Evidence that Adenine Methylation Influences DNA-Protein Interactions in *Escherichia coli*

NAT STERNBERG

Central Research and Development Department Experimental Station, E. I. Du Pont de Nemours and Co., Inc., Wilmington, Delaware 19898

In this review, most of the information presented will be derived from studies with *Escherichia coli* K-12 (E. coli) and its related bacteriophages, simply because more is known about methylation in these organisms than in any other. The two methylated bases that have been detected in *E. coli* are 6-methyladenine (6-meAde) and 5-methylcytosine (5-meC). Since little is known about the biological function of 6-methylcytosine, I will deal exclusively here with the studies on 6-meAde.

6-meAde is the product of a reaction catalyzed by a methylase coded for by the bacterial dam gene (21). In the absence of that methylase, less than 1% of the 6-meAde normally found in DNA remains (2, 21, 22). The specific sequence recognized by the methylase is 5'-GATC-3' (4, 9, 14). Dam− mutants exhibit a variety of phenotypes that include increased spontaneous mutability (5, 23), increased sensitivity to agents such as UV light or 2-aminopurine (22), hyperrecombination (20), and increased transposon-mediated rearrangement of DNA (13; D. Roberts and N. Kleckner, unpublished communication). A lack of 6-meAde has also been correlated with either an increase or decrease in the expression of genes (8, 12, 19; N. Sternberg, unpublished data), a decreased replication of replicons (24, 30), and a decreased packaging of viral DNAs (31; N. Sternberg, unpublished data).

In Dam− cells, both strands of the DNA are unmethylated. In normal cells, however, this situation rarely exists, because the old template strand in a replicating DNA duplex is already methylated, and the newly synthesized DNA strand is methylated soon after its replication (16, 17, 32). Thus, whatever effects are attributed to the absence of 6-meAde in DNA, they must be manifested during the short period of time when the newly synthesized DNA strand is transiently unmethylated, namely, when the DNA is hemimethylated. A general idea that appears to be emerging from studies on the biological role of 6-meAde is that gene expression can be coupled to DNA replication and cell division if the gene is transcribed poorly from fully methylated DNA and efficiently from hemimethylated DNA. Examples of biological systems in which gene expression and protein function are affected by adenine methylation are described below. I will not discuss in this review the extensive literature that deals with the role of Dam methylation in the repair of mismatched DNA. Reviews that discuss this subject have recently appeared or are about to appear in the literature (18; M. Marinus, personal communication).

The SOS response. When the DNA metabolism of *E. coli* is perturbed, the cells respond by inducing a set of genes, collectively referred to as the SOS regulon, whose function is necessary for cell survival under those stressful conditions (15, 33). An important step in that induction process is the inactivation of a repressor of the SOS regulon that is encoded by the *lexA* gene. *lexA* inactivation requires a functional *recA* gene. K. R. Peterson, K. F. Wettman, D. W. Mount, and M. G. Marinus have recently found that the genes of the SOS regulon could be induced (1.7- to 6-fold) by introducing a dam mutation into *E. coli*. The affected genes include *recA*, *lexA*, *uvrA*, *uvrB*, *suLA*, *dinD*, and *dinF*. The only SOS gene not affected is *uvrD*. Except for *lexA* and *uvrB*, none of the genes are fully induced. In essence, the cell adjusts the basal level of the SOS response to that needed without fully inducing the regulon. In this case, the control of gene expression is probably an indirect effect of the methylation defect. Presumably, the absence of methylation produces a cascade of events that includes random mismatch repair, increased single-stranded breaks, inactivation of *lexA* repressor, and SOS induction. Moreover, since *recA* dams, *lexA* dam, and *uvr* dam double mutants are not viable (21, 22), it would appear that the induction of at least one SOS regulon gene (the *rav* gene) is essential for the viability of dam mutants.

