1-1-2013

Investigation Of Bacterial Rna-Directed Dna Methylation Via Dcm And Hfq

Dandan Li
Wayne State University,

INVESTIGATION OF BACTERIAL RNA-DIRECTED DNA METHYLATION VIA DCM AND HFQ

by

DANDAN LI

THESIS

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

In partial fulfillment of the requirements

For the degree of

MASTER OF SCIENCE

2013

MAJOR: CHEMISTRY (Biochemistry)

Approved by:

_____

Advisor Date

DEDICATION

To my Beloved Family.

ACKNOWLEDGEMENTS

I thank my PI Andrew Feig, for his constant guidance through my research. I appreciate the help and encouragement from Dr. Christine Chow and Dr. Matthew Allen throughout my graduate study. I also want to thank Dr. Ashok Bhagwat, who gave me much advice for my research, and Dr. David Rueda, Dr. Louis Romano, who I have been a TA with.

I appreciate all the help I got from my labmates and other students in our department. Taewoo Lee, who had done much of the preliminary research; and Rebecca Swett, who designed the bioinformatic search program for the genomic search; as well as Abeykoon J Rathnayake, who helped to construct pDCM-V5 plasmid. I would like to thank Honn lab members for allowing me to use their real-time PCR instrument and sonicator, and Daya Kharel and Jun Jiang in the Chow lab who helped with UV-melt experiments.

My thesis could not be done without the kind assistance from people outside of our department: Skye M. Barendt and Malcolm E. Winkler in Indiana University. They helped run the fluorescence microscopy experiments and gave me many suggestions. My husband, Chengzhi Xie, and his PI Yubin Ge, led me use their real-time PCR instrument and gave me valuable technical support on real-time PCR and Co-IP.

Special love goes to my mother Shilan Li, who brought me up well and taught me how to love, and my stepfather, Ron Kohler, who kindly supported us in many ways. I would like to show my respect to my parents-in-law, Yaqin Xuan

and Quan Xie for their great job on arising my husband. Last but not least, I thank my dear son, Leo Xie, who taught me the value of my life, who gave me enormous strength to face difficulties, and who filled my heart with sweetness and happiness.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION OF SMALL RNAS AND DNA METHYLATION IN GENE REGULATION

1.1 Small RNAs in Post-transcriptional Gene Regulation

Regulatory RNAs have garnered attention due to their significant roles in gene regulation and potential uses as therapeutic drugs (1-3). More and more regulatory small RNAs (sRNAs) and long non-coding RNAs were discovered to control gene expression (4,5). Regulatory RNAs can modulate gene expression at many biological steps, like post-transcriptional level and transcriptional level via chromatin remodeling (6).

The best-known examples of eukaryotic small regulatory RNAs are those involved in RNA interferences (RNAi), pathways that control post-transcriptional gene silencing (PTGS) (Figure 1.1A) (7-9). Briefly, long-dsRNAs are trimmed by the Dicer protein to form short interfering RNAs (siRNAs) or micro RNAs (miRNAs) (10,11). One strand of the small dsRNAs is degraded while the other strand binds to the RNA-induced silencing complex (RISC) (12). The short RNA then serves as a guide for the RISC to target specific mRNAs. Two possible outcomes can occur. One is mRNA degradation, which silences translation from the mRNA. This outcome is usually associated with perfect pairing between the miRNA and its target. The other outcome, occurring when base paring is imperfect, is translation inhibition (7,13).

