

## **Dam methylation: coordinating cellular processes**

Anders Løbner-Olesen<sup>1</sup>, Ole Skovgaard<sup>1</sup> and M.G. Marinus<sup>2</sup>

<sup>1</sup>Department of Life Sciences and Chemistry, Roskilde University, DK-4000 Roskilde, Denmark

<sup>2</sup>Department of Biochemistry and Molecular Pharmacology, University of Massachusetts

Medical School, Worcester MA 01605 USA

Corresponding Author:

Dr. M.G. Marinus

Biochemistry and Molecular Pharmacology

University of Massachusetts Medical School

364 Plantation Street, Worcester MA 01605 USA

Tel: 508 856 3330

Fax: 508 856 3036

[Martin.Marinus@umassmed.edu](mailto:Martin.Marinus@umassmed.edu)

Other Authors:

Dr. Anders Løbner-Olesen,

Department of Life Sciences and Chemistry, 18-1

Roskilde University,

Universitetsvej 1

P.O.Box 260

DK-4000 Roskilde, Denmark

Tel: 45 46 74 26 15

Fax: 45 46 74 30 11

lobner@ruc.dk

Dr. Ole Skovgaard,

Department of Life Sciences and Chemistry, 18-1

Roskilde University,

Universitetsvej 1

P.O.Box 260

DK-4000 Roskilde, Denmark

Tel: 45 46 74 24 05

Fax: 45 46 74 30 11

olesk@ruc.dk

## Summary

GATC sequences in *Escherichia coli* DNA are methylated at the adenine residue by DNA adenine methyltransferase (DamMT). These methylated residues and/or the level of DamMT can influence cellular functions such as gene transcription, DNA mismatch repair, initiation of chromosome replication and nucleoid structure. In certain bacteria, unlike *E. coli*, DamMT is essential for viability perhaps due to its role in chromosome replication. DamMT has also been implicated as a virulence factor in bacterial pathogenesis. The origin and phylogeny of DamMT, based on sequenced genomes, has been deduced.

## Introduction

In *E. coli* DNA almost all GATC sequences contain N<sup>6</sup>-methyladenine due to the action of Dam methyltransferase (DamMT), encoded by the *dam* gene. Deletion of the *dam* gene produces a variety of phenotypes indicating multiple functions for GATC methylation in modulating gene expression, DNA mismatch repair and chromosome initiation and nucleoid stabilization among others (Table 1). The recent contributions in these, and some additional, areas will be covered here to update the last available reviews on Dam methylation [1;2].

Replication of the fully methylated chromosome generates a transient wave of hemimethylated GATC sites (methylated on the parental strand, but not the daughter strand) behind the replication fork. In some chromosomal regions, such as the origin of chromosome replication, *oriC*, and the *dnaA* gene promoter, hemimethylated DNA persists for a large part of the cell

cycle. This likely results from the high density of GATC sites in these regions, which provides multiple binding sites for the SeqA protein. SeqA exist as a homotetramer in solution and each tetramer binds a pair of hemimethylated GATC sites separated by up to 31 bp [3;4](Fig. 1). The SeqA protein also binds fully methylated DNA albeit with lower affinity. Although the atomic structure of the SeqA protein has been determined it does not reveal the reason for high-affinity binding to hemimethylated DNA [5].

The level of DamMT in the cell, 130 molecules, is tightly regulated at the transcriptional level, involves multiple promoters and one terminator and is growth-rate regulated [2]. DamMT is a substrate for Lon protease suggesting another possible regulatory mechanism [6]. The methylation reaction itself has long been thought to occur by base flipping but in the only atomic structure available, the phage T4 DamMT-DNA co-crystal (Fig. 2) had not yet initiated this event [7]. DamMT acts in a processive manner; that is, for each binding event about 55 methylated GATC sites are formed before the enzyme dissociates from DNA [8].

DamMT is one of three (DcmMT, EcoKMT) dispensable methyltransferases in *E. coli* K-12 but it is required for viability in *Yersinia pseudotuberculosis* and *Vibrio cholerae* [1]. The reason for the growth requirement is unknown but for *V. cholerae*, which has two chromosomes, DamMT is required for initiation of replication of both chromosomes [9]. DamMT is also essential for bacteriophage P1 replication in part due to its role in packaging genomes into phage capsids but it may function in the control of immunity as well [10]. The *E. coli* K-12 CcrM (cell cycle regulated methyltransferase) is cell cycle regulated and essential for viability [11]. CcrM homologues are found in bacteria closely related to *E. coli* (Fig. 4). A different protein, also

called CcrM, is present in the alpha- proteobacteria and extensive work in *Caulobacter crescentus* suggests that CcrM is required for multiple cell cycle events including proper DNA replication initiation, DNA methylation, cell division, and cell wall metabolism [12;13].

### **DamMT modulation of gene expression**

GATC-methylation can influence gene expression by two different mechanisms. First, methylated GATC sequences in gene promoter regions can increase (e.g., *dnaA*), decrease (e.g., *sulA*) or have no effect on transcription initiation. Second, DamMT and regulatory proteins, such as Cap, Lrp or OxyR, compete for overlapping sites in, or near promoters. The *pap* genes have two GATC sequences located between the Pap divergent promoters to which there is alternative pair wise binding of DamMT, Lrp and PapI-Lrp to control off-to-on switching of pili on the cell surface (Fig. 3) [1;14]. This switching between the on and off states constitutes an epigenetic switch. A similar epigenetic switching occurs at three GATCs in the *agn43 (flu)* gene in *E. coli* which specifies an outer membrane protein that can cause autoaggregation ("fluffing") of bacterial cells and is important for mature biofilm formation and increased resistance to antimicrobial agents [15-17].