Transposon function. The most extensively studied case of Dam-mediated gene regulation involves the transposon Tn10. Tn10-mediated transposition is elevated 10-fold, and Tn10-mediated deletion or inversion is elevated 100- to 250-fold in a Dam− cell (13; D. Roberts and N. Kleckner, personal communication). Other transposons which show a 5- to 10-fold elevation in transposition in Dam− cells are Tn903 (13), Tn5 (13; J. Yin and W. Reznikoff, personal communication), and IS50 (J. Yin and W. Reznikoff, personal communication). The Dam methylase appears to operate at two levels to repress transposon function (D. Roberts, N. Kleckner, J. Yin, and W. Reznikoff, personal communication). First, methylation of critical Dam sites in the −10 region of the promoter for the transposase gene reduces transcription of the gene and also the amount of transposase made, by about 5- to 10-fold. Second, methylation of Dam sites at the inside ends of insertion elements (IS50 for Tn10 and IS50 for Tn5) decreases the ability of the transposase to utilize those ends for transposition or rearrangement. The more pronounced effect of the Dam− phenotype on Tn10-mediated inversion or deletion, compared with transposition, reflects the preferential use of the inside end of the insertion element for the former class of rearrangement.

The promoter sequences of the Tn10 and Tn5 transposase genes are shown in Fig. 1A (6; J. Yin and W. Reznikoff, personal communication). Both promoters contain a 6-meAde nucleotide just 5' to the −10 region. Tn5 also has a second 6-meAde within the −10 region. The importance of these nucleotides is indicated by a Tn10 mutation that converts the G residue upstream of the −10 region to a A (Fig. 1A). This mutation (DR33) eliminates the Dam methylation site and renders transcription mediated by this promoter both in vivo and in vitro, insensitive to Dam methylation (D. Roberts, N. Kleckner, B. Hoopes, and W. McClure, personal communication). The sequences of the
inside ends of IS10 and IS50 are shown in Fig. 1B (3, 6). Both sequences contain Dam sites in the regions that are thought to interact with the transposase. In an experiment in which the Tn10 transposase was provided by a Dam-sensitive promoter, transposition of elements utilizing the inside end of IS10 was shown to be elevated 6- to 10-fold in a Dam⁻ host (D. Roberts and N. Kleckner, personal communication). This effect was eliminated by a mutation that converts the GATC sequence at the inside end of IS10 to GATT (DR7; Fig. 1B).

In Dam⁺ cells, the stimulation of transposition produced when the DNA is undermethyalted most probably occurs only during the short period after the transposon has been replicated and before the newly replicated DNA strand has been methylated. If this is true, one might expect to find that undermethylation of one of the two DNA strands is more critical than that of the other. These predictions were confirmed in experiments carried out by D. Roberts and N. Kleckner (personal communication). DNA in which only one of the two strands is methylated (hemimethylated DNA) was generated in vivo by single-stranded Hfr transfer followed by complementary strand synthesis in the recipient cell. Transfer of two IS10 elements inserted at the same site, but in opposite orientations, results in hemimethylated IS10 elements which are methylated on opposite strands. Transposition from the newly transferred and replicated DNA was measured. A 10-fold difference was observed in the Dam⁺ recipient for the two orientations. This difference was enhanced only an additional twofold in the Dam⁻ recipient. These results argue that most of the transposition events occurred during the hemimethylated state created by the DNA transfer process. This conclusion is supported by the observation that a similar orientation difference was not observed with an isogenic mutant of IS10 that lacked the GATC sites (Tn10 DR33 and DR7).

The single-strand transfer experiment supports the contention that transposition is modulated by the natural loss of full methylation after replication of the transposon. After replication, there will be a short interval in which the two hemimethylated copies of the transposon exist. One of these will be activated for transposition until methylation occurs, at which point transposition will return to a low level. For this model to be valid, the transposase or its ability to act must be limiting. This is, in fact, the case. It is known that the Tn10 transposase promoter is weak and that transposition can be increased by replacing it with a stronger promoter (26).