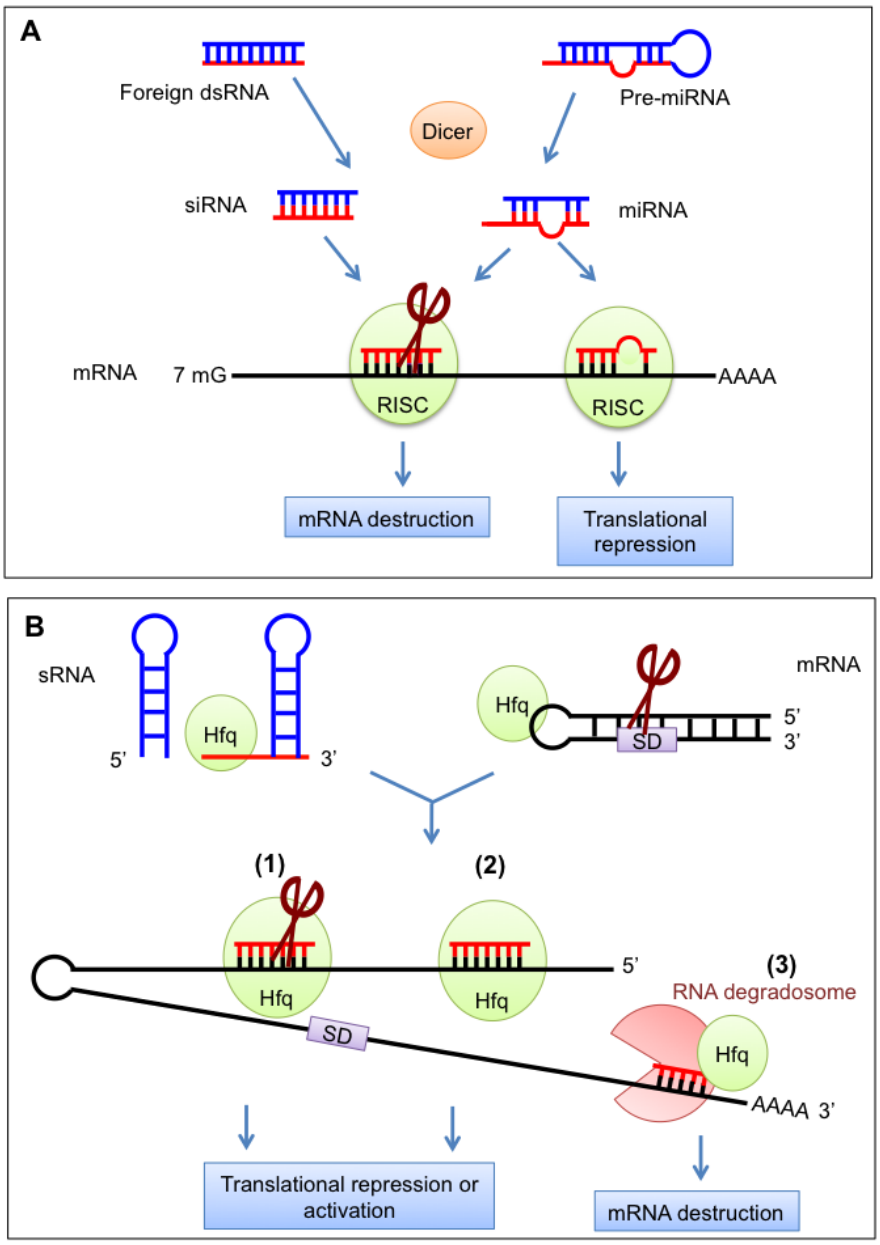


Fig 1.1. Comparison of eukaryotic RNAi mechanism and bacterial sRNA-mediated post-transcriptional gene regulation. (A) Eukaryotic RNAi pathways. Short siRNAs and miRNAs are cut from long dsRNAs by Dicer in the cytoplasm. One strand is selectively loaded on RISC, leading to post-transcriptional gene regulation. Perfect base pairing between a small RNA and an mRNA leads to the degradation of the mRNA, whereas an imperfect base pair typically results in translation inhibition. (B) Bacterial sRNA-mediated post-transcriptional gene regulation. Hfq binds to an sRNA or an mRNA, and brings those two close enough to base pair. Similar to RNAi, this leads to degradation of the mRNA (pathway 3) or inhibition of its translation (pathway 2). Hfq can also lead to up-regulation by either preventing mRNA degradation or freeing the Shine-Dalgarno sequence for ribosome binding (pathway 1).

Research on bacterial small RNAs in gene regulation is not as intense as RNAi; however, the fundamental similarities between the two systems and their essential functions require our attention to understand both of them comprehensively. Similar to eukaryotic short regulatory RNAs, bacterial sRNAs serve as guides for protein complexes to either degrade the mRNA targets, or inhibit their translation by initiating refolding of the mRNA (14). Besides down-regulation, bacterial sRNAs sometimes are involved in up-regulation of genes (15).

On the other hand, some necessary differences in bacteria make their systems simpler yet efficient. Unlike short miRNAs or siRNAs in eukaryotic cells, bacterial sRNAs are relatively long (typically 50-300nt), and have folded secondary structures (14,16,17). Usually, a strong G-C rich stem at the 3' end of an sRNA is necessary for Rho-independent transcriptional termination, and a single strand region in between hairpins is an important site for protein binding (Figure 1.1B) (18,19).