Two studies employing high density microarray technology have measured changes in gene expression in *dam* mutants and each study showed increased expression of the SOS response genes which had previously been shown by genetic means [18;19]. Apart from the SOS regulon and a few other genes, the two studies differ markedly in the number of genes whose expression is affected by GATC methylation from 359 genes [18] to about 20 [19]. The reason(s) for this

discrepancy is not known but could include different genetic backgrounds, different *dam* alleles, different cultural conditions and different array preparation methods. The Oshima et al [18] study also included a proteome analysis where 93 proteins differed in concentration between the wildtype and *dam* mutant.

A comparison of global gene expression in *dam*, *seqA* and DamMT overproducing strains showed that the profiles for the SeqA-depleted and DamMT overproducing strains were almost identical and distinct from that in *dam* cells [19]. This result is compatible with a model in which DamMT and SeqA compete for GATC binding sites in hemimethylated DNA behind the replication fork. Even a 10-fold overproduction of DamMT reduces SeqA binding leading to a global alteration in transcription which is best explained by proposing that SeqA is an important factor in forming and/or maintaining chromosome structure (Fig. 1) [19].

### **DNA methylation and chromosome replication**

A subset of hemimethylated GATC sites are substrates for both DamMT and SeqA and the latter acts on them first unless there is overproduction of DamMT [19]. SeqA foci observed in living cells are dependent on ongoing replication and likely indicate the localization of newly replicated DNA and the replication factories within the cell [20;21]. Binding of SeqA to newly replicated DNA seems to serve a least two functions.

The first of these is to limit initiation of chromosome replication to once and only once per cell cycle. Newly replicated hemimethylated origins, which contain an over-representation of GATC

sites, remain bound by SeqA ("sequestered") and inactive for about one third of the cell cycle. During sequestration at least three mechanisms operate to lower the activity of the initiator protein, DnaA [22]. First, the *dnaA* promoter, which also contains an excess of GATC sequences, is sequestered for the same period of time as the origin, *oriC*. Second, new DnaA binding sites outside *oriC* are generated by replication which serve to titrate free DNA protein. Third, after initiation DnaA-ATP is converted to inactive DnaA-ADP by a process called RIDA, which is dependent on the beta-clamp of DNA polymerase III holoenzyme and the Hda protein [23]. Disruption of any of these three mechanisms leads to re-initiation at origins within the same cell cycle [24;25].

The second function of SeqA, in vivo, is the organization of nascent nucleoids behind the replication fork. SeqA tetramers bound to distant pairs of hemimethylated GATC sites can, concurrent with replication, organize nascent daughter chromosomes into nucleoid domains (Fig. 1) [26] which may require also the ability of SeqA to generate positive supercoils [27]. SeqA is also involved in proper chromosome segregation as both *seqA* mutants and SeqA overproducers shows gross defects in nucleoid positioning within the cell [28;29].

SeqA tetramers do not bind separate DNA molecules [26] indicating that they act strictly in *cis* and therefore presumably are not involved in cohesion of sister chromosomes [30]. Sister chromosome cohesion, therefore, most likely results from catenation of daughter chromosomes behind the replication fork [31]. Interestingly, SeqA protein bound to hemimethylated DNA may play a role in their decatenation as it interacts with, and stimulates, the decatenating activity of topoisomerase IV [32].

## **DamMT as a virulence factor for bacterial pathogenesis**

The involvement of DamMT as a virulence factor was first described for *Salmonella enterica* serovar Typhimurium where the *dam* mutant was out competed by wildtype in establishing fatal infections in mice and where mice previously infected with the *dam* mutant were less susceptible to superinfection by the wildtype (reviewed in [1]). A similar situation may occur in *Haemophilus influenzae* infections [33]. Changes in transcription of genes in pathogenicity islands in *dam* mutants are likely the cause of virulence attenuation by altering protein composition on the cell surface and/or protein secretion as demonstrated in *Salmonella enterica* serovar Typhimurium *dam* mutants [34]. An additional mechanism affecting virulence of *Salmonella enterica* that is influenced by DamMT, is conjugal transfer of its virulence plasmid, which is promoted by Lrp and inhibited by DamMT [35].

A role for DamMT as a virulence factor has also been proposed for *E. coli*, *Erwinia chrysanthemi*, *Yersinia pseudotuberculosis* and *Vibrio cholerae* and possible mechanisms are reviewed by Low et al [1]. Overproduction of DamMT in *Yersinia pseudotuberculosis* causes attenuation of virulence, secretion of several outer proteins (Yops) and heightened immunity [36]. It is possible that these effects may not be due to the high level of DamMT, but instead are the result of inhibition of SeqA binding as described for *E. coli* [19]. A similar rationale probably applies to *dam* plasmid attenuation of virulence in *Pasteurella multocida* which causes bovine respiratory disease [37]. A *Shigella flexneri dam* mutant shows no attenuation of virulence [38].

