin structure at CpG islands in cell lines. Cell *62*, 503–14.

Antequera, F., Macleod, D., and Bird, A. P. (1989) Specific protection of methylated CpGs in mammalian nuclei. Cell *58*, 509–517.

Bestor, T. (1990) DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher eukaryotes. Phil. Trans. R. Soc. Lond. B *326*, 179–187.

Bestor, T. (1992) Activation of mammalian DNA methyltransferase by cleavage of a Zinc-binding regulatory domain. EMBO J., *11*, 2611–2617.

Bestor, T., Hellewell, S. B., and Ingram, V. M. (1984) Differentiation of two mouse cell lines is accompanied by demethylation of their genomes. Mol. Cell. Biol. *4*, 1800–1806.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J. Mol. Biol. *203*, 971–83.

Bestor, T. H., and Ingram, V. M. (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. Proc. Natl. Acad. Sci. USA *80*, 5559–63.

Bestor, T. H., and Ingram, V. M. (1985) Growth-dependent expression of multiple species of DNA methyltransferase in murine erythroleukemia cells. Proc. Natl. Acad. Sci. USA *82*, 2674–8.

Bolden, A., Ward, C., Siedlecki, J. A., and Weissbach, A. (1984) DNA methylation. Inhibition of *de novo* and maintenance methylation *in vitro* by RNA and synthetic polynucleotides. J. Biol. Chem. *259*, 12437–43.

Chaillet, J. R., Vogt, T. F., Beier, D. R., and Leder, P. (1991) Parental-specific methylation of an imprinted transgene is established during gametogenesis and progressively changes during embryogenesis. Cell *66*, 77–83.

Chen, L., MacMillan, A. M., Chang, W., Ezaz, N. K., Lane, W. S., and Verdine, G. L. (1991) Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. Biochemistry *30*, 11018–25.

Czank, A., Häuselman, R., Page, A. W., Leonhardt, H., Bestor, T. H, Schaffner, W., and Hergersberg, M. (1991) Expression in mammalian cells of a cloned gene encoding murine DNA methyltransferase. Gene *109*, 259–263.

Engler, P., Haasch, D., Pinkert, C. A., Doglio, L., Glymour, M., Brinster, R., and Storb, U. (1991) A strain-specific modifier on mouse chromosome 4 controls the methylation of independent transgene loci. Cell *65*, 939–947.

Flatau, E., Gonzales, F. A., Michalowsky, L. A., and Jones, P. A. (1984) DNA methylation in 5-aza-2′-deoxycytidine-resistant variants of C3H 10T1/2 C18 cells. Mol. Cell. Biol. *4*, 2098–102.

Fox, K. R. (1986) The effect of HhaI methylation on DNA local structure. Biochem. J. *234*, 213–216.

Gruenbaum, Y., Stein, R., Cedar, H., and Razin, A. (1981) Methylation of CpG sequences in eukaryotic DNA. FEBS Lett. *124*, 67–71.

Gruenbaum, Y., Cedar, H., and Razin, A. (1982) Substrate and sequence specificity of a eukaryotic DNA methylase. Nature *295*, 620–622.

Holliday, R., and Pugh, J. (1975) DNA modification mechanisms and gene activity during development. Science *187*, 226–232.

Howlett, S. K., and Reik, W. (1991) Methylation levels of maternal and paternal genomes during preimplantation development. Development *113*, 119–127.

Hubrich, K. K., Buhk, H. J., Wagner, H., Kroger, H., and Simon, D. (1989) Non-C-G recognition sequences of DNA cytosine-5-methyltransferase from rat liver. Biochem. Biophys. Res. Commun. *160*, 1175–1182.

Ingrosso, D., Fowler, A. V., Bleibaum, J., and Clarke, S. (1989) Sequence of the D-aspartyl/L-isoaspartyl protein methyltransferase from human erythrocytes. Common sequence motifs for protein, DNA, RNA, and small molecule S-adenosylmethionine-dependent methyltransferases. J. Biol. Chem. *264*, 20131–20139.

Klimasauskas, S., Nelson, J. L., and Roberts, R. J. (1991) The methylase specificity domain of cytosine C5 methylases. Nucl. Acids Res. *19*, 6183–6190.

Jahner, D., and Jaenisch, R. (1985) Chromosomal position and specific demethylation in enhancer sequences of germ line-transmitted retroviral genomes during mouse development. Mol. Cell. Biol. *5*, 2212–2220.