An RNA chaperone protein, Hfq, binds to the single strand region of sRNAs for further gene regulation (Figure 1.1B) (20). Hfq binds to mRNAs as tightly as it does to sRNAs (21,22), and facilitates paring between the two RNAs by increasing their local concentration, as well as inducing local structural changes (22,23).

Small regulatory RNAs have biological significance to quickly adjust gene expression patterns to cope with various stresses. Usually, bacterial regulatory sRNAs control gene regulation in pathways such as stress adaptation and

virulence (24,25). For example, when under cold stress and oxidative stress, *E.coli* can increase the transcription of small RNAs DsrA and ArcZ, respectively, to up-regulate a RNA polymerase sigma factor, RpoS, to cope with the stress (26). This adaption is much quicker than other regulation methods, and it also saves energy compared to constantly making and degrading proteins to respond to potentially short-term environmental changes.

Besides the differences mentioned earlier, some other contrasts of small RNAs mediated regulation in eukaryotes and bacteria are notable. First, eukaryotic RNA transcription occurs inside nuclei, and post-transcriptional regulatory functions of RNAs takes place in the cytoplasm. This spatial restriction limits the direct interaction between components of transcription and translation machineries, and only permits newly discovered co-transcripitonal gene regulation mechanism through small RNAs in nucleus (27). In bacteria, this separation is not present, which provides unique conditions that may lead to some potential relations, such as co-translation-and-transcriptional gene regulation. Second, bacterial sRNAs usually do not require further post-transcriptional processing to be active. This might be related to the aforementioned fact that bacterial sRNAs do not need to be transported outside of the nuclei as long dsRNAs. Omitting the transporting and dicing steps, bacterial sRNAs can respond to stresses faster than human sRNAs, which is essential for short-lived bacteria whose normal doubling time is about 20 minutes.

1.2 DNA Methylation in Gene Regulation

Besides regulatory RNAs, DNA methylation is another mechanism of gene expression found in both eukaryotic and bacterial systems. Two types of DNA methylation are known as 5-methyl cytosine (m5C) and N6-methyl adenine in bacteria, carried out by DNA cytosine methyltransferases (Dcm) and DNA adenine methyltransferase (Dam), respectively (Figure 1.2) (28). In contrast, higher eukaryotes use only m5C modification, carried DNA methyltransferases (DNMTs). Only a small number of m6A presences in lower eukaryotes(29). Dcm recognizes the 5'-CCWGG-3' sites, where W is A or T, and methylates the inner cytosine on both strands, whereas DNMTs recognize and methylate CpG islands; Dam methylates the adenine in the sequence of 5'-GATC-3' (28,30-32). Dcm, together with Dam, are involved in restriction-modification system (33).

Dynamic methylation cycles are controlled by the combination of methylation and demethylation. An unmethylated DNA can become fully methylated through *de novo* DNA methylation. A hemi-methylated DNA can be methylated through the mechanism of DNA methylation maintenance. In eukaryotes, the job of methylation is distributed to *de novo* methyltransferases DNMT3A and DNMT3B and maintenance methyltransferase DNMT1 (32), while in bacteria, both *de novo* methylation and maintenance methylation are both catalyzed by Dcm since it is the only DNA cytosine methyltransferase (34). There are two ways to demethylate a site of DNA modification. One way is through the passive demethylation pathway. Here an unmethylated daughter strand is

synthesized after one round of replication to make a hemi-methylated DNA. If replication occurs prior to the double strand methylation, the progeny of this cell