Phase variation in *Haemophilus influenzae* results in high-frequency switching in the expression of surface pili. Alteration in the number of di- or tetranucleotide repeats in the promoter region or coding sequence of a gene is responsible for phase variation. A *Haemophilus influenzae dam* mutant does not alter the frequency of tetranucleotide switching but increased dinucleotide switching by 14-fold. Dinucleotide, but not tetranucleotide, switching was increased 30-fold in *mutH*, *mutL* and *mutS* derivatives in which the DNA mismatch repair system is inactivated [39].

### **GATC methylation and mismatch repair**

DNA mismatch repair corrects base mismatches that occur as replication errors in the newly synthesized unmethylated strand of hemimethylated DNA following the replication fork. The mismatches are bound by MutS, which recruits MutL and MutH to form a ternary complex on the DNA [40]. Formation of the ternary complex activates the latent endonuclease activity of MutH which cleaves the unmethylated strand at a nearby GATC site. Excision of the region containing the mismatch and subsequent resynthesis restores the correct nucleotide sequence. In *dam* mutants the lack of strand methylation leads to repair occurring on either the new or the old strand behind the replication fork and to repair events at mismatches in unreplicated DNA. This leads to accumulation of single-strand breaks which can be converted to double-strand breaks which are repaired by recombination [41].

*E. coli dam* mutants are much more sensitive than wildtype to the cytotoxic action of agents such as cisplatin and methylating agents [42] and hydrogen peroxide [43]. Inactivation of mismatch

repair genes *mutS*, *mutL* or *mutH* in a *dam* mutant, however, confers resistance to these agents. In other words, the mismatch repair system sensitizes *dam* cells to killing by these agents but the mechanism by which this occurs is currently unknown. DNA double-strand breaks are produced by cisplatin, methylating agents [42] and oxidative damage [44] and these need to be repaired by homologous recombination. Inhibition of recombination at drug-induced lesions in a *dam* mutant exposed to these agents, therefore, is expected to be lethal and for cisplatin lesions, MutS blocks the action of RecA protein during the initiation of recombination repair [42]. Loss of mismatch repair capability in *dam mut* bacteria reduces the number of endogenous strand breaks thereby increasing recombinational capacity to allow repair of drug-induced strand breaks.

The DamMT and MutH proteins have hemimethylated GATC sequences as a common substrate. Overproduction of DamMT prevents the action of MutH but the reverse is not true [2]. The atomic structure of MutH has been solved but the structure is at odds with mutational studies [45;46]. These data will presumably be reconciled when the structure of the protein specifically bound to DNA is solved.

### **Origin and phylogeny of DamMT**

An analysis of fully sequenced bacterial genomes reveals that homologs to the DamMT are present in one clade of bacteria which consists of the orders *Enterobacteriales*, *Vibrionales*, *Aeromonadales*, *Pasteurellales*, and *Alteromonadales* (Fig. 4). Members of the DamMT-clade share the following characteristics: (a) The phylogenetic inference based on the DamMT sequence shows a similar derivation to the phylogenetic inference based on 16S RNA. (b) The

*dam* gene is organized in an operon with *aroK* and *aroB*. (c) They have homologs to SeqA and MutH [47]. (d) They have separated the replication initiator gene (*dnaA*) from the origin of replication (*oriC*) and (e) GATC sites are approximately 10-fold over-represented in their *oriC* and in their *dnaA* promoters.

The characteristics above suggest that the *dam* gene was acquired only once in the ancestor of the DamMT-clade and that the evolutionary success of this clade was completed with the acquisition of SeqA and MutH which happened either simultaneously or in a different event (Fig. 4). The HN-S and StpA histone-like proteins and the PriB primosomal protein are also common to the DamMT-clade. It was suggested [47] that the replacement of SMC with MukFEB happened concomitantly with the acquisition of DamMT but since *Shewanella oneidensis* still has SMC but not MukFEB proteins, it must have occurred later. Another hemimethylated DNA specific binding protein encoded by *yccV* is only present in the *Enterobacteriae* [48].

The sequences most similar to DamMT outside the DamMT-clade are found in *Neisseria meningitidis* and *Wautersia metallidurans*. In *N. meningitidis*, a *dam* mutation does not influence mutability nor phase variation [49]. DamMT has also a strong similarity to the modification methylases MboIA and DpnIIA, suggesting that DamMT was acquired from such a restriction-modification system.

## Conclusions

From the topics discussed above it is clear that *E. coli dam* bacteria, or strains over expressing DamMT, remain useful tools to investigate basic molecular biological processes in bacteria. Inactivation of the *dam* gene in other bacteria will undoubtedly lead to more functions being uncovered. For bacteria in which *dam* is essential for viability, future work will obviously be aimed at discovering what cellular function(s) DamMT is critical for. The recent discovery of the cell-cycle regulated CcrM methyltransferase needs to be investigated further to determine the essential function of this protein. SeqA's role in initiating nucleoid organization needs to be clarified in relation to known nucleosome organizing proteins such as MukFEB, SMC, HN-S, HU, Fis, etc., and how systems involving DamMT, SeqA and MutH modulate their roles during the short existence of hemimethylated DNA. Although much insight has been obtained into the mechanism by which DnaA, DamMT, SeqA and other proteins coordinate synchronous initiation of chromosome replication with the cell cycle in *E. coli*, the details still need to be worked out. The mechanism of chromosome initiation in *dam* mutants, in which there should be an altered level of DnaA and less SeqA binding than wildtype, needs further attention.