Jost, J.-P., Saluz, H.-P., McEwan, I., Feavers, I. M., Hughes, M., Reiber, S., Liang, H. M., and Vaccaro, M. (1990) Tissue specific expression of avian vitellogenin gene is correlated with DNA hypomethylation and *in vivo* specific protein-DNA interactions. Phil. Trans. Roy. Soc. Lond. B *326*, 53–63.

Lahm, A., and Suck, D. (1991) DNase I-induced DNA conformation. J. Mol. Biol. *221*, 645–667.

Lange, C., Jugel, A., Walter, J., Noyer, W. M., and Trautner, T. A. (1991) 'Pseudo' domains in phage-encoded DNA methyltransferases. Nature *352*, 645–648.

Lauster, R., Trautner, T. A., and Noyer-Weidner, M. (1989) Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains. J. Mol. Biol. *206*, 305–312.

Li, E., Bestor, T. H., and Jaenisch, R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915–926.

Mansour, S. L., Thomas, K. R., and Capecchi, M. R. (1988) Disruption of the protooncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature *336*, 348–352.

Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L., and Bird, A. P. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell *58*, 499–507.

Monk, M. (1990). Changes in DNA methylation during mouse embryonic development in relation to X-chromosome activity and imprinting. Phil. Trans. Roy. Soc. Lond. B. *326*, 179–187.

Monk, M., Boubelik, M., and Lehnert, S. (1987) Temporal and regional changes in DNA methylation in the embryonic, extraembryonic, and germ cell lineages during mouse embryo development. Development *99*, 371–382.

Pfeifer, G. P., and Drahovsky, D. (1986) DNA methyltransferase polypeptides in mouse and human cells. Biochim. Biophys. Acta *868*, 238–242.

Pfeifer, G. P., Grunwald, S., Palitti, F., Kaul, S., Boehm, T. L., Hirth, H. P., and Drahovsky, D. (1985) Purification and characterization of mammalian DNA methyltransferases by use of monoclonal antibodies. J. Biol. Chem. *260*, 13787–13793.

Pfeifer, G. P., Grunwald, S., Boehm, T. L., and Drahovsky, D. (1983) Isolation and characterization of DNA cytosine 5-methyltransferase from human placenta. Biochim. Biophys. Acta *740*, 323–330.

Posfai, J., Bhagwat, A. S., Posfai, G., and Roberts, R. J. (1989) Predictive motifs derived from cytosine methyltransferases. Nucl. Acids Res. *17*, 2421–2435.

Renbaum, P., Abrahamove, D., Fainsod, A., Wilson, G. G., Rottem, S., and Razin, A. (1990) Cloning, characterization, and expression in Escherichia coli of the gene coding for the CpG DNA methylase from Spiroplasma sp. strain MQ1 (M. *Sssl*). Nucl. Acids Res. *18*, 1145–1152.

Riggs, A. D. (1975) X inactivation, differentiation, and DNA methylation. Cytogenet. Cell Genet. *14*, 9–25.

Santi, D. V., Garrett, C. E., and Barr, P. J. (1983) On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. Cell *33*, 9–10.

Simon, D., Stuhlmann, H., Jahner, D., Wagner, H., Werner, E., and Jaenisch, R. (1983) Retrovirus genomes methylated by mammalian but not bacterial methylase are non-infectious. Nature *304*, 275–277.

Schwartz, M. N., Trautner, T. A., and Kornberg, A. (1962) Enzymatic synthesis of DNA. J. Biol. Chem. *237*, 1961–1967.

Sznyter, L. A., Slatko, B., Moran, L., O'Donnell, K. H., and Brooks, J. E. (1987) Nucleotide sequence of the DdeI restriction-modification system and characterization of the methylase protein. Nucl. Acids Res. *15*, 1846–1866.

Uehara, Y., Ono, T., Kurishita, A., Kokuryu, H., and Okada, S. (1989) Age-dependent and tissue-specific changes of DNA methylation within and around the c-*fos* gene in mice. Oncogene *4*, 1023–1028.

Wu, J. C., and Santi, D. V. (1987) Kinetic and catalytic mechanism of HhaI methyltransferase. J. Biol. Chem. *262*, 4778–4786.

Zucker, K. E., Riggs, A. D., and Smith, S. S. (1985) Purification of human DNA (cytosine-5)-methyltransferase. J. Cell. Biochem. *29*, 337–349.