Molecular Microbiology (2018) 110(1), 1–10

doi:10.1111/mmi.14079
First published online 11 July 2018

MicroReview

The impact of DNA methylation in *Alphaproteobacteria*

Annabelle Mouammine and Justine Collier *
Faculty of Biology and Medicine, Department of Fundamental Microbiology, University of Lausanne, Quartier UNIL/Sorge, Lausanne, CH 1015, Switzerland.

Abstract

***Alphaproteobacteria* include bacteria with very different modes of life, from free-living to host-associated and pathogenic bacteria. Their genomes vary in size and organization from single circular chromosomes to multipartite genomes and are often methylated by one or more adenine or cytosine methyltransferases (MTases). These include MTases that are part of restriction/modification systems and so-called orphan MTases. The development of novel technologies accelerated the analysis of methylomes and revealed the existence of epigenetic patterns in several *Alphaproteobacteria*. This review describes the known functions of DNA methylation in *Alphaproteobacteria* and also discusses its potential drawbacks through the accidental deamination of methylated cytosines. Particular emphasis is given to the strong connection between the cell cycle-regulated orphan MTase CcrM and the complex network that controls gene expression and cell cycle progression in *Alphaproteobacteria*.**

Introduction

Methylated bases can be found on the genomes of organisms from all kingdoms of life. The role of these epigenetic marks has been mostly studied in eukaryotes, where they are involved in a variety of processes such as cellular differentiation, embryogenesis, genomic imprinting and cancer development (Egger *et al.*, 2004; Schubeler, 2015; Smith and Meissner, 2013). Methylated cytosines and adenines are however also frequently found in bacteria and can represent more than 2% of the bases in a bacterial genome (Marinus and Lobner-Olesen, 2014; Sanchez-Romero *et al.*, 2015). Methylated bases (m5C, m4C and m6A) are often used by bacteria to discriminate their own DNA from exogenous DNA entering the cell during horizontal gene transfers (HGT). In this case, motifs regularly spread on self-DNA are methylated by a DNA methyltransferase (MTase) to protect them from cleavage by a restriction endonuclease (RE) partner. The MTase and the RE then belong to a so-called restriction-modification (R/M) system (Vasu and Nagaraja, 2013; Loenen *et al.*, 2014). RE cleave non-methylated motifs detected on exogenous DNA, thereby acting as 'immigration controllers' limiting the efficiency of HGT.

In addition to these R/M-associated MTases, bacteria often encode orphan MTases that are no longer associated with an RE partner. The most recent estimate comparing 230 bacterial and archaeal genomes is that ~half of these bacteria encode minimum one active orphan MTase (Blow *et al.*, 2016). Although these have lost their capacity to control HGT, there is more and more evidence that several of these are very important, if not essential, for cell cycle progression, virulence or genome maintenance (Collier, 2009; Marinus and Casadesus, 2009; Sanchez-Romero *et al.*, 2015; Adhikari and Curtis, 2016). The two best known examples are the conserved Dam MTase found in many *Gammaproteobacteria* and the CcrM MTase found in most *Alphaproteobacteria* (Adhikari and Curtis, 2016). These two MTases methylate adenines found in specific motifs, creating m6A epigenetic marks that can affect the activity of specific DNA binding proteins such as transcription factors, endonucleases or initiators of DNA replication. Dam-dependent methylation is, for example, essential for the replication of one of the chromosomes of *Vibrio cholerae* (Marinus and Lobner-Olesen, 2014) and required for the phase-variable expression of several virulence factors in clonal populations of *Salmonella enterica* (Garcia-Pastor *et al.*, 2018). Interestingly, many bacterial pathogens also express other phase-variable orphan MTases that control randomly switching regulons, called phasevarions, encoding proteins involved

Accepted 7 July 2018. *For correspondence. E-mail Justine.collier@unil.ch; Tel. +41 21 692 5610; Fax +41 21 692 5605.

in pathogenesis, host-adaptation and antibiotics resistance (Atack *et al*., 2018).

This review focuses on the functions of DNA methylation in *Alphaproteobacteria.* This class of bacteria displays very diverse life cycles as it includes free-living, host-associated and intracellular bacteria that sometimes act as plant, animal or human pathogens.

Methylated DNA motifs and DNA methyltransferases in Alphaproteobacteria

Until recently, the detection of methylated bases involved the use of complex methods such as the use of methylation-sensitive restriction enzymes and Southern blotting to detect the methylation state of specific DNA motifs on bacterial genomes. Although time-consuming, these methods proved efficient to discover the function of the best characterized orphan and cell cycle-regulated MTase CcrM. Initially discovered in *Caulobacter crescentus,* CcrM homologs are however found in nearly every *Alphaproteobacteria*, except *Rickettsiales* and *Magnetococcales* (Wright *et al*., 1997; Brilli *et al*., 2010; Gonzalez *et al*., 2014). *C. crescentus* is a free-living bacterium commonly used to study the regulation of the bacterial cell cycle because its cell cycle can be synchronized (Collier, 2016). It replicates asymmetrically into two distinct cell types: a stalked cell that immediately replicates its chromosome (S-phase) while it elongates into a pre-divisional cell, and a swarmer cell that cannot replicate its chromosome (G1-phase) before it differentiates into a stalked cell (Fig. 1A). CcrM methylates adenines in 5′-GANTC-3′ motifs in late pre-divisional cells, when the replication of the unique circular chromosome of *C. crescentus* has just ended (Zweiger *et al*., 1994). Considering that DNA replication is semi-conservative, GANTC motifs stay in a hemi-methylated state (only the old DNA strand is methylated) for some time after their replication and until cells reach the pre-divisional stage. The duration of this period depends on the location of the motif on the chromosome. Indeed, motifs located close to the origin of replication stay hemi-methylated during most of the S-phase of the cell cycle, while motifs located near the chromosomal terminus are methylated immediately after their replication (Fig. 1A). This model was recently confirmed at the genome scale thanks to the use of Single Molecule Real Time (SMRT) sequencing, which allows the detection of methylated motifs on bacterial genomes (Flusberg *et al*., 2010; Davis *et al*., 2013). This technique revealed that most GANTC motifs are efficiently methylated by CcrM on the *C. crescentus* genome, highlighting the efficiency of this enzyme (Kozdon *et al*., 2013; Gonzalez *et al*., 2014). It is unclear whether CcrM is a processive enzyme that remains attached to DNA and methylates multiple GANTC motifs before dissociating