The realization that recombination is essential for *E. coli dam* mutant viability has allowed for new interpretations of the phenomenon whereby mismatch repair-proficient *dam* bacteria and mammalian cells are sensitive to certain anticancer agents and base analogues. The initial studies suggest that drug-modified DNA is recognized as "foreign" and because of this mismatch repair prevents recombination repair mechanisms from operating. This proposal needs to be investigated further especially the mechanism(s) that initiate recombinational repair and to determine if a similar mechanism operates in mammalian cells.

## **Acknowledgements**

A.L.O. and O.S. were supported by grants from the Danish Natural Research Foundation and the Danish Medical Research Council. Research conducted in the M.G. M. laboratory during the review period (2002-2004) was supported by grant GM63790 from the National Institutes of Health.

## References and recommended reading

1. Low DA, Weyand NJ, Mahan MJ: **Roles of DNA adenine methylation in regulating bacterial gene expression and virulence.** *Infect.Immun.* 2001, **69**:7197-7204.
2. Marinus MG: **Methylation of DNA.** In *Escherichia coli and Salmonella: Cellular and Molecular Biology.* Edited by Neidhardt FC, Curtiss R, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umberger HE. Washington DC: ASM Press; 1996:782-791.
3. Brendler T, Sawitzke J, Sergueev K, Austin S: **A case for sliding SeqA tracts at anchored replication forks during *Escherichia coli* chromosome replication and segregation.** *EMBO J* 2000, **19**:6249-6258.
4. Han JS, Kang S, Kim SH, Ko MJ, Hwang DS: **Binding of SeqA protein to hemimethylated GATC sequences enhances their interaction and aggregation properties.** *J Biol.Chem.* 2004, **279**:30236-30243.
- \*\*5. Guarne A, Zhao Q, Ghirlando R, Yang W: **Insights into negative modulation of *E. coli* replication initiation from the structure of SeqA-hemimethylated DNA complex.** *Nat.Struct.Biol.* 2002, **9**:839-843.

The crystal structure of the C-terminal DNA-binding domain of SeqA complexed to a hemimethylated GATC site was determined. It was proposed that a tetramer of SeqA is capable of forming pseudo stem-loop structures between hemimethylated GATC sites on the same DNA helix. The structure was not informative as to why SeqA has a higher binding affinity for hemimethylated versus fully methylated DNA.

6. Calmann MA, Marinus MG: **Regulated expression of the *Escherichia coli* dam gene.** *J Bacteriol.* 2003, **185**:5012-5014.
- \*7. Yang Z, Horton JR, Zhou L, Zhang XJ, Dong A, Zhang X, Schlagman SL, Kossykh V, Hattman S, Cheng X: **Structure of the bacteriophage T4 DNA adenine methyltransferase.** *Nat.Struct.Biol.* 2003, **10**:849-855.

The first, and still the only, atomic structure for a DamMT. The structure obtained had two DamMT monomers per GATC site which was unexpected and this prevented the detection of a base-flipping mechanism. The enzyme has two domains, one of which binds S-adenosyl-homocysteine and the other for binding DNA. The five helix bundle and a beta-hairpin that comprise the DNA binding domain are conserved in the MTs recognizing GATC substrate sequences.

- \*\*8. Urig S, Gowher H, Hermann A, Beck C, Fatemi M, Humeny A, Jeltsch A: **The *Escherichia coli* dam DNA methyltransferase modifies DNA in a highly processive reaction.** *J.Mol.Biol.* 2002, **319**:1085-1096.

This paper showed that DamMT is a monomer and methylates only one strand of the DNA in each binding event. DamMT has a highly processive reaction mechanism - about 55 sites per binding event. The highly processive mechanism can explain why only small amounts of the enzyme are needed to maintain the methylation state. The processive mode of action is in contrast to the distributive mechanism used by MTs which are part of restriction-modification systems.

- \*\*9. Egan ES, Waldor MK: **Distinct replication requirements for the two *Vibrio cholerae* chromosomes.** *Cell* 2003, **114**:521-530.

There are essential genes on each of the two *Vibrio* chromosomes and the origins of replication (*oriCI* and *II*) of the chromosomes were cloned as autonomously-replicating sequences. The origins are not related to each other although both require DamMT for initiation. One of the origins, *oriCI*, shares conserved properties with that of *E. coli oriC* but *oriCII* requires, in addition to DnaA, an additional DNA binding protein, a repeat sequence, and an RNA.

10. Yarmolinsky M: **Bacteriophage P1 in retrospect and prospect.** *J Bacteriol.* 2004, **186**:7025-7028.

- \*11. Kossykh VG, Lloyd RS: **A DNA adenine methyltransferase of *Escherichia coli* that is cell cycle regulated and essential for viability.** *J.Bacteriol.* 2004, **186**:2061-2067.

CcrM is the only known methyltransferase in *E. coli* that is essential for growth. It is cell cycle-regulated which may explain why it was not detected previously. Complete methylation of ATGCAT substrate sites in the chromosome can be achieved by overproducing the enzyme but the overproduction slows growth and produces cells with various sizes and DNA content.

12. Reisenauer A, Shapiro L: **DNA methylation affects the cell cycle transcription of the CtrA global regulator in *Caulobacter*.** *EMBO J.* 2002, **21**:4969-4977.

