Commentary

DNA demethylation

Alan P. Wolffe*, Peter L. Jones, and Paul A. Wade

Laboratory of Molecular Embryology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892-5431

Cytosine 5’ methylation of CpG dinucleotides within and around genes exerts a major influence on transcription in many plants and animals (1–5). DNA methylation can be causal for transcriptional silencing (4–5) and targets the machinery necessary to assemble specialized chromatin enriched in deacetylated histones (6–8). Once established in somatic cells, CpG methylation patterns within the genome are very stable and provide an attractive mechanism for segregating a large fraction of stably repressed chromatin (9–11). In contrast, DNA methylation is remarkably dynamic during early mammalian development and in certain tumor cells (12, 13). Alterations in the methylation status of the entire genome (14, 15), individual chromosomes (16), and specific genes (17–20) are essential for normal development (21, 22) and can promote tumorigenesis (23, 24). Understanding how these important transitions might be regulated requires the biochemical definition of the enzymatic processes that both methylate and demethylate the genome.

Two mammalian DNA methyltransferases have been functionally defined (25, 26) from a family of related proteins (26, 27). Dnmt1 is essential for inactivation of the X chromosome and genomic imprinting in the mammalian embryo (21). This large enzyme (1,620 amino acids) is targeted to replication foci consistent with the rapid remethylation of DNA in somatic cells (28–30). Like the prokaryotic cytokine-5 methyltransferases with which it shares homology, the enzyme makes use of the Michael addition mechanism to carry out the reaction by first increasing the reactivity of the C-5 position of cytosine (31–33). An enzyme cysteinyl thiolate forms a covalent linkage with C-6 of cytosine, and the carboxyl group of an invariant glutamyl residue protonates the N-3 position to create a cytosine 4, 5 enamine. This reactive moiety attacks the sulfonium linked methyl group of S-adenosyl L-methionine. After transfer of the methyl group, the proton is abstracted from the C-5 of cytosine to reform the 5, 6 double bond and release the enzyme by β-elimination (34, 35). In contrast to the well-defined molecular genetics, cell biology, and biochemistry of the Dnmt1 methyltransferase, the enzymatic basis of demethylation of 5-methylcytosine in vivo has been mysterious. Although strategies for demethylating DNA have been proposed (Figs. 1 and 2), none have yet been proven to operate under relevant physiological circumstances in vivo. This issue has now been brought into sharper focus by the recent reports of a mammalian protein with specific demethylation activity for methyl CpG dinucleotides (36), together with a demethylase enzyme complex that acts processively (37) and as published in this issue of Proceedings the finding that this complex converts 5-methylcytosine to cytosine and methanol (38).

A Demethylase?

The “amazing demethylase,” as it was recently described by Cedar and Verdine (ref. 39, p. 579), is a member of a conserved family of MBD (methyl CpG-binding domain) proteins (36, 40) related to the MeCP2 transcriptional repressor (41). The demethylase enzyme is identical to MBD2b (36, 40), and a longer MBD2a variant of unknown enzymatic activity also exists. MBD2a is found in vivo whereas there is so far no evidence for naturally occurring MBD2b. The small basic MBD2b protein (262 amino acids, pI 9.1) shows selectivity (10–100 fold) for binding to DNA containing methylated CpG dinucleotides (40). As described, the demethylase has remarkable biochemistry and enzymatic activities. The enzyme was first identified by empirical testing for 5-methylcytosine demethylase activity after the in vitro translation of mRNA derived from an expressed sequence tag cDNA with homology to MeCP2 (36). The polyhistidine-tagged enzyme retained activity after purification on a Ni2+ -charged agarose resin, and transfection of human embryonal kidney cells with the MBD2b expressing clones generated detectable demethylase that fractionated on sucrose gradients in the 160–190 K range (36). It is suggested that the MBD2b protein might multimerize (36). Although the molecular identity of a demethylase would appear to be clearly established by these experiments as MBD2b, the demethylase enzyme complex used for mechanistic studies has been fractionated from human cells and is not defined at a molecular level (36–38). Thus, the relationship of this cellular demethylase complex to MBD2b remains unclear. An additional complication is the highly unusual biochemistry of the cellular enzyme fraction compared with the known properties of MBD2b. The initial purification step, gradient elution from the weak ammonium matrix DEAE-Sephadex A-50, provides 850-fold enrichment with activity reported at 4.9–5.0 M NaCl. Two subsequent chromatographic steps result in a further 75-fold purification by successive elution (also at 4.8–5.0 M NaCl) from both a strong cation exchanger (SP-Sepharose) and a strong anion exchanger (Q-Sepharose). Remarkably, a further 10-fold enrichment is achieved by applying the 4.8–5.0 M NaCl pool from Q-Sepharose, without prior reduction of salt concentration, to DEAE-Sephasel and eluting with 10 mM TrisHCl (pH 7.5) and 10 mM MgCl2. A silver stained SDS polyacrylamide gel description coelution of activity with three polypeptides of 38–40 kDa, clearly distinct from the known electrophoretic migration of recombinant MBD2b (29 kDa; ref. 38). The repeated elution of demethylase at 4.8–5.0 M NaCl from both strong and weak anion exchangers as well as a strong cation exchanger differ substantially from the chromatographic properties of recombinant MBD2b and from the archetypical MBD family member MeCP2 (40, 41). These discrepancies remain to be resolved; it is conceivable that MBD2b might be post-translationally modified in vivo or that it is associated with a cofactor that substantially alters chromatographic properties. Whatever the precise nature of the demethylase, the 500,000-fold purification of this cellular enzyme allows effective comparison with previously defined “demethylase” activities (36–38).