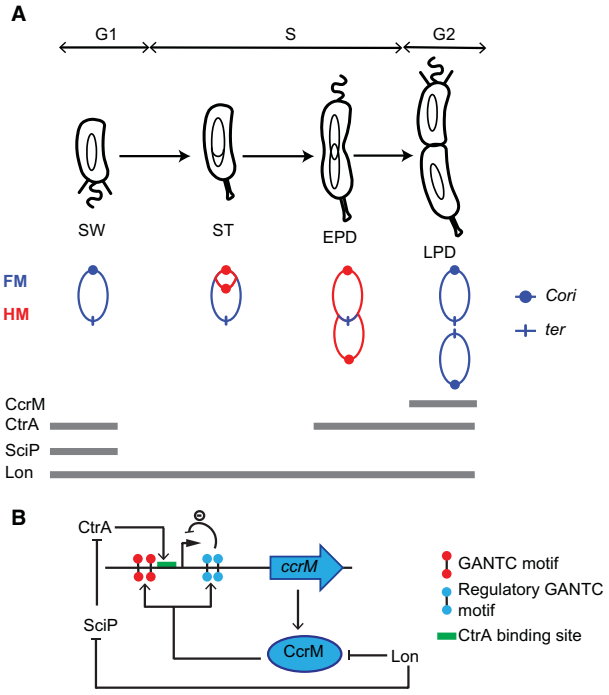

Fig. 1. DNA methylation by CcrM during the *C. crescentus* cell cycle.
A. Diagram of the *C. crescentus* cell cycle showing the methylation state of its chromosome and the abundance of key regulators at different stages. Top panel: Diagram of the cell cycle. SW, ST, EPD and LPD indicate swarmer, stalked, early pre-divisional and late pre-divisional cells respectively. Middle panel: Methylation state of most GANTC motifs on the chromosome as a function of the cell cycle. FM and HM indicate fully-methylated and hemi-methylated GANTC motifs respectively. *Cori* and *ter* represent the chromosomal origin and terminus respectively. Fully-methylated GANTC motifs (DNA colored in blue) become hemi-methylated (DNA colored in red) upon replication. CcrM converts hemi-methylated GANTC motifs into fully-methylated GANTC motifs in late pre-divisional cells when DNA replication is finished. Bottom panel: Schematic representing the intracellular levels of CcrM and of regulators of CcrM as a function of the cell cycle. A grey bar indicates that the protein is abundant at that time of the cell cycle.
B. Transcriptional and post-transcriptional control of CcrM levels. The *ccrM* promoter region carries four GANTC motifs (each represented by a lollipop): two before and two after the transcriptional start site (Reisenauer *et al*., 1999). The two blue motifs are supposedly important for the feedback regulation of *ccrM* transcription by CcrM. The blue arrow indicates the *ccrM* coding sequence. [Colour figure can be viewed at wileyonlinelibrary.com]

from the DNA (Berdis *et al*., 1998; Shier *et al*., 2001; Woodcock *et al*., 2017) or a distributive enzyme that dissociates after each methylation reaction (Albu *et al*., 2012; Maier *et al*., 2015).

Analysis of the *C. crescentus* methylome also revealed the existence of four other functional DNA MTases: two are adenine MTases predicted to be part of R/M systems, while the two others are supposedly orphan cytosine MTases (Kozdon *et al*., 2013). Interestingly, motifs methylated by these two orphan MTases appear as only partially methylated, suggesting the possible existence of

methylation patterns or phasevarions controlled by these MTases, although their function remains elusive. So far, the methylomes of ~20 *Alphaproteobacteria* have been characterized using SMRT sequencing (Kozdon *et al*., 2013; Gonzalez *et al*., 2014; Blow *et al*., 2016) and bioinformatics searches indicate that genes encoding putative DNA MTases are frequently found in alphaproteobacterial genomes (Roberts *et al*., 2015; Oliveira *et al*., 2016). The function of some of these MTases has now been characterized in more details. These include six DNA MTases that are part of R/M systems in the marine bacterium *Agrobacterium gelatinovorum* (Suzuki *et al*., 1996b), in the methylotrophic bacterium *Paracoccus aminophilus* (Dziewit *et al*., 2011), in the photosynthetic bacterium *Rhodobacter sphaeroides* (Scavetta *et al*., 2000; Szegedi and Gumport, 2000; Szegedi *et al*., 2000), in the plant symbiont *Rhizobium leguminosarum* (Rochepeau *et al*., 1997), and in the bioethanol and acetic acid producing *Zymomonas mobilis* (Kerr *et al*., 2011) and *Acetobacter pasteurianus* (Suzuki *et al*., 1996a) bacteria.

Regulation of CcrM during the cell cycles of Alphaproteobacteria

As mentioned earlier, the activity of CcrM is cell cycle-regulated in *C. crescentus*, so that it happens only once DNA replication is finished in late pre-divisional cells. Interestingly, if CcrM is over-produced throughout the whole cell cycle, cells from *C. crescentus* and from the plant symbiont *Sinorhizobium meliloti* display cell division defects and over-initiate DNA replication (Wright *et al*., 1996; Wright *et al*., 1997). CcrM over-production also leads to nodulation defects with *Mesorhibozium loti* (Ichida *et al*., 2009) and to intracellular replication defects in the animal and human pathogen *Brucella abortus* (Robertson *et al*., 2000). All together, these observations demonstrate that the temporal control of CcrM activity is critical in many *Alphaproteobacteria*.

The *C. crescentus* CcrM protein is only detectable in pre-divisional cells due to tight transcriptional and post-transcriptional control mechanisms (Fig. 1). A common feature found in *C. crescentus* (Reisenauer *et al*., 1999), *B. abortus* (Bellefontaine *et al*., 2002; Francis *et al*., 2017), *S. meliloti* (Pini *et al*., 2015) and *Mesorhizobium huakuii* (Peng *et al*., 2014) is that the conserved CtrA regulator is responsible for the timed activation of *ccrM* transcription (Fig. 1B). This response regulator is the most abundant and active in swarmer and pre-divisional *C. crescentus* cells (Fig. 1A), where it controls the expression of hundreds of genes (Laub *et al*., 2002; Collier, 2016). In addition, the SciP regulator prevents the activation of *ccrM* by CtrA in swarmer cells (Fig. 1) (Gora *et al*., 2010; Tan *et al*., 2010; Collier, 2016). Noteworthy, there is also evidence suggesting that CcrM represses its own expression through the methylation of the *ccrM* promoter region (Fig. 1B) (C. M. Stephens *et al*., 1995). Once synthesized, CcrM is a particularly unstable protein degraded by the Lon protease (Fig. 1) (Wright *et al*., 1996), enabling variations of CcrM levels as *ccrM* transcription fluctuates during the *C. crescentus* cell cycle.

The variable essentiality of DNA MTases in Alphaproteobacteria

CcrM was the first DNA MTase identified as being potentially essential for the survival of at least certain *Alphaproteobacteria*, including *C. crescentus* (C. Stephens *et al*., 1996), *B. abortus* (Robertson *et al*., 2000), *S. meliloti* (Wright *et al*., 1997), *R. sphaeroides* (Burger *et al*., 2017), the plant pathogen *Agrobacterium tumefaciens* (Curtis and Brun, 2014) and the phototrophic environmental bacterium *Rhodopseudomonas palustris* (Pechter *et al*., 2015) cultivated in standard laboratory conditions. This finding was however challenged more recently concerning *C. crescentus.* Indeed, a *ccrM* deletion mutant could be constructed when *C. crescentus* was cultivated in slow growing conditions instead of fast growing conditions (Gonzalez and Collier, 2013). The resulting Δ*ccrM* cells were significantly elongated, but their viability was comparable to wild-type cells. Since then, CcrM was also found to be dispensable in a few other *Alphaproteobacteria*, including the environmental bacterium *Brevundimonas subvibrioides* (Curtis and Brun, 2014) and, most likely, the pollutant-degrading bacterium *Sphingomonas wittichii* (Roggo *et al*., 2013).