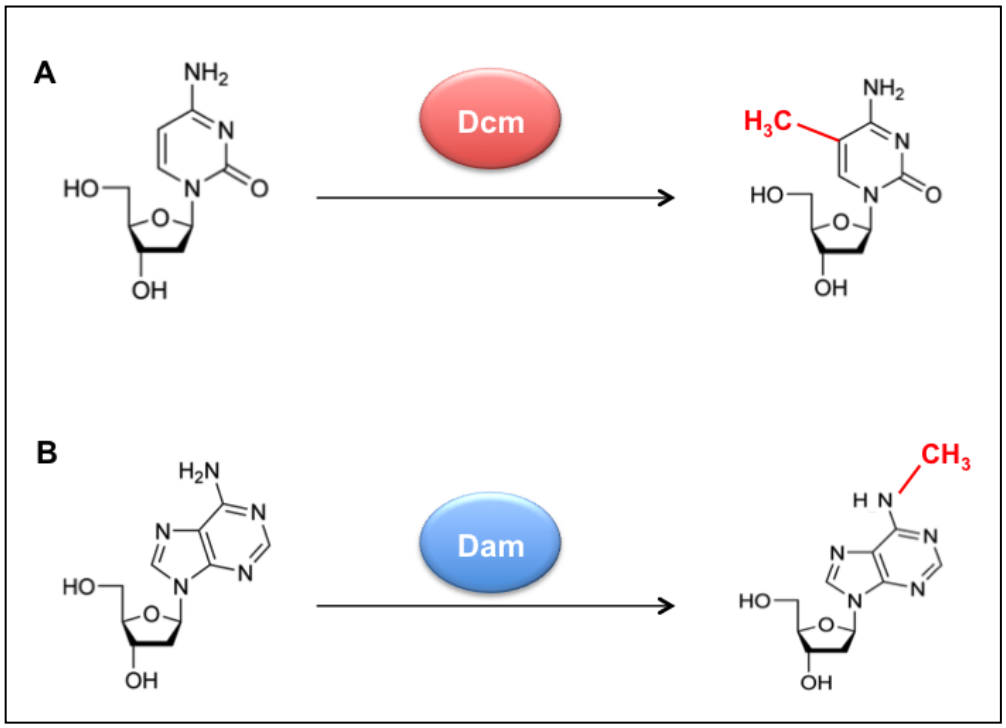


Fig 1.2. DNA Methylation reactions carried out by Dcm and Dam. (A) Cytosine is methylated to m5C by Dcm, and (B) adenine is methylated to m6A by Dam.

will lose the methylation mark. Bacteria have only passive demethylation. The second approach is the active removal of a methyl group from a cytosine, which has been recently found in mammals, but the mechanism is still unclear (35,36). One possibility is the oxidation of m5C by TET1 protein to produce the intermediate of 5-hydroxymethylcytosine (5hmC) (37), which can be further oxidized into 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (38). These modifications hinder the recognition by methylation maintenance machinery on the possible methylation sites (39). Alternatively, other proteins in the DNA deamination pathway could actively deaminate m5C and create T:G mutation, which will further convert back to C:G pair by DNA repair machinery (39,40).

Dam has essential functions in gene regulation affecting DNA replication, chromosome segregation, mismatch repair and transcriptional gene regulation (41-43). Transcriptional control by Dam is exemplified by the well-studied case of *pap* operon (44-46). PapA and PapB proteins in the *pap* operon are essential proteins for pili formation for *E.coli* to infect urinary tracts (47). A cluster of six conserved 5'-GATC-3' sites is found in their promoter region. Work of David Low showed that these sites can be evenly divided into proximal sites and distal sites, and the binding of the transcription factor leucine response protein (Lrp) at proximal sites inhibits the Dam methylation and blocks the RNA polymerase binding, so the genes cannot be transcribed and bacteria do not grow pili (41,44,48). Methylation on the proximal sites repels Lrp, leading it to bind at the unmethylated distal sites together with PapI protein; hence, the RNA polymerase can bind to the proximal sites and transcribe PapA and PapB (41,44,48). From

this example, we can learn that DNA adenine methylation is essential in transcriptional gene regulation, and this regulation can have either positive or negative influence in bacteria.

In contrast to Dam, little is known about Dcm-dependent gene regulation. Deletion of *dcm* causes no obvious defects (34). We would not expect to see a regulatory role for Dcm, since DNA cytosine methylation in eukaryotes is often involved in the chromatin remodeling to silence genes at transcriptional level, and plays a significant role in many pathways, such as genomic imprinting (49), development of cancer (50), and silencing repetitive elements (6). Two recent studies showed that Dcm may be involved in the gene regulation during the stationary phase (51,52). One of the studies showed that Dcm is highly conserved in many *E.coli* strains and it may affect ribosomal RNA expression (51). The other, a bisulfite sequencing study, showed that Dcm methylation is not saturated in the exponential growth phase, but it reaches saturation in the stationary phase; and a global gene regulator, RpoS, correlated with cytosine methylation level changes at its promoter during the transition to the stationary phase (52).

1.3 RNA-directed DNA Methylation in Eukaryotic Cells

As introduced earlier, regulatory RNAs, especially short RNAs, control gene expression at the post-transcriptional level, whereas DNA methylation controls gene at the transcriptional level through chromatin remodeling. While it may seem that these two regulatory pathways are distinct and separate at first

glance, actually they are quite intertwined in some cases of transcriptional gene silencing (TGS). The best-known mechanism for this is RNA-directed DNA methylation (RdDM). RdDM was first found in plants, and later similar pathways were also found in yeast and humans (53-55).