Other Dam-sensitive genes. The following four genes have been identified that also show a sensitivity to Dam methylation: E. coli trpR (M. Marinus, personal communication), E. coli sulA (25; S. Gottesman, personal communication), bacteriophage P1 cre (unpublished data), and bacteriophage Mu mom (8, 10, 12, 28). The first three genes contain Dam methylation sites in the −35 region of the promoters (Fig. 1A). Their transcription is elevated two- to sixfold in a Dam⁻ host compared with the isogenic Dam⁺ host. For sulA, part of the induction (two- to threefold) is indirect and is due to partial SOS induction, but the remainder (two- to threefold) is due to a direct effect of methylation at the promoter (S. Gottesman and M. Marinus, personal communication). While the critical experiments have not yet been done, the activation of these three promoters, like that of the transposase genes, probably involves the production by replication of hemimethylated promoter DNA. In this respect, Marinus (personal communication) has argued that the location of the tryptophan operon repressor gene (trpR), between the bacterial origin of replication and the tryptophan operon, ensures a burst of repressor synthesis before the operon is replicated. The P1 Cre protein is a site-specific recombinase (31) that acts to cyclize the linear P1 viral DNA after its injection into E. coli. Cre must act quickly before the DNA is degraded. While the cre promoter is quite weak, it is expressed three to four times more efficiently in a Dam⁺ strain than in a Dam⁻ strain (N. Sternberg, unpublished data). This might be biologically relevant if the cre gene were rapidly replicated after viral injection. The resulting hemimethylated or perhaps even unmethylated state of the cre promoter would result in an increased transcription of the gene and a sufficient level of recombinase for the cyclization reaction.

The Mu mom gene encodes a DNA modification function, the expression of which requires the host Dam function. Hattman (8) has shown that this regulation operates at the transcription level since mom transcripts are reduced 20-fold in a Dam⁺ host. The proposed mom promoter has no Dam methylation sites but is just downstream from a region that contains a cluster of three Dam sites (10). Removal of some of these Dam sites eliminates the dependency on Dam
function (10, 12). Based on these and other observations, a model for the Dam regulation of mom has been proposed (10) in which mom is repressed by a protein that binds to the unmethylated mom promoter region but not to the methylated promoter. It is argued that methylation stabilizes or induces a DNA conformation (12, 28) that prevents access of the protein to the site. However, this model appears to be ruled out by the demonstration (A. Seiler and R. Kahmann, personal communication) that mutations in the mom regulatory region which do not affect the proposed DNA conformation make mom expression independent of dam function. An alternative hypothesis is suggested by the observation (A. Seiler and R. Kahmann, personal communication) that a mutH mutation can suppress the inhibition of mom expression in a dam" host. Preliminary experiments (K. Welsh, A.-L. Lu, and P. Modrich, personal communication) suggest that the MutH protein cleaves the unmethylated DNA strand at Dam sites in hemimethylated DNA, but not in methylated DNA. Taken together, these results point to a model in which the primary cause of the inhibition of mom expression in a Dam" host is the interaction between unmethylated DNA upstream of the mom gene and MutH protein.

In addition to affecting gene expression, Dam methylation appears to affect other protein-DNA interactions as well. Examples of two such interactions, DNA replication and DNA packaging, are discussed below.

**Replication initiation.** The E. coli origin of replication, oriC, has been localized to a region of 250 base pairs that contains 12 to 14 Dam methylation sites (27). Evidence for the importance of these sites in origin function comes from the following three sources. (i) Eight of the oriC Dam sites are conserved among origins from five gram-negative bacteria (34). (ii) Plasmids that use oriC for initiation of replication transform Dam" strains either not at all or very inefficiently. Moreover, oriC DNA isolated from Dam" cells functions two-to-fourfold less well in an oriC-specific in vitro initiation system than does oriC DNA from Dam" cells. Methylation of Dam" DNA in vitro restores normal template activity (31). (iii) Increased levels of Dam methylase in cells, due to a dam" gene on a high-copy-number plasmid, reduce the spacing between replication initiation events (24). The latter result supports the contention that the spacing of replication forks is regulated by the time required for the methylation of the newly replicated hemimethylated DNA. A model to account for this, based on the stabilization of DNA intrastrand secondary structure at oriC by Dam methylation, has been proposed (30).