Besides *ccrM*, the gene encoding another putative orphan DNA methyltransferase from *R. palustris* has been recently shown to be un-disruptable during a transposon mutagenesis followed by deep sequencing (Tn-Seq) experiment. Interestingly, this gene shows similarities with the Dcm cytosine MTase that is mostly found in *Gammaproteobacteria* where it is usually dispensible (Pechter *et al*., 2015). This finding suggests that cytosine methylation can sometimes play essential roles in *Alphaproteobacteria.* This is however not always the case, since cytosine MTases found in *C. crescentus* and *B. subvibrioides* were shown to be dispensable using similar Tn-Seq experiments (Christen *et al*., 2011; Curtis and Brun, 2014).

Impact of CcrM-mediated methylation on the C. crescentus *transcriptome*

Methylated bases can serve as epigenetic signals influencing gene expression in prokaryotes and eukaryotes. Then, an attractive hypothesis is that many orphan DNA MTases found in bacteria, which have lost their use as ‘immigration controllers’ may instead be used to

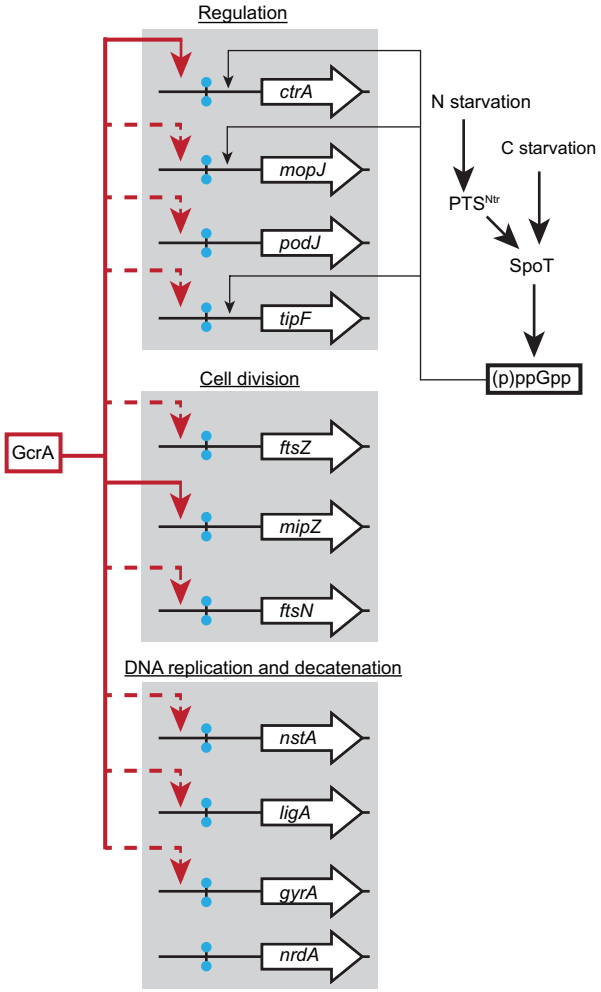

Fig. 2. Selection of genes activated by CcrM-mediated methylation in *C. crescentus*. Each gene included in this diagram is transcribed from a promoter methylated and activated by CcrM (Gonzalez *et al*., 2014). To simplify, a blue lollipop indicates that the promoter region carries one or more GANTC motifs. Many of these promoters are also activated by GcrA (Holtzendorff *et al*., 2004; Fioravanti *et al*., 2013; Haakonsen *et al*., 2015) or by the (p)ppGpp alarmone (Boutte and Crosson, 2011; Sanselicio & Viollier, 2015). Dashed red arrows indicate that binding of GcrA to the promoter region was detected *in vivo* during chromatin-immunoprecipitation assays described in (Fioravanti *et al*., 2013) or (Haakonsen *et al*., 2015). Red arrows indicate that the binding of purified GcrA onto the promoter was confirmed *in vitro* (Fioravanti *et al*., 2013). PTS^{Ntr} indicates the nitrogen-related phosphoenolpyruvate phosphotransferase system sensitive to glutamine levels (Ronneau *et al*., 2016). Note that this diagram does not include all the genes regulated by CcrM, GcrA or (p)ppGpp. [Colour figure can be viewed at wileyonlinelibrary.com]

modulate gene expression. *C. crescentus* CcrM is one of these examples. Indeed, an early study demonstrated that the activity of the *ctrAP1* promoter controlling the expression of the CtrA master regulator, is influenced by its methylation state (Fig. 2) (Reisenauer and Shapiro, 2002). Considering that GANTC motifs are under-represented in the genomes of bacteria that have a *ccrM* homolog but over-represented in intergenic regions in these genomes, CcrM-mediated DNA methylation could potentially have a much broader impact on gene expression in *Alphaproteobacteria* (Gonzalez *et al*., 2014). Proof of this came from a study comparing the transcriptomes of wild-type, Δ*ccrM* and *ccrM*-overexpressing *C. crescentus* cells (Gonzalez *et al*., 2014). Strikingly, the expression of ~10% of the *C. crescentus* genes appeared as modulated in response to the absence of *ccrM,* with a significant enrichment for genes that have minimum one GANTC motif in their promoter region and for genes whose expression is cell cycle-regulated or essential. Examples of genes fitting each of these criteria notably included genes encoding proteins involved in DNA replication (*ligA, gyrA* or *nrdA*), chromosome decatenation (*nstA*) and cell division (*ftsZ*, *mipZ* or *ftsN*), uncovering a strong connection between CcrM-mediated methylation and cell cycle progression (Fig. 2). A complementary study demonstrated that the activity of the *ftsZ* promoter is very low when its GANTC motif remains un-methylated, providing insufficient levels of FtsZ for the division of *C. crescentus* cells cultivated in fast-growing conditions (Gonzalez and Collier, 2013). Supporting this model, an experimental evolution approach showed that Δ*ccrM* cells must accumulate mutations that promote *ftsZ* transcription to survive under fast-growing conditions (Gonzalez and Collier, 2015).

Whether CcrM-mediated DNA methylation is required for the expression of cell division genes in other *Alphaproteobacteria* remains untested. It however appears that GANTC motifs are not systematically found in *ftsZ* promoter regions from other *Alphaproteobacteria*, even in bacteria where *ccrM* is essential for viability such as *Rhizobiales* (Gonzalez and Collier, 2013). This observation indicates that CcrM-mediated DNA methylation affects the transcription of other genes or other processes required for cell cycle progression in certain *Alphaproteobacteria*.