13. Chen SL, Shapiro L: **Identification of long intergenic repeat sequences associated with DNA methylation sites in *Caulobacter crescentus* and other alpha-proteobacteria.** *J.Bacteriol.* 2003, **185**:4997-5002.

- \*14. Hernday AD, Braaten BA, Low DA: **The mechanism by which DNA adenine methylase and PapI activate the *pap* epigenetic switch.** *Mol.Cell* 2003, **12**:947-957.

The PapI regulator was shown to increase the affinity of the global regulatory protein Lrp to bind to ACGATC sequences in the promoter region. Mutational analysis showed that DamMT methylation blocks PapI-dependent binding of Lrp suggesting that such binding to hemimethylated DNA in vivo is important for the switching mechanism.

15. Wallecha A, Correnti J, Munster V, van der Woude M.: **Phase variation of Ag43 is independent of the oxidation state of OxyR.** *J Bacteriol.* 2003, **185**:2203-2209.

16. Correnti J, Munster V, Chan T, van der Woude M: **Dam-dependent phase variation of Ag43 in *Escherichia coli* is altered in a *seqA* mutant.** *Mol.Microbiol.* 2002, **44**:521-532.
17. Waldron DE, Owen P, Dorman CJ: **Competitive interaction of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in *Escherichia coli*.** *Mol.Microbiol.* 2002, **44**:509-520.
- \*18. Oshima T, Wada C, Kawagoe Y, Ara T, Maeda M, Masuda Y, Hiraga S, Mori H: **Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*.** *Mol.Microbiol.* 2002, **45**:673-695.

Transcriptional profile and proteomics of a *dam* mutant using microarray analysis. A very large number of genes were found with increased transcription including those involved in aerobic respiration, stress responses, amino acid and nucleotide metabolism. There was decreased transcription of genes for anaerobic respiration, flagellar biosynthesis, chemotaxis and motility. Many of the DamMT-controlled genes had overlapping fumarate nitrate reduction (Fnr) and catabolite activator protein (CAP) recognition sequences.

- \*19. Lobner-Olesen A, Marinus MG, Hansen FG: **Role of SeqA and Dam in *Escherichia coli* gene expression: a global/microarray analysis.** *Proc.Natl Acad.Sci.U.S.A* 2003, **100**:4672-4677.

Global transcription patterns were determined for *dam*, *seqA* and DamMT-overproducing strains. The patterns for *seqA* and DamMT-overproducing bacteria were almost identical indicating that Dam and SeqA compete for binding sites not only at *oriC* but in hemimethylated DNA trailing the replication fork. The results are consistent with a role for SeqA in forming ~50 transcriptional domains in the chromosome. The global transcription pattern for the *dam* mutant was different from the other two strains.

20. Hiraga S: **Dynamic localization of bacterial and plasmid chromosomes.** *Annu.Rev.Genet.* 2000, **34**:21-59.
21. Molina F, Skarstad K: **Replication fork and SeqA focus distributions in *Escherichia coli* suggest a replication hyperstructure dependent on nucleotide metabolism.** *Mol.Microbiol.* 2004, **52**:1597-1612.
22. Boye E, Lobner-Olesen A, Skarstad K: **Limiting DNA replication to once and only once.** *EMBO Rep.* 2000, **1**:479-483.
- \*\*23. Kato J, Katayama T: **Hda, a novel DnaA-related protein, regulates the replication cycle in *Escherichia coli*.** *EMBO J* 2001, **20**:4253-4262.

The Hda (homology to DnaA) protein was identified and characterized. The Hda protein is necessary for the replication dependent conversion of DnaA-ATP, the active form for initiation, to DnaA-ADP, which is inactive. This conversion prevents excessive initiation events from occurring at the same *oriC*.

- \*24. Ogawa T, Yamada Y, Kuroda T, Kishi T, Moriya S: **The *datA* locus predominantly contributes to the initiator titration mechanism in the control of replication initiation in *Escherichia coli*.** *Mol.Microbiol.* 2002, **44**:1367-1375.

Of the eight major high affinity chromosomal binding sites for DnaA, only inactivation of the *datA* locus results in over-initiation of chromosome replication. This shows that the *datA* locus is important in adjusting the levels of free and bound DnaA to ensure that initiation occurs only once per cell cycle. The 1 kb *datA* region contains five repeated sequences capable of binding hundreds of DnaA molecules