Much of our understanding of TGS mechanisms derives from pioneering studies done in plants (Figure 1.3). In *Arabidopsis,* the DNA methylation loci are often near repetitive sequences interspersed throughout the genome, and about one-third of the methylated DNA loci are close to siRNAs (56). The synthesis of siRNAs requires Pol IV, a DNA-dependent RNA polymerase (57,58). Pol IV transcribes long single stranded RNA, and then an RNA-dependent RNA polymerase RDR2 uses the long ssRNA as template to synthesize a double stranded RNA (57). Short RNAs are cleaved from the long dsRNA by a Dicer-like protein DCL3 (57), and then methylated at 3'-terminals by HUA ENHANCER 1 (HEN1) (59,60). The siRNAs then load onto AGO4 or AGO6 (61,62). Recent studies suggested the mechanism that siRNAs then base pair with an untranslated scaffold RNA (63). This scaffold RNA is transcribed from a non-coding region of a target gene by another plant-specific enzyme DNA-dependent RNA polymerase V (63), which would then recruit domain rearranged methyltransferase 2 (DRM2) to methylate the adjacent DNA. Other chromatin remodeling proteins, like chromatin-remodeling protein DRD1, are then recruited to cause chromatin remodeling and gene silencing (64,65). Unlike the RNAi pathways, the small regulatory RNAs in RdDM-dependent pathways regulate gene expression at the transcriptional level. This extension allows sRNAs to