The coordinated action of CcrM and GcrA regulates gene expression as a function of the cell cycle

Analysis of the CcrM regulon in *C. crescentus* revealed that it contained a strong over-representation of GcrA-regulated genes (Gonzalez *et al*., 2014). This finding suggested that the GcrA global cell cycle regulator could act as an epigenetic regulator, whose activity may be affected by the methylation state of GANTC motifs in its target promoters (Fig. 2). In addition, the phenotypes of Δ*gcrA* and Δ*ccrM* mutants displayed striking similarities, being incapable of dividing under fast-growing conditions (Gonzalez and Collier, 2013, 2015; Murray *et al*., 2013). GcrA is a cell cycle-regulated transcriptional regulator that controls the expression of hundreds of genes when

it is present in S-phase cells (Holtzendorff *et al*., 2004; Fioravanti *et al*., 2013; Haakonsen *et al*., 2015) and it is often co-conserved with CcrM in *Alphaproteobacteria* (Brilli *et al*., 2010; Murray *et al*., 2013). Its mechanism of action remained unclear for over a decade after its discovery, in large part due to its unique domains that did not resemble canonical DNA binding domains. Still, several recent studies shed some light on how GcrA modulates gene transcription in *C. crescentus*. GcrA appears to be a dual function protein that can interact with the Sigma^{73} housekeeping factor (RpoD) loaded onto the RNA polymerase and that can bind to methylated promoter regions through an atypical but dedicated DNA binding domain (Fioravanti *et al*., 2013; Haakonsen *et al*., 2015; Wu *et al*., 2018). A crystal structure of this domain revealed the existence of two protein pockets that are probably responsible for the detection of m6A modifications (Wu *et al*., 2018). Consistent with its role as an epigenetic transcriptional regulator, purified GcrA appears to have more affinity for certain promoter regions when they are in a methylated, rather than in a non-methylated state (Fig. 2) (Fioravanti *et al*., 2013). Interestingly, the affinity of GcrA for hemi-methylated motifs on a few tested promoter regions (including the *mipZ* promoter) is dependent on the DNA strand that is methylated, suggesting that GcrA might regulate certain genes asymmetrically in early pre-divisional cells. Chromatin-immunoprecipitation followed by deep sequencing (ChIP-Seq) experiments confirmed that GcrA binds preferentially, but not exclusively, to Sigma^{73}-dependent promoters that harbor GANTC motifs *in vivo.* Also, GcrA has lower affinity for chromosomal regions in cells that lack the CcrM MTase, confirming the influence of GANTC methylation on GcrA activity (Fioravanti *et al*., 2013; Haakonsen *et al*., 2015). Overall, the current model, supported by the analysis of crystal structures of GcrA domains, is that GcrA preferentially activates Sigma^{73}-dependent promoters that harbor fully-methylated GANTC motifs through a direct interaction with the domain of Sigma^{73} that recognizes the -10 element of core promoters (Haakonsen *et al*., 2015; Wu *et al*., 2018). GcrA binding sites can be located in core promoter regions, but also at a significant distance, including downstream of transcriptional start sites (TSS) (Wu *et al*., 2018).

Considering that GcrA and Sigma^{73} are bound to the *ftsZ* promoter *in vivo* (Fioravanti *et al*., 2013; Haakonsen *et al*., 2015) and that the *ftsZ* promoter is activated in response to its methylation by CcrM (Gonzalez and Collier, 2013), it is likely that GcrA is the epigenetic regulator that activates *ftsZ* transcription in stalked and pre-divisional *C. crescentus* cells, although this was not directly tested *in vitro*. If this epigenetic mechanism controlling cell division is conserved in some other *Alphaproteobacteria*, it may explain why *gcrA* homologs sometimes appear to be required for the fitness of certain *Alphaproteobacteria* such as *R. palustris* (Pechter *et al*., 2015).

Interestingly, a connection between GcrA/CcrM-mediated regulation and (p)ppGpp mediated regulation has unexpectedly been uncovered in *C. crescentus* (Fig. 2) (Sanselicio and Viollier, 2015). (p)ppGpp is the alarmone that is produced in response to glucose and nitrogen starvation (Ronneau *et al*., 2016; Hallez *et al*., 2017;). It controls key cellular processes and regulates the expression of many genes in *C. crescentus* through mechanisms that are not yet fully understood (Boutte and Crosson, 2011; Gonzalez and Collier, 2014). Among these, several are co-regulated by the GcrA epigenetic regulator, such as *ctrA* and *mopJ* (encoding a pleiotropic regulator). Supporting this connection, it was shown that *gcrA* can be disrupted in fast-growing *C. crescentus* cells over-expressing (p)ppGpp (Haakonsen *et al*., 2015). This finding suggests that epigenetic control by GcrA/CcrM can be modulated in response to environmental conditions in *Alphaproteobacteria*, although the precise mechanism is still unknown.

Although the connection between GcrA and CcrM now appears as obvious, there are still many *C. crescentus* genes whose expression is different in Δ*ccrM* cells compared to wild-type cells, but that are not regulated by GcrA or that are not under the control of the housekeeping Sigma^{73} factor (Narayanan *et al*., 2015; Zhou *et al*., 2015; Adhikari and Curtis, 2016; Collier, 2016). There are also other *Alphaproteobacteria* that display an essential CcrM MTase but that lack a *gcrA* homolog, as it is the case for certain strains of *A. tumefaciens* (Curtis and Brun, 2014). These observations open up new research directions to better understand how epigenetic signals can affect gene expression and cellular processes in *Alphaproteobacteria*.

Hypomethylated DNA motifs in Alphaproteobacteria *and epigenetic patterns*

In *Gammaproteobacteria*, Dam-mediated regulation often depends on the protection of bi-stable promoters by specific transcription factors so that GATC motifs recognized by Dam are found non-methylated in a subset of cells within clonal populations. OxyR, Lrp and Fur, for example, have been shown to create such 'epigenetic patterns' mediating phase variations in *Escherichia coli* and *Salmonella enterica* (Wion and Casadesus, 2006). Thus, the presence of hypo-methylated promoter regions can be an indication that such promoters may be under epigenetic control mechanisms.

A first indication that epigenetic patterns may exist in *Alphaproteobacteria* came from a study characterizing the *M. loti* genome using restriction landmark genome scanning (RLGS) to detect GANTC motifs that are

hypo-methylated on this genome (Ichida *et al*., 2007). This study revealed that ~170 GANTC motifs were protected from methylation by CcrM in free-living cells, while much fewer were detectable in bacteroids within nodules. The fact that CcrM-dependent epigenetic patterns change during the establishment of symbiosis suggests that epigenetics may contribute to the control of plant-microbe interactions. Using the same method, CcrM-dependent epigenetic patterns were also identified on the genomes of two other plant interacting bacteria: *Bradyrhozobium japonicum* and *A. tumefaciens* (Ichida *et al*., 2007).