25. Camara JE, Skarstad K, Crooke E: **Controlled initiation of chromosomal replication in *Escherichia coli* requires functional Hda protein.** *J Bacteriol.* 2003, **185**:3244-3248.
26. Han JS, Kang S, Lee H, Kim HK, Hwang DS: **Sequential binding of SeqA to paired hemi-methylated GATC sequences mediates formation of higher order complexes.** *J Biol.Chem.* 2003, **278**:34983-34989.
- \*27. Klungsoyr HK, Skarstad K: **Positive supercoiling is generated in the presence of *Escherichia coli* SeqA protein.** *Mol.Microbiol.* 2004, **54**:123-131.
- SeqA protein in vitro generates positive supercoils in covalently-closed or nicked DNA, suggesting that it participates in organizing the bacterial nucleoid at the level of hemimethylated DNA trailing the replication fork.
28. von Freiesleben U, Krekling MA, Hansen FG, Lobner-Olesen A: **The eclipse period of *Escherichia coli*.** *EMBO J* 2000, **19**:6240-6248.
29. Bach T, Krekling MA, Skarstad K: **Excess SeqA prolongs sequestration of *oriC* and delays nucleoid segregation and cell division.** *EMBO J* 2003, **22**:315-323.
30. Sunako Y, Onogi T, Hiraga S: **Sister chromosome cohesion of *Escherichia coli*.** *Mol.Microbiol.* 2001, **42**:1233-1241.
31. Postow L, Crisona NJ, Peter BJ, Hardy CD, Cozzarelli NR: **Topological challenges to DNA replication: conformations at the fork.** *Proc.Natl Acad.Sci.U.S.A* 2001, **98**:8219-8226.
32. Kang S, Han JS, Park JH, Skarstad K, Hwang DS: **SeqA protein stimulates the relaxing and decatenating activities of topoisomerase IV.** *J Biol.Chem.* 2003, **278**:48779-48785.
33. Watson ME, Jr., Jarisch J, Smith AL: **Inactivation of deoxyadenosine methyltransferase (*dam*) attenuates *Haemophilus influenzae* virulence.** *Mol.Microbiol.* 2004, **53**:651-664.
34. Pucciarelli MG, Prieto AI, Casadesus J, Garcia-del Portillo F: **Envelope instability in DNA adenine methylase mutants of *Salmonella enterica*.** *Microbiology* 2002, **148**:1171-1182.

- \*35. Camacho EM, Casadesus J: **Conjugal transfer of the virulence plasmid of *Salmonella enterica* is regulated by the leucine-responsive regulatory protein and DNA adenine methylation.** *Mol.Microbiol.* 2002, **44**:1589-1598.

DamMT and Lrp regulate the transfer operon of the virulence plasmid by opposing actions. Lrp acts by promoting *traJ* transcription thereby promoting transfer. DamMT activates FinP RNA synthesis which acts as a conjugational repressor. The rate of plasmid transfer may be influenced by global regulators responding to environmental cues. The results suggest that a similar regulatory circuit also operates in other F-like plasmids.

36. Julio SM, Heithoff DM, Sinsheimer RL, Low DA, Mahan MJ: **DNA adenine methylase overproduction in *Yersinia pseudotuberculosis* alters YopE expression and secretion and host immune responses to infection.** *Infect.Immun.* 2002, **70**:1006-1009.
37. Chen L, Paulsen DB, Scruggs DW, Banes MM, Reeks BY, Lawrence ML: **Alteration of DNA adenine methylase (Dam) activity in *Pasteurella multocida* causes increased spontaneous mutation frequency and attenuation in mice.** *Microbiology* 2003, **149**:2283-2290.
38. Honma Y, Fernandez RE, Maurelli AT: **A DNA adenine methylase mutant of *Shigella flexneri* shows no significant attenuation of virulence.** *Microbiology* 2004, **150**:1073-1078.
39. Bayliss CD, Sweetman WA, Moxon ER: **Mutations in *Haemophilus influenzae* mismatch repair genes increase mutation rates of dinucleotide repeat tracts but not dinucleotide repeat-driven pilin phase variation rates.** *J.Bacteriol.* 2004, **186**:2928-2935.
40. Modrich P, Lahue R: **Mismatch repair in replication fidelity, genetic recombination, and cancer biology.** *Annu.Rev.Biochem.* 1996, **65**:101-133.
41. Marinus MG: **Recombination is essential for viability of an *Escherichia coli dam* (DNA adenine methyltransferase) mutant.** *J.Bacteriol.* 2000, **182**:463-468.
- \*42. Calmann MA, Marinus MG: **MutS inhibits RecA-mediated strand exchange with platinated DNA substrates.** *Proc.Natl.Acad.Sci.U.S.A* 2004, **101**:14174-14179.

This paper presents evidence for a mechanism to explain the sensitization effect of mismatch repair in *E. coli dam* cells exposed to the antitumor agent, cisplatin. The authors suggest that during recombinational repair, platinated DNA is perceived as homeologous (similar but not identical in sequence) and the formation of recombination intermediates is blocked preventing repair and promoting cytotoxicity.

43. Wyrzykowski J, Volkert MR: **The *Escherichia coli* methyl-directed mismatch repair system repairs base pairs containing oxidative lesions.** *J.Bacteriol.* 2003, **185**:1701-1704.

44. Spek EJ, Vuong LN, Matsuguchi T, Marinus MG, Engelward BP: **Nitric oxide-induced homologous recombination in *Escherichia coli* is promoted by DNA glycosylases.** *J.Bacteriol.* 2002, **184**:3501-3507.
45. Wu TH, Loh T, Marinus MG: **The function of Asp70, Glu77 and Lys79 in the *Escherichia coli* MutH protein.** *Nucleic Acids Res.* 2002, **30**:818-822.
46. Friedhoff P, Thomas E, Pingoud A: **Tyr212: a key residue involved in strand discrimination by the DNA mismatch repair endonuclease MutH.** *J.Mol.Biol.* 2003, **325**:285-297.
- \*47. Hiraga S, Ichinose C, Onogi T, Niki H, Yamazoe M: **Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of *oriC* copies in *Escherichia coli*.** *Genes Cells* 2000, **5**:327-341.  
  