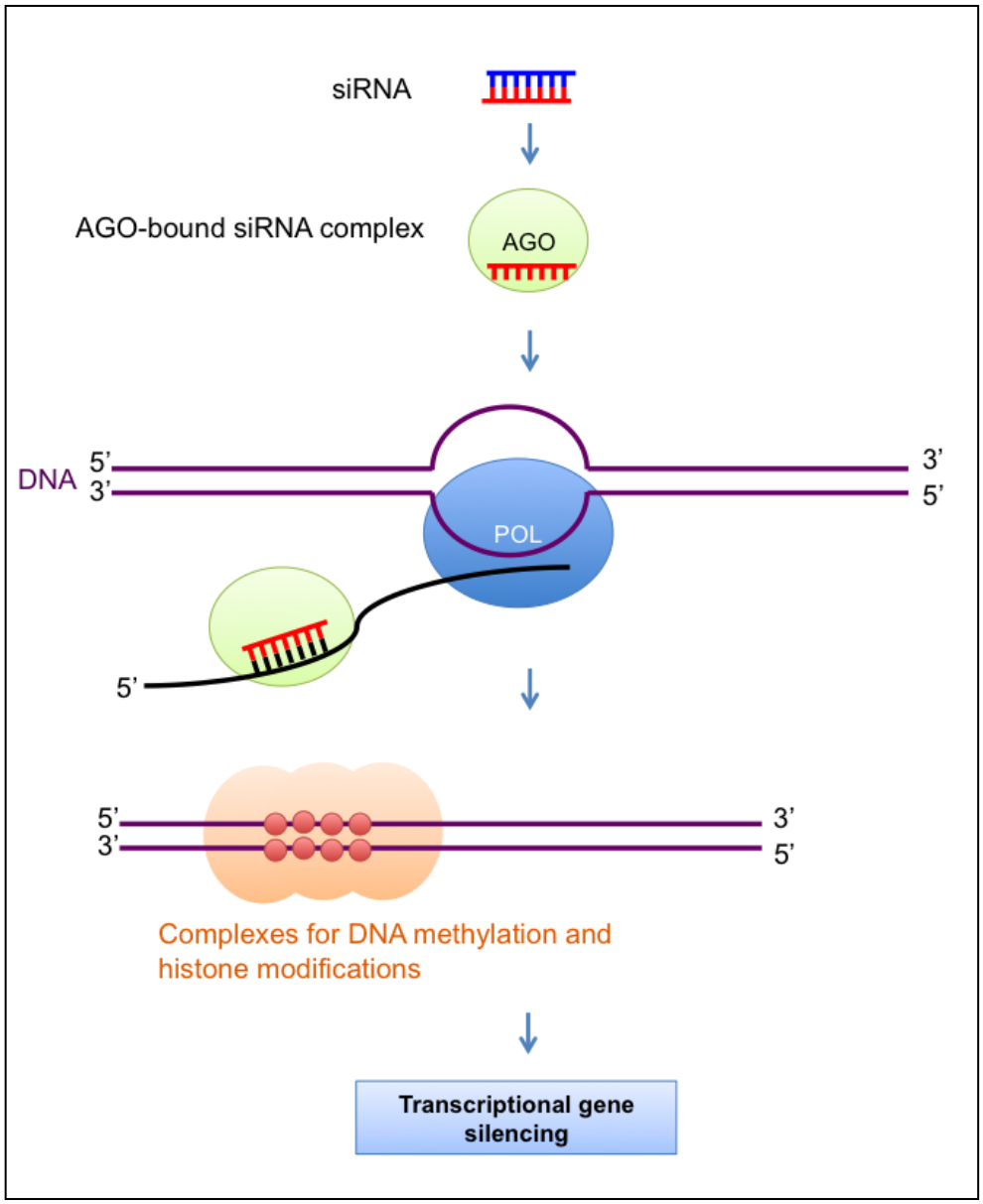


Figure 1.3. The current model for the RNA-directed DNA methylation pathway. The siRNAs, bound to AGO4 or AGO6, base pair with mRNA targets, which are mRNA scaffolds transcribed from nuclear RNA polymerase V or polymerase II (64,66,67). Consequently, this complex recruits DNA cytosine methyltransferase, and guides chromatin modifications at promoter regions, and then leads to subsequent transcriptional gene silencing.

have long-term influence compared to the PTGS pathways.

The lack of the long-term regulatory mechanisms of regulatory RNAs in bacteria motivated us to search for additional regulatory functions of sRNAs. In my study, I initiated the study of the potential regulatory roles of bacterial sRNA and Hfq at the transcriptional level through promoter methylation. The role of Dam in transcriptional gene regulation has been well studied, but the research on Dcm has been lagging behind. Therefore, we were interested in the investigation of the Dcm on such a promoter methylation event.

CHAPTER TWO

EXPERIMENTAL AND BIOINFORMATIC SEARCHES FOR POTENCIAL REGULATORY ROLES OF DCM

2.1 Introduction

Hfq is a key player in the network of sRNA-mediated gene regulation. Regulatory sRNAs are intertwined by Hfq to form a network of regulations under various growth conditions, and those RNAs compete against each other for the relatively limited amount of Hfq (14,68). It binds to different sRNAs expressed under specific stress conditions to modulate gene expressions in response to certain stresses (Figure 2.1) (20,69). Previous studies from our lab and other labs have shown that Hfq binds to groups of proteins that have distinct functions, like proteins in RNA degradasome and proteins involved in RNA modification (70). Based on these results, our lab hypothesized that Hfq would associate with different proteins and sRNAs to achieve distinct functions in various pathways.

To identify proteins that may compose the Hfq-sRNA-protein complexes, a sequential two-step purification method was used by Taewoo Lee, a previous member of our lab. Three known Hfq binding sRNAs, DsrA, SgrS and RydB were used as bait in the first purification column, and complexes associated with an sRNA were pulled out using an RNA affinity column that bound to an artificial tail of the sRNA. To further purify proteins that specifically associate with an sRNA and Hfq, a Co^{2+} column was then used to bind to the His-tagged Hfq. Identities of the proteins were assessed by mass spectrometry. The result showed that many