The more recent use of the SMRT sequencing technology also revealed the existence of hypo-methylated GANTC motifs on genomes of *Alphaproteobacteria* that encode a CcrM MTase. The estimation was that ~30–50 GANTC motifs are frequently hypo-methylated on the *C. crescentus and Celeribacter marinus* genomes (Kozdon *et al*., 2013; Gonzalez *et al*., 2014; Yang *et al*., 2016). A more quantitative method based on the use of a methylation sensitive restriction enzyme (*HinfI*) followed by Illumina-based deep-sequencing (REC-Seq) also confirmed the existence of hypo-methylated sites on the *C. crescentus* and the *S. meliloti* genomes (Ardissone *et al*., 2016). In *C. crescentus*, these epigenetic patterns do not change as a function of the cell cycle (Kozdon *et al*., 2013), but do respond to environmental changes such as phosphate availability (Ardissone *et al*., 2016) or switches from rich to minimal media (Gonzalez and Collier, unpublished observation). In *ccrM* over-expressing cells, very few of these GANTC motifs remain hypo-methylated (Gonzalez *et al*., 2014; Ardissone *et al*., 2016). This observation indicates that CcrM may compete with DNA binding proteins that occlude GANTC motifs from their methylation by CcrM in wild-type cells. A recent study has shown that the conserved MucR1 and MucR2 proteins, which regulate a transcriptional switch during the S-to-G1 phase transition in *C. crescentus* (Fumeaux *et al*., 2014), prevent the efficient methylation of about half of the hypo-methylated GANTC motifs of the *C. crescentus* genome, most likely through a competition with CcrM (Fig. 3) (Ardissone *et al*., 2016). Similar observations were made for the MucR regulator of *S. meliloti*. There is however no direct evidence that certain promoters regulated by MucR1/2 display bi-stable activities or are subject to phase variation as it is the case for some hypo-methylated promoters in *Gammaproteobacteria* (Casadesus and Low, 2013; Sanchez-Romero *et al*., 2015). Still, other hypo-methylated promoters may have such properties. Hypo-methylated GANTC motifs may also play a role in cell division, DNA replication or nodulation control in *Alphaproteobacteria* since *ccrM* overproduction most often leads to defects in these processes (Wright *et al*., 1997; Robertson *et al*., 2000; Kahng and Shapiro, 2001; Ichida *et al*., 2009). It is nevertheless worth noting

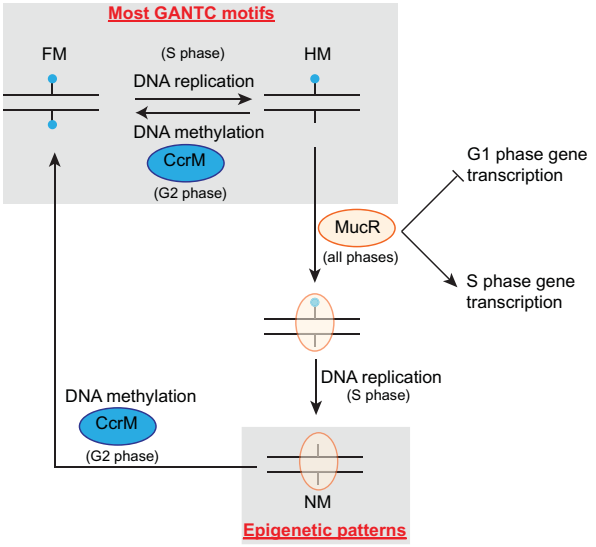

Fig. 3. Creation of epigenetic patterns by MucR in *Alphaproteobacteria*. MucR proteins participate in a S-to-G1 phase transcriptional switch controlling cell cycle progression (Fumeaux *et al*., 2014). Most GANTC motifs are either hemi- (HM) or fully-methylated (FM) in *C. crescentus* (Kozdon *et al*., 2013; Gonzalez *et al*., 2014). Binding of MucR to promoter regions that carry GANTC motifs can block their methylation by CcrM (Ardissone *et al*., 2016), generating hypo-methylated GANTC motifs that can become non-methylated (NM) after two rounds of replication, creating MucR-dependent epigenetic patterns. Such patterns have been directly detected in *C. crescentus* and *S. meliloti* (Ardissone *et al*., 2016). [Colour figure can be viewed at wileyonlinelibrary.com]

that hypo-methylated GANTC motifs are generally not conserved among different *Alphaproteobacteria* (Blow *et al*., 2016), suggesting that epigenetics controls different genes/processes in different bacteria or that the hypo-methylation of many of these motifs has no impact on gene expression.

Influence of DNA methylation on genome maintenance and evolution

Adenine methylation is sometimes involved in DNA mismatch repair (MMR) or in controlling the initiation of DNA replication in *Gammaproteobacteria* that encode a Dam methylase (Adhikari and Curtis, 2016). Such effects have been rarely tested directly in *Alphaproteobacteria*. It was however shown that a *C. crescentus* Δ*ccrM* strain has a spontaneous mutation rate that is similar to a wild-type strain, showing that CcrM-dependent methylation is not required for MMR in this *Alphaproteobacterium* (Gonzalez *et al*., 2014). Similarly, this mutant does not display obvious defects in DNA replication control. Then, there is, so far, no experimental indication that DNA methylation is involved in DNA replication control or in MMR in *Alphaproteobacteria*.

Although R/M systems usually reduce HGT (Fig. 4), R/M systems can sometimes be shut-down in response

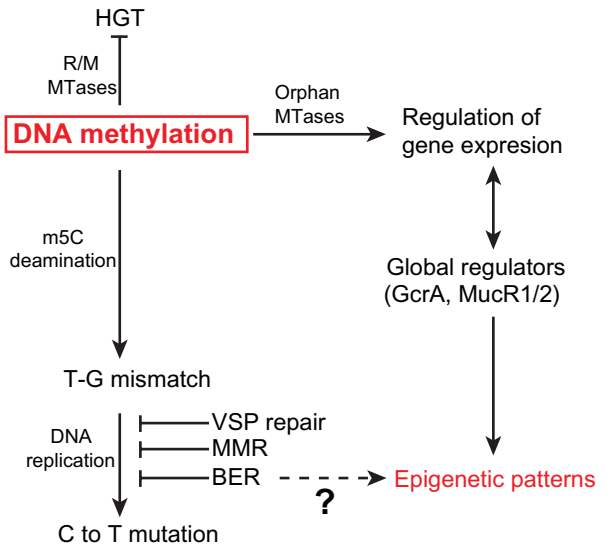

Fig. 4. Overview of the impact of DNA methylation on genome maintenance and expression in *Alphaproteobacteria*. HGT indicates horizontal gene transfers. R/M indicates restriction/modification systems. MTases indicates DNA methyl-transferases. VSP indicates very small patch repair. MMR indicates DNA mismatch repair. BER indicates base excision repair. HGT and mutations drive bacterial evolution, while DNA repair and R/M systems contribute to maintain genome integrity. [Colour figure can be viewed at wileyonlinelibrary.com]