It was shown that SeqA focus formation is dependent on ongoing replication and most likely represents binding to hemimethylated DNA at the replication fork. Daughter copies of *oriC* were found linked to each other for an extended period following replication.
48. d'Alencon E, Taghbalout A, Bristow C, Kern R, Aflalo R, Kohiyama M: **Isolation of a new hemimethylated DNA binding protein which regulates *dnaA* gene expression.** *J.Bacteriol.* 2003, **185**:2967-2971.
49. Richardson AR, Stojiljkovic I: **Mismatch repair and the regulation of phase variation in *Neisseria meningitidis*.** *Mol.Microbiol.* 2001, **40**:645-655.

## Figure Legends

Figure 1. The SeqA protein organizes nascent nucleoids behind the replication fork. A SeqA tetramer is composed of two dimers. One monomer of each dimer binds a hemimethylated GATC site (open circles) at the same face of the DNA (blue) to form a stable complex (A). Stepwise interaction of SeqA tetramers bound to hemimethylated GATC sites in adjacent regions of the same DNA helix results in looping out of the intervening DNA (B). At the replication fork, fully methylated GATC sites (filled circles) are converted to their hemimethylated counterparts. These sites are bound by SeqA molecules to condense and loop the chromosome, thereby forming a nucleoid structure that can be maintained by other proteins (not shown). Note that SeqA tetramers aggregate strictly in *cis* which is important for maintaining separation of nascent nucleoids. The figure is adapted from reference [3].

Figure 2. The loose ternary structure of T4 DamMT with S-adenosyl-L-homocysteine (the methyl donor product) bound to 12-mer oligonucleotides containing a single GATC sequence. The oligonucleotides are stacked head-to-head forming a pseudo duplex. One enzyme molecule binds to the oligonucleotide and the second binds to the joint of the two DNA molecules. The enzyme molecules below the oligonucleotides are rotated 180° compared to those at the top. This figure is reproduced with permission of Dr. Xiaodong Chen and Nature Structural Biology [7].

Figure 3. Switching at the *pap* promoter region. Lrp binds co-operatively to either Lrp binding sites 1-3 (in the OFF-state, non-piliated) or to the Lrp binding sites 4-6 (ON-state, pilated). Lrp binding site 3 overlaps the *papB* promoter and Lrp binding to site 3 inhibits transcription. Lrp binding sites 2 and 5 overlaps with GATC sites and Lrp binding to either site prevents

methylation of that site by DamMT. Lrp binding to sites 1-3 mutually excludes binding to sites 4-6. When in OFF-state, each DNA replication produce one hemimethylated GATC site (in Lrp site 5) and one unmethylated GATC site (in Lrp site 2) and dissociate Lrp from its binding sites. The OFF-state is preserved by rebinding of Lrp to the same binding sites around the unmethylated Lrp binding site. A shift from phase OFF to ON may occur if PapI mediates Lrp binding to the hemimethylated Lrp site 5, followed by Lrp binding to sites 4 and 6. The shift is further stabilized by full methylation of site 2 by DamMT and conversion of the hemimethylated site 5 to unmethylated by subsequent DNA replications. This figure is adapted from reference [14].

Figure 4. Phylogeny of the DamMT clade. 4311 annotated *E. coli* protein sequences were searched with TBLASTN against selected genomic sequences with and without the *dam* gene. The genes encoding SeqA, MutH, HN-S, PriB, and 75 other proteins were found in all of the selected genomes with the *dam* gene and in none of the genomes without the *dam* gene (*P. aeruginosa*, *P. putida*, *Xylella fastidiosa*, *Xanthomonas campestris*, and *Coxiella burnetii*). The acquisition of selected genes and the loss of genes for SMC are indicated on a 16S RNA based phylogenetic tree. The bar indicates phylogenetic distance.

Table 1. Physiological systems influenced by GATC methylation and/or the level of DamMT

System	Protein	Function
Chromosome replication	SeqA	Binds to hemimethylated <i>oriC</i> DNA
	DnaA	Initiation of replication
	<i>oriC</i>	Replication origin
Gene expression	LexA	Repressor of the SOS regulon
	TrpR	Tryptophan operon repressor
Mismatch repair	MutH	Strand-specific endonuclease
Nucleoid structure	SeqA	Binds to hemimethylated DNA at replication forks

The table shows the major physiological systems affected by GATC methylation in *E. coli*. The *oriC* region has an over-representation of GATC sequences and one of the *dnaA* promoters needs to be GATC methylated for maximal expression. The SeqA protein binds to hemimethylated *oriC* and to a subset of hemimethylated GATC sites behind the replication fork. The SOS regulon is partially induced in *dam* mutants including the gene for LexA. Methylation of a GATC sequence in the *trpR* promoter decreases transcription initiation. MutH cleaves at GATC sites in the unmethylated strand of hemimethylated DNA. A more detailed list of physiological properties of *dam* mutants can be found in reference [2]. SeqA's role in nucleoid structure is diagrammed in Fig. 1.