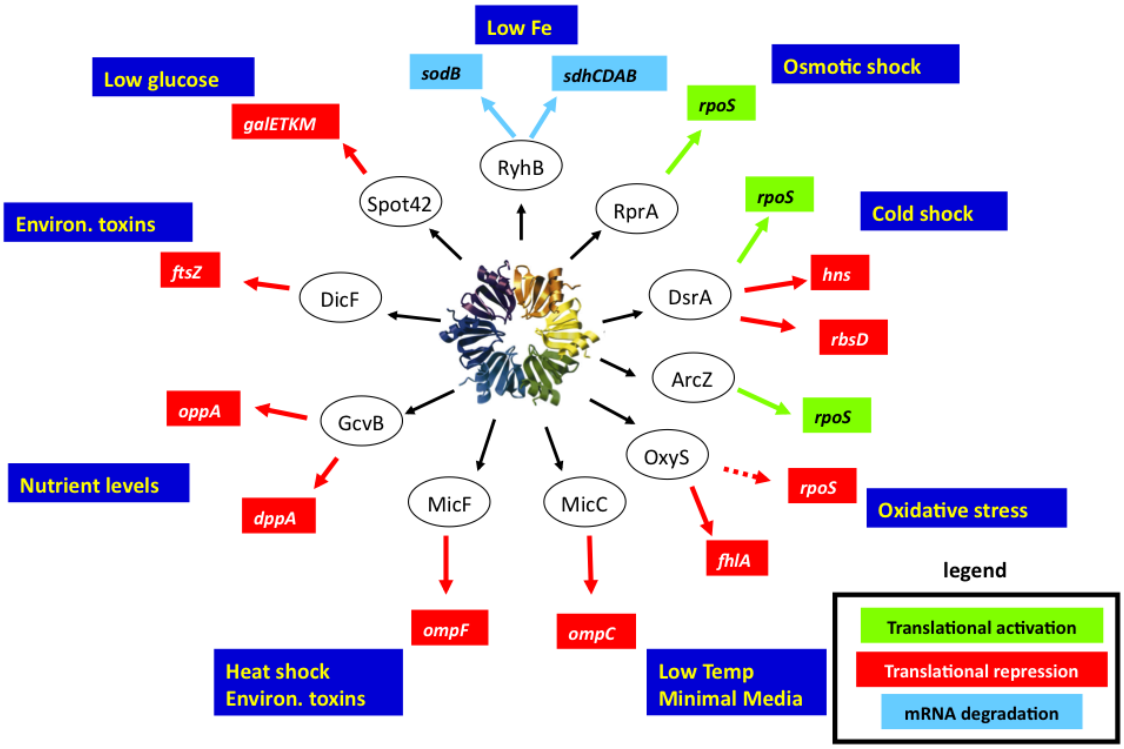


Fig 2.1. Hfq-centered, sRNA-mediated post-transcriptional gene regulation network. Structure of Hfq protein is shown in the center, and surrounding circles represent different sRNAs that expressed under different stress conditions as indicated. The mRNA targets can be degraded (light blue squares), translational activated (green squares) or repressed (red squares). The figure is an updated version of a figure prepared by Taewoo; I updated the regulatory network on *rpoS*.