to unfavorable environmental conditions in bacteria. This may promote DNA uptake and integration by recombination in these conditions, influencing HGT for adaptation (Jeltsch, 2003). In addition, it has been known for a long time that cytosine methylation by orphan or R/M MTases can have an impact on genome evolution. Indeed, methylated cytosines, although sometimes useful for the regulation of gene expression in bacteria, are mutation hot-spots due to their accidental deamination into thymines (Jeltsch, 2003; Marinus and Lobner-Olesen, 2014). TG mismatches must be repaired before the next round of DNA replication, otherwise they turn into undetectable C-to-T mutations. Bacteria can use up to three systems to detect and repair such DNA mismatches (Fig. 4): the highly conserved MMR process (Lenhart *et al.*, 2016), the bacterial very small patch (VSP) repair process (Marinus and Lobner-Olesen, 2014), and DNA glycosylases that can remove damaged bases by base excision repair (BER) (Jacobs and Schar, 2012). Interestingly, such DNA glycosylases might act as enzymes to de-methylate some bases in bacterial genomes in order to modulate epigenetic patterns (Fig. 4), as it is the case in eukaryotes (Schuermann *et al.*, 2016).

Perspectives

There are still many outstanding questions awaiting answers, notably to understand how m6A marks can influence gene expression, whether methylated cytosines can modulate gene expression in *Alphaproteobacteria*, whether there are phase variation mechanisms mediated by epigenetic switches in *Alphaproteobacteria*, what are the roles of poorly conserved orphan MTases and whether there are de-methylating processes in bacteria to modulate or re-set epigenetic patterns. With the advent of novel technologies to analyze epigenomes and their expression (Beaulaurier *et al.*, 2015; Blow *et al.*, 2016), we anticipate that major discoveries are likely to be made in the near future.

Acknowledgments

The Swiss National Science Foundation provided financial support (SNSF fellowships 31003A_173075 and CRSII3_160703).

References

Adhikari, S., and Curtis, P.D. (2016) DNA methyltransferases and epigenetic regulation in bacteria. *FEMS Microbiol Rev.* **40**: 575–591.

Albu, R.F., Jurkowski, T.P., and Jeltsch, A. (2012) The *Caulobacter crescentus* DNA-(adenine-N6)-methyltransferase CcrM methylates DNA in a distributive manner. *Nucleic Acids Res* **40**: 1708–1716.

Ardissone, S., Redder, P., Russo, G., Frandi, A., Fumeaux, C., *et al.* (2016) Cell Cycle Constraints and Environmental Control of Local DNA Hypomethylation in *alpha-Proteobacteria*. *PLoS Genet* **12**: e1006499.

Atack, J.M., Tan, A., Bakaletz, L.O., Jennings, M.P., and Seib, K.L. (2018) Phasevarions of Bacterial Pathogens: Methylomics Sheds New Light on Old Enemies. *Trends Microbiol* **26**: 715–726.

Beaulaurier, J., Zhang, X.S., Zhu, S., Sebra, R., Rosenbluh, C., Deikus, G., and Fang, G., (2015) Single molecule-level detection and long read-based phasing of epigenetic variations in bacterial methylomes. *Nat Commun* **6**: 7438.

Bellefontaine, A.F., Pierreux, C.E., Mertens, P., Vandenhaute, J., Letesson, J.J., and De Bolle, X. (2002) Plasticity of a transcriptional regulation network among *alpha-proteobacteria* is supported by the identification of CtrA targets in *Brucella abortus*. *Mol Microbiol* **43**: 945–960.

Berdis, A.J., Lee, I., Coward, J.K., Stephens, C., Wright, R., Shapiro, L., and Benkovic, S.J. (1998) A cell cycle-regulated adenine DNA methyltransferase from *Caulobacter crescentus* processively methylates GANTC sites on hemimethylated DNA. *Proc Natl Acad Sci U S A* **95**: 2874–2879.

Blow, M.J., Clark, T.A., Daum, C.G., Deutschbauer, A.M., Fomenkov, A., Fries, R., and Roberts, R.J., (2016) The Epigenomic Landscape of Prokaryotes. *PLoS Genet* **12**: e1005854.

Boutte, C.C., and Crosson, S. (2011) The complex logic of stringent response regulation in *Caulobacter crescentus*: starvation signalling in an oligotrophic environment. *Mol Microbiol* **80**: 695–714.

Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., and Biondi, E.G. (2010) The diversity and evolution of cell cycle regulation in *alpha-proteobacteria*: a comparative genomic analysis. *BMC Syst Biol* **4**: 52.

Burger, B.T., Imam, S., Scarborough, M.J., Noguera, D.R., and Donohue, T.J. (2017). Combining Genome-Scale Experimental and Computational Methods To Identify Essential Genes in Rhodobacter sphaeroides. *mSystems*, **2**. pii: e00015-17.

Casadesus, J., and Low, D.A. (2013) Programmed heterogeneity: epigenetic mechanisms in bacteria. *J Biol Chem* **288**: 13929–13935.

Christen, B., Abeliuk, E., Collier, J.M., Kalogeraki, V.S., Passarelli, B., Coller, J.A., and Shapiro, L., (2011) The essential genome of a bacterium. *Mol Syst Biol* **7**: 528.

Collier, J. (2009) Epigenetic regulation of the bacterial cell cycle. *Curr Opin Microbiol* **12**: 722–729.

Collier, J. (2016) Cell cycle control in *Alphaproteobacteria*. *Curr Opin Microbiol* **30**: 107–113.

Curtis, P.D., and Brun, Y.V. (2014) Identification of essential alphaproteobacterial genes reveals operational variability in conserved developmental and cell cycle systems. *Mol Microbiol* **93**: 713–735.

Davis, B.M., Chao, M.C., and Waldor, M.K. (2013) Entering the era of bacterial epigenomics with single molecule real time DNA sequencing. *Curr Opin Microbiol* **16**: 192–198.

Dziewit, L., Kuczkowska, K., Adamczuk, M., Radlinska, M., and Bartosik, D. (2011) Functional characterization of the type II PamI restriction-modification system derived from plasmid pAMI7 of *Paracoccus aminophilus JCM 7686*. *FEMS Microbiol Lett* **324**: 56–63.

Egger, G., Liang, G., Aparicio, A., and Jones, P.A. (2004) Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429**: 457–463.

Fioravanti, A., Fumeaux, C., Mohapatra, S.S., Bompard, C., Brilli, M., Frandi, A., and Biondi, E.G. (2013) DNA Binding of the Cell Cycle Transcriptional Regulator GcrA Depends on N6-Adenosine Methylation in *Caulobacter crescentus* and Other *Alphaproteobacteria*. *PLoS Genet* **9**: e1003541.

Flusberg, B.A., Webster, D.R., Lee, J.H., Travers, K.J., Olivares, E.C., Clark, T.A., and Turner, S.W. (2010) Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* **7**: 461–465.