Distinct Pathways for “Active” Demethylation

The chemistry necessary to directly demethylate 5-methylcytosine is challenging, requiring the disruption of a C-C bond.

The companion to this Commentary begins on page 6107.

*To whom reprint requests should be addressed. e-mail: awlme@helix.nih.gov.

PNAS is available online at www.pnas.org.
Earlier work had demonstrated that 5-methylcytosine was replaced by labeled cytosine during the demethylation reaction in erythroleukemia cells, indicative of replacement of the entire nucleotide or base alone (42). One potential mechanism to achieve this aim is through the action of 5-methylcytosine DNA glycosylase, which removes the methylated cytosine from DNA, leaving the deoxyribose intact (ref. 43 and Fig. 1). Local DNA repair then eventually adds back the cytosine in nucleotide form. It has been suggested that targeting of the glycosylase requires an RNA moiety (44). A more active role for RNA as an acceptor of the entire 5-methylcytosine nucleotide in an in vitro demethylation assay has been suggested (45) but later reevaluated (46). The demethylase enzyme complex characterized in the most recent series of papers lacks DNA glycosylase or nuclease activity (36), is resistant to RNase (38), and hydrolyses 5-methylcytosine to cytosine and methanol (38). Therefore, the novel 5-methylcytosine demethylase enzyme complex appears distinct from previously characterized activities (Fig. 1). This confusing situation will only be clarified by the molecular cloning and characterization of the genes encoding the demethylase present in these extracts after their biochemical purification. Knockout experiments should eventually establish whether a particular protein is important for the demethylation of DNA sequences that occurs during development or in tumor cells.

Meanwhile, if MBD2b proves to be the key demethylase and other unidentified activities in the in vitro translation extract or transfected cell are not contributing to demethylation, a whole new chemical problem requires analysis. This hydrolysis of 5-methylcytosine to cytosine and methanol will have a very

**Fig. 1.** Active DNA demethylation. Two pathways are illustrated. (**A**) The 5-methylcytosine demethylase hydrolyzes 5-methylcytosine to cytosine and water (38). (**B**) The 5-methylcytosine DNA glycosylase abstracts 5-methylcytosine from the phosphodiester backbone, which is then repaired by using endonuclease (43, 44).

**Fig. 2.** Replication-coupled DNA demethylation. (**A**) Normally, each strand of a symmetrically methylated CpG dinucleotide (M) would be segregated after replication daughter chromatids, leaving one strand of DNA methylated (M) and the other not (●). The maintenance methyltransferase activity of Dnmt1 then would restore symmetrical methylation. (**B**) Regulatory nucleoprotein complexes might occlude Dnmt1 leading to loss of methylation (48, 49). (**C**) Histone acetylation might repel or inhibit Dnmt1, leading to loss of methylation (47).
large activation energy because O-H and C-C bonds have to be broken to allow the formation of C-O and C-H bonds. Cedar and Verdine (39) suggest that demethylation by water should be thermodynamically possible and that comparable molecular events to those occurring during DNA methylation would facilitate the process. This remains to be shown. As noted by the authors (36), their results “identify a new enzyme and biochemical reaction that have not yet been described in any organism” (p. 583).

Replication-Coupled Demethylation

Aside from the evidence for active demethylation by dedicated demethylase enzymes and use of the DNA repair pathway, it is probable that DNA replication also will contribute to eliminating 5-methylcytosine from DNA. We have described the targeting of the Dnmt1 methyltransferase to the replication fork and the rapid maintenance methylation of CpG dinucleotides (28–30). Chromatin structures and nucleoprotein complexes associated with transcriptional activity might interfere with the maintenance methyltransferase activity of Dnmt1, thereby progressively reducing the level of DNA methylation with each cell division event. Examples of this phenomenon include the selective loss of DNA methylation in *Neurospora* after growth of cells in Trichostatin A, an inhibitor of histone deacetylase (47), and the requirements for protein association during DNA replication to demethylate specific DNA sequences (48, 49). As illustrated in Fig. 2, during replication, both strands of DNA making up a symmetrically methylated CpG dinucleotide in the parental chromosome will be segregated to daughter chromatids. This leads both daughter chromatids to contain hemimethylated DNA. Under normal circumstances, the maintenance methyltransferase activity of Dnmt1 will restore symmetric methylation. It is possible that sequence-specific transacting factors bound over the sites of DNA methylation might prevent access of the methyltransferase enzyme to these sites, leading to the progressive demethylation of DNA, potentially in a sequence specific manner. With respect to more global genome demethylation, particular histone modifications such as acetylation might exert a “teflon effect,” preventing efficient access of the methyltransferase enzyme. These models predict that DNA replication will be an integral component of the process.

Significance

DNA methylation is a major force in mammalian development and tumorigenesis. It provides a defense mechanism against the expression of exogenous DNA in plants and animals that constrains the expression of transgenes (50–52) and, when misdirected, contributes to tumorigenesis (53). Understanding the biochemistry of DNA methylation dynamics is a problem of tremendous biomedical importance. To achieve this goal, the characterization of the molecular mechanisms that target both DNA methylation and demethylation is essential. Such knowledge should enable the manipulation of these events effectively in the hope of achieving the most effective expression of exogenous DNA. DNA methylation dynamics present a mystery that needs to be solved.

We thank Adrian Bird, Howard Cedar, Brian Hendrich, Marylin Monk, and Susan Wallace for useful discussions and Thuy Vo for manuscript preparation.