**Keywords:** DNA Methylation, *Escherichia coli*, Transcription, DNA replication, Cell division, Chromosome, Virulence, Mismatch repair,

## Glossary

AroK	Shikimate kinase I (aromatic amino acid biosynthesis)
AroB	3-Dehydroquinate synthase
Cap	Global positive regulator of transcription
CcrM	Cell cycle regulated DNA adenine methyltransferase
DamMT	DNA adenine methyltransferase
DcmMT	DNA cytosine methyltransferase
DnaA	Positive initiator of chromosome replication
EcoKMT	Host specificity DNA methyltransferase
Fis	Maintenance of nucleoid structure
HN-S	Maintenance of nucleoid structure
HU	Maintenance of nucleoid structure
LexA	Repressor of the SOS response genes
Lon	A major protease
Lrp	Global regulatory protein
MukFEB	Proteins involved in chromosome partitioning
MutH	Mismatch repair, binds hemimethylated DNA
MutL	Mismatch repair
MutS	Mismatch repair, mismatch recognition
<i>oriC</i>	Origin of replication
OxyR	Regulatory protein and sensor for oxidative stress
Pap	Pili on uropathogenic <i>E. coli</i>

PriB	Required for initiation of chromosome replication at non- <i>oriC</i> sites
RecA	Initiates strand transfer during genetic recombination
SeqA	Negative initiator of chromosome replication, maintenance of nucleoid structure
SMC	Maintenance of nucleoid structure
StpA	Maintenance of nucleoid structure
TrpR	Repressor of tryptophan biosynthesis
YccV	Hypothetical protein
Yops	<i>Yersinia</i> outer membrane proteins

Mutations in the genes encoding the DNA binding proteins listed in the Table as maintaining nucleoid structure alter nucleoid structure and indirectly affect processes in the cell such as transcription initiation, transposition, initiation of chromosome replication, recombination, etc.

Figure 1

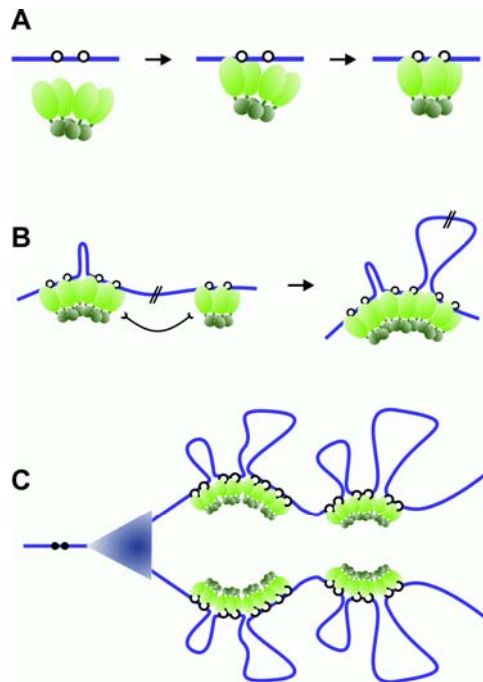


Figure 2

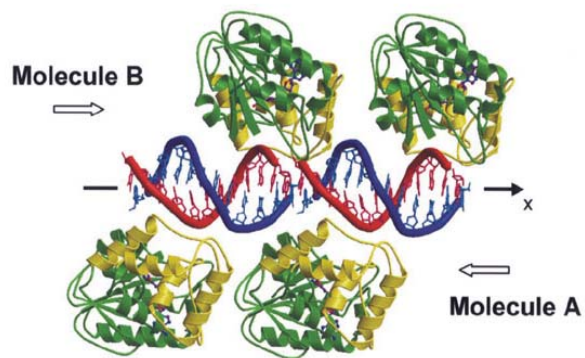


Figure 3

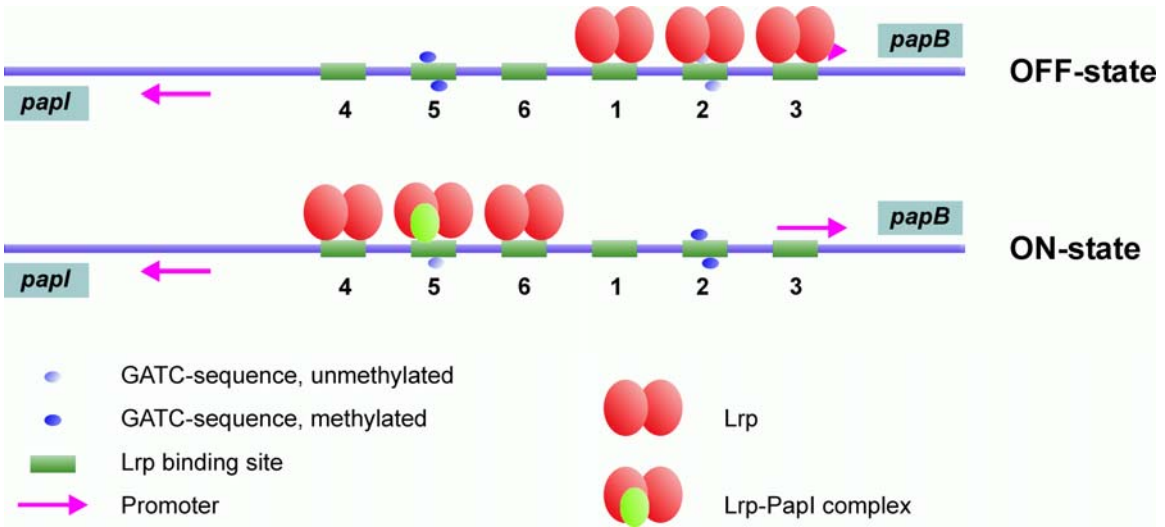


Figure 4