Francis, N., Poncin, K., Fioravanti, A., Vassen, V., Willemart, K., Ong, T.A., and De Bolle, X. (2017) CtrA controls cell division and outer membrane composition of the pathogen *Brucella abortus*. *Mol Microbiol* **103**: 780–797.

Fumeaux, C., Radhakrishnan, S.K., Ardissone, S., Theraulaz, L., Frandi, A., Martins, D., and Viollier, P.H. (2014) Cell cycle transition from S-phase to G1 in *Caulobacter* is mediated by ancestral virulence regulators. *Nat Commun* **5**: 4081.

Garcia-Pastor, L., Puerta-Fernandez, E., and Casadesus, J. (2018) Bistability and phase variation in *Salmonella enterica*. *Biochim Biophys Acta*.

Gonzalez, D., and Collier, J. (2013) DNA methylation by CcrM activates the transcription of two genes required for the division of *Caulobacter crescentus*. *Mol Microbiol* **88**: 203–218.

Gonzalez, D., and Collier, J. (2014) Effects of (p)ppGpp on the progression of the cell cycle of *Caulobacter crescentus*. *J Bacteriol* **196**: 2514–2525.

Gonzalez, D., and Collier, J. (2015) Genomic Adaptations to the Loss of a Conserved Bacterial DNA Methyltransferase. *MBio* **6**: e00952.

Gonzalez, D., Kozdon, J.B., McAdams, H.H., Shapiro, L., and Collier, J. (2014) The functions of DNA methylation by CcrM in *Caulobacter crescentus*: a global approach. *Nucleic Acids Res* **42**: 3720–3735.

Gora, K.G., Tsokos, C.G., Chen, Y.E., Srinivasan, B.S., Perchuk, B.S., and Laub, M.T. (2010) A cell-type-specific protein-protein interaction modulates transcriptional activity of a master regulator in *Caulobacter crescentus*. *Mol Cell* **39**: 455–467.

Haakonsen, D.L., Yuan, A.H., and Laub, M.T. (2015) The bacterial cell cycle regulator GcrA is a sigma70 cofactor that drives gene expression from a subset of methylated promoters. *Genes Dev* **29**: 2272–2286.

Hallez, R., Delaby, M., Sanselicio, S., and Viollier, P.H. (2017) Hit the right spots: cell cycle control by phosphorylated guanosines in *alphaproteobacteria*. *Nat Rev Microbiol* **15**: 137–148.

Holtzendorff, J., Hung, D., Brende, P., Reisenauer, A., Viollier, P.H., McAdams, H.H., and Shapiro, L. (2004) Oscillating global regulators control the genetic circuit driving a bacterial cell cycle. *Science* **304**: 983–987.

Ichida, H., Matsuyama, T., Abe, T., and Koba, T. (2007) DNA adenine methylation changes dramatically during establishment of symbiosis. *FEBS J* **274**: 951–962.

Ichida, H., Yoneyama, K., Koba, T., and Abe, T. (2009) Epigenetic modification of rhizobial genome is essential for efficient nodulation. *Biochem Biophys Res Commun* **389**: 301–304.

Jacobs, A.L., and Schar, P. (2012) DNA glycosylases: in DNA repair and beyond. *Chromosoma* **121**: 1–20.

Jeltsch, A. (2003) Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems? *Gene* **317**: 13–16.

Kahng, L.S., and Shapiro, L. (2001) The CcrM DNA methyltransferase of *Agrobacterium tumefaciens* is essential, and its activity is cell cycle regulated. *J Bacteriol* **183**: 3065–3075.

Kerr, A.L., Jeon, Y.J., Svenson, C.J., Rogers, P.L., and Neilan, B.A. (2011) DNA restriction-modification systems in the ethanologen, *Zymomonas mobilis ZM4*. *Appl Microbiol Biotechnol* **89**: 761–769.

Kozdon, J.B., Melfi, M.D., Luong, K., Clark, T.A., Boitano, M., Wang, S., and McAdams, H.H. (2013) Global methylation state at base-pair resolution of the *Caulobacter* genome throughout the cell cycle. *Proc Natl Acad Sci U S A*.

Laub, M.T., Chen, S.L., Shapiro, L., and McAdams, H.H. (2002) Genes directly controlled by CtrA, a master regulator of the *Caulobacter* cell cycle. *Proc Natl Acad Sci U S A* **99**: 4632–4637.

Lenhart, J.S., Pillon, M.C., Guarne, A., Biteen, J.S., and Simmons, L.A. (2016) Mismatch repair in Gram-positive bacteria. *Res Microbiol* **167**: 4–12.

Loenen, W.A., Dryden, D.T., Raleigh, E.A., Wilson, G.G., and Murray, N.E. (2014) Highlights of the DNA cutters: a short history of the restriction enzymes. *Nucleic Acids Res* **42**: 3–19.

Maier, J.A., Albu, R.F., Jurkowski, T.P., and Jeltsch, A. (2015) Investigation of the C-terminal domain of the bacterial