Smith et al. (30) have shown that the Dam sites in oriC DNA isolated from Dam" mutants exhibit some sensitivity to digestion by DpnI, a restriction endonuclease specific for methylated GATC sites (4). They conclude that these mutant strains (one of which contains a Tn9 insertion in the dam gene) must retain an activity that can methylate at least some oriC Dam sites. The source of that methylase still needs to be determined. P. Janczyk and D. Smith (personal communication) have recently uncovered another regulatory circuit involving the dam gene. Some dnaATs mutants at 38°C exhibit 5 to 10 times more transcription from the dam gene promoter than does the isogenic dnaA" cell. This result suggests that the DnaA protein (an essential oriC initiation protein) may also be a repressor of dam.

**The regulation of P1 pac cleavage.** The packaging of bacteriophage P1 DNA into a viral capsid is initiated by the cleavage of the DNA at a unique site (pac) on a concatamer that consists of repeating P1 monomers (1, 31). Packaging proceeds in one direction from the cut end until a headful of DNA (more than a P1 monomer) has been encapsidated. The DNA is then cut again, and the next headful initiates at the point where the first one terminated. An essential feature of such a processive headful packaging mechanism is that no more than one or, at most, a few pac sites per concatemer should be cleaved. If each pac site were cut at the same time, the concatemer would be converted to monomer-sized pieces, and packaging would not be able to generate DNA with homologous sequences at both ends. Those homologous ends are necessary for cyclizing the viral DNA after its injection into a bacterial host.

The pac site has been localized to a 150- to 160-base-pair segment of DNA that contains seven Dam methylation sites, three at one end of the pac site and four at the other (N. Sternberg, unpublished data). Deletion of either of these two clusters of Dam sites inactivates pac. Moreover, unmethylated pac-containing DNA is insensitive to cleavage in vivo until it is methylated. J. Coulby and N. Sternberg (unpublished data) have recently shown that P1 encodes its own Dam-like methylase. Based on these observations, a model for the regulation of pac cleavage has been proposed in which a limiting step in the process is the activation of pac by the methylation of unmethylated Dam sites. During the early stages of the P1 vegetative cycle, pac is largely unmethylated and therefore inactive, because replication of the P1 DNA is occurring rapidly and the host methylation system cannot fully methylate it. Later in the infection cycle, some of the pac sites became methylated, due to the synthesis and action of the virus-encoded methylase. As soon as that occurs, pac becomes a substrate for the initiation of packaging, and packaging itself then blocks further utilization of the other pac sites. By requiring multiple methylation steps to activate pac, the system ensures that two pac sites on the same concatemer will rarely become activated at the same time.

The data summarized in this review clearly demonstrates that DNA methylation at 5'-GATC-3' sites is important in regulating gene expression and DNA processing in E. coli and in its bacteriophages. The fact that the methyl group at the N6 position of adenine is a critical part of this system suggests that methylation may play a significant role in the recognition of Dam sites by various proteins. In particular, it may affect specific interactions between RNA polymerase and promoters containing Dam sites. Those interactions need to be studied in greater detail. Also of particular importance will be studies that deal with the function of those Dam methylases, such as that encoded by bacteriophage T4 (29), whose biological role has not yet been identified. The next few years promise to contain some exciting additions to the methylease-gene function story.

**ACKNOWLEDGMENTS**

I thank D. Roberts, N. Kleckner, B. Hoopes, W. McClure, W. Reznikoff, R. Kahmann, D. Smith, W. Messer, M. Marinus, P. Modrich, and S. Hattman for communicating information before publication. I also thank R. Hoess for critical reading of the manuscript.

**ADDITION IN PROOF**

Robert Braun (personal communication) has recently shown that Dam methylation regulates the transcription of the E. coli dnaA gene. In a Dam" strain, transcription of that gene is reduced twofold relative to a Dam" strain.
LITERATURE CITED