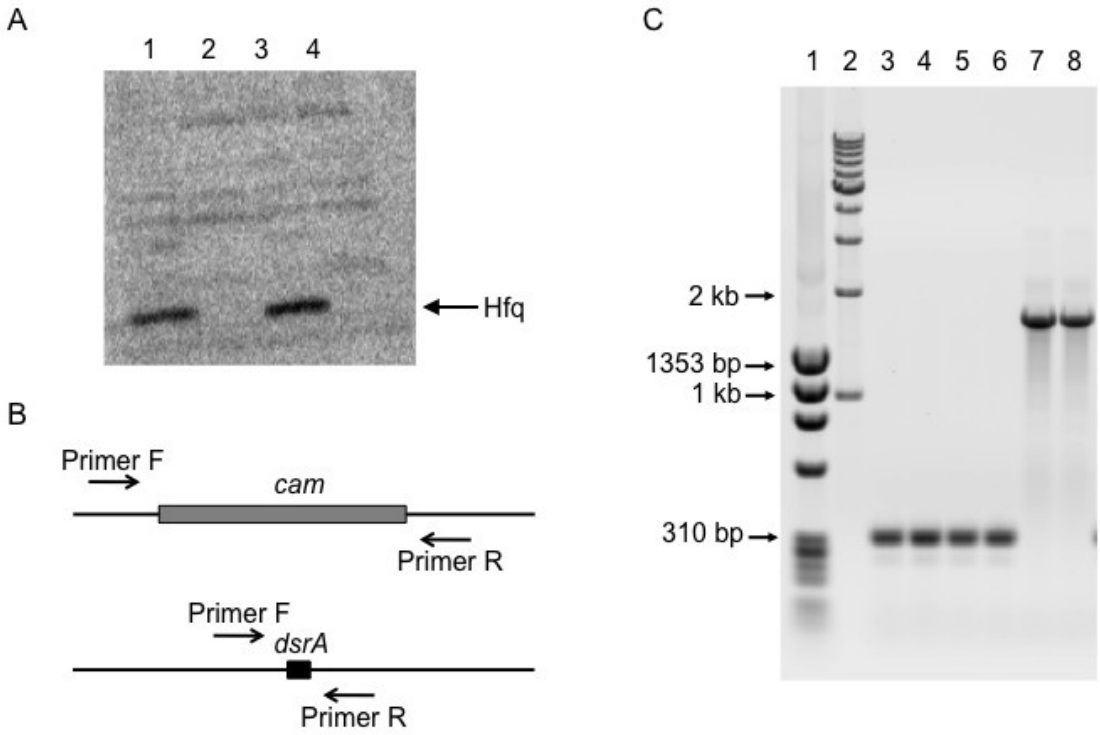


Fig. 2.2. Comfirmation of *hfq* and *dsrA* knockout strains. (A) Confirmation of *hfq* knockout using western blotting. Hfq proteins were present in GM30 (lane 1) and GM31 (lane 3), but not in DL2 (lane 2) and DL1 (lane 4). (B) Primers for confirmation of *dsrA* knockouts. Primer F and primer R target *dsrA* upstream and downstream regions, and would give a PCR product of 323 bp if *dsrA* gene is in the genome. A PCR product of 1729 bp would be detected if *cam* cassette replaces *dsrA* gene in the correct genome site. (C) PCR amplification to confirm *dsrA* knockout. The *dsrA* gene was intact in GM30 (lane 3), GM31 (lane 4), DL1 (lane 5), DL2 (lane 6), but was replaced by *cam* cassettes in DL7 (lane 7) and DL8 (lane 8).

previously known Hfq-binding proteins or RNA binding proteins were successfully pulled out, such as RNase E. Interestingly, a couple of new proteins were identified, one of whom was Dcm. The unexpected discovery of Dcm in the pool could be a promising result, if it was not a non-specific false positive result from the experiment.

The first question that needed to be answered was whether the potential Dcm and Hfq interaction in the previous pull down assay was biologically relevant. To answer this question was not easy, because we needed to select the biological phenotypes from a large number of possible phenotypes that may be affected by either Dcm or Hfq. Previous studies in our lab tried to find out if Hfq affected known Dcm functions, and find no defect cytosine methylation in vivo when Hfq was disrupted (Taewoo's thesis). I investigated the biological phenotypes of *dcm* and *hfq* knockout strains to find out unknown roles of Dcm on the known Hfq functions.

2.2 Results and Discussion

A classic role of Hfq is gene regulation under stress, so we investigated whether Dcm was also involved in such a gene regulation mechanism. These phenotypes can be tested using *dcm* and *hfq::kan* single knockout strains, and the double knockout strain; their effects can be tested by monitoring cell growth rates under different stress conditions. Several typical Hfq-related stresses were tested, including cold stress, oxidative stress, osmotic stress and nutrient starvation. Strains used in my research are listed in Table 2.1, and the knockout

Table 2.1 E.coli strains used in the study.

Strain	Genotypes	Source
GM30	F-, *thr⁻1, araC14, leuB6 (Am), fhuA31, lacY1, tsx⁻78, glnV44 (AS), galK2 (Oc), galT22, LAM⁻, hisG4 (Oc), rpsL136 (strR), xylA5, mtl⁻1, thi⁻1*	(40)
GM31	GM30, *dcm⁻6*	(40)
DL1	GM30, *dcm⁻6, hfq::kan*	This study
DL2	GM30, *hfq::kan*	This study
DL7	GM30, *dsrA::cam*	This study
DL8	GM30, *dcm⁻6, dsrA::cam*	This study
DL9	GM30, *dsrA* (cam cassette from DL7 was flipped out)	This study

strains constructed by me were confirmed using Western blots and PCR (Figure 2.2).

2.2.1 Dcm affects bacteria growth phenotypes in nutrient starvation, but not in oxidative stress or osmotic stress

Since Hfq was known to regulate gene expression under nutrient limitation conditions (72,73), I tested whether Dcm may affect cell growth phenotype through Hfq under this condition. The strains were grown in M9 media as described in the Methods and Materials section. As expected, the *hfq::kan* strain grew slower than the wt strain. Interestingly, the *dcm* strain grew much slower than the wt cells and the *hfq::kan* strain (Figure 2.3). This phenotype implies that Dcm functions in bacterial adaptation to nutrient starvation. This result supports our hypothesis that Dcm is somehow involved in gene regulation under non-optimal growth conditions.

To evaluate whether Hfq and Dcm work together to affect growth during nutrient starvation, we compared the effect of *hfq* in the strains with or without *dcm*. The difference in growth rate between the *dcm* and wt strains, and the difference between the *hfq::kan* and *dcm hfq::kan* double mutant strains were about the same (Figure 2.3). This result