DNA-(adenine N6)-methyltransferase CcrM. *Biochimie* **119**: 60–67.
Marinus, M.G., and Casadesus, J. (2009) Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol Rev* **33**: 488–503.
Marinus, M.G., and Lobner-Olesen, A. (2014). DNA Methylation. *EcoSal Plus*, **6**.
Murray, S.M., Panis, G., Fumeaux, C., Viollier, P.H., and Howard, M. (2013) Computational and genetic reduction of a cell cycle to its simplest, primordial components. *PLoS Biol* **11**: e1001749.
Narayanan, S., Janakiraman, B., Kumar, L., and Radhakrishnan, S.K. (2015) A cell cycle-controlled redox switch regulates the topoisomerase IV activity. *Genes Dev* **29**: 1175–1187.
Oliveira, P.H., Touchon, M., and Rocha, E.P. (2016) Regulation of genetic flux between bacteria by restriction-modification systems. *Proc Natl Acad Sci U S A* **113**: 5658–5663.
Pechter, K.B., Gallagher, L., Pyles, H., Manoil, C.S., and Harwood, C.S. (2015) Essential Genome of the Metabolically Versatile *Alphaproteobacterium Rhodopseudomonas palustris*. *J Bacteriol* **198**: 867–876.
Peng, J., Hao, B., Liu, L., Wang, S., Ma, B., Yang, Y., and Li, Y., (2014) RNA-Seq and microarrays analyses reveal global differential transcriptomes of *Mesorhizobium huakuii 7653R* between bacteroids and free-living cells. *PLoS One* **9**: e93626.
Pini, F., De Nisco, N.J., Ferri, L., Penterman, J., Fioravanti, A., Brilli, M., and Biondi, E.G. (2015) Cell Cycle Control by the Master Regulator CtrA in *Sinorhizobium meliloti*. *PLoS Genet* **11**: e1005232.
Reisenauer, A., Quon, K., and Shapiro, L. (1999) The CtrA response regulator mediates temporal control of gene expression during the *Caulobacter* cell cycle. *J Bacteriol* **181**: 2430–2439.
Reisenauer, A., and Shapiro, L. (2002) DNA methylation affects the cell cycle transcription of the CtrA global regulator in *Caulobacter*. *Embo J* **21**: 4969–4977.
Roberts, R.J., Vincze, T., Posfai, J., and Macelis, D. (2015) REBASE–a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res, 43(Database issue)*: D298–D299.
Robertson, G.T., Reisenauer, A., Wright, R., Jensen, R. B., Jensen, A., *et al*. (2000) The *Brucella abortus* CcrM DNA methyltransferase is essential for viability, and its overexpression attenuates intracellular replication in murine macrophages. *J Bacteriol* **182**: 3482–3489.
Rochepeau, P., Selinger, L.B., and Hynes, M.F. (1997) Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum VF39SM*. *Mol Gen Genet* **256**: 387–396.
Roggo, C., Coronado, E., Moreno-Forero, S.K., Harshman, K., Weber, J., and van der Meer, J.R. (2013) Genome-wide transposon insertion scanning of environmental survival functions in the polycyclic aromatic hydrocarbon degrading bacterium *Sphingomonas wittichii RW1*. *Environ Microbiol* **15**: 2681–2695.
Ronneau, S., Petit, K., De Bolle, X., and Hallez, R. (2016) Phosphotransferase-dependent accumulation of (p)ppGpp in response to glutamine deprivation in *Caulobacter crescentus*. *Nat Commun* **7**: 11423.
Sanchez-Romero, M.A., Cota, I., and Casadesus, J. (2015) DNA methylation in bacteria: from the methyl group to the methylome. *Curr Opin Microbiol* **25**: 9–16.
Sanselicio, S., and Viollier, P.H. (2015). Convergence of Alarmone and Cell Cycle Signaling from Trans-Encoded Sensory Domains. *MBio*, **6**.
Scavetta, R.D., Thomas, C.B., Walsh, M.A., Szegedi, S., Joachimiak, A., Gumport, R.I., and Churchill, M.E. (2000) Structure of RsrI methyltransferase, a member of the N6-adenine beta class of DNA methyltransferases. *Nucleic Acids Res* **28**: 3950–3961.
Schubeler, D. (2015) Function and information content of DNA methylation. *Nature* **517**: 321–326.
Schuermann, D., Weber, A.R., and Schar, P. (2016) Active DNA demethylation by DNA repair: Facts and uncertainties. *DNA Repair (Amst)* **44**: 92–102.
Shier, V.K., Hancey, C.J., and Benkovic, S.J. (2001) Identification of the active oligomeric state of an essential adenine DNA methyltransferase from *Caulobacter crescentus*. *J Biol Chem* **276**: 14744–14751.
Smith, Z.D., and Meissner, A. (2013) DNA methylation: roles in mammalian development. *Nat Rev Genet* **14**: 204–220.
Stephens, C., Reisenauer, A., Wright, R., and Shapiro, L. (1996) A cell cycle-regulated bacterial DNA methyltransferase is essential for viability. *Proc Natl Acad Sci U S A* **93**: 1210–1214.
Stephens, C.M., Zweiger, G., and Shapiro, L. (1995) Coordinate cell cycle control of a *Caulobacter* DNA methyltransferase and the flagellar genetic hierarchy. *J Bacteriol* **177**: 1662–1669.
Suzuki, T., Sugimoto, E., Tahara, Y., and Yamada, Y. (1996a) Cloning and nucleotide sequence of ApaLI restriction-modification system from *Acetobacter pasteurianus IFO 13753*. *Biosci Biotechnol Biochem* **60**: 1401–1405.
Suzuki, T., Sugimoto, E., Tahara, Y., and Yamada, Y. (1996b) Cloning and nucleotide sequence of the AgeI methylase gene from *Agrobacterium gelatinovorum IAM 12617*, a marine bacterium. *Biosci Biotechnol Biochem* **60**: 444–447.
Szegedi, S.S., and Gumport, R.I. (2000) DNA binding properties *in vivo* and target recognition domain sequence alignment analyses of wild-type and mutant RsrI [N6-adenine] DNA methyltransferases. *Nucleic Acids Res* **28**: 3972–3981.
Szegedi, S.S., Reich, N.O., and Gumport, R.I. (2000) Substrate binding in vitro and kinetics of RsrI [N6-adenine] DNA methyltransferase. *Nucleic Acids Res* **28**: 3962–3971.
Tan, M.H., Kozdon, J.B., Shen, X., Shapiro, L., and McAdams, H.H. (2010) An essential transcription factor, SciP, enhances robustness of *Caulobacter* cell cycle regulation. *Proc Natl Acad Sci U S A* **107**: 18985–18990.
Vasu, K., and Nagaraja, V. (2013) Diverse functions of restriction-modification systems in addition to cellular defense. *Microbiol Mol Biol Rev* **77**: 53–72.
Wion, D., and Casadesus, J. (2006) N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. *Nat Rev Microbiol* **4**: 183–192.
Woodcock, C.B., Yakubov, A.B., and Reich, N.O. (2017) *Caulobacter crescentus* Cell Cycle-Regulated DNA

Methyltransferase Uses a Novel Mechanism for Substrate Recognition. *Biochemistry* **56**: 3913–3922.

Wright, R., Stephens, C., and Shapiro, L. (1997) The CcrM DNA methyltransferase is widespread in the alpha subdivision of *proteobacteria*, and its essential functions are conserved in *Rhizobium meliloti* and *Caulobacter crescentus*. *J Bacteriol* **179**: 5869–5877.

Wright, R., Stephens, C., Zweiger, G., Shapiro, L., and Alley, M.R. (1996) *Caulobacter* Lon protease has a critical role in cell-cycle control of DNA methylation. *Genes Dev* **10**: 1532–1542.

Wu, X., Haakonsen, D.L., Sanderlin, A.G., Liu, Y.J., Shen, L., Zhuang, N., and Zhang, Y. (2018) Structural insights into the unique mechanism of transcription activation by *Caulobacter crescentus* GcrA. *Nucleic Acids Res*.

Yang, J.A., Kang, I., Moon, M., Ryu, U.C., Kwon, K.K., Cho, J.C., and Oh, H.M. (2016) Complete genome sequence of *Celeribacter marinus IMCC12053*(T), the host strain of marine bacteriophage P12053L. *Mar Genomics* **26**: 5–7.

Zhou, B., Schrader, J.M., Kalogeraki, V.S., Abeliuk, E., Dinh, C.B., Pham, J.Q., and Shapiro, L., (2015) The global regulatory architecture of transcription during the *Caulobacter* cell cycle. *PLoS Genet* **11**: e1004831.

Zweiger, G., Marczynski, G., and Shapiro, L. (1994) A *Caulobacter* DNA methyltransferase that functions only in the predivisional cell. *J Mol Biol* **235**: 472–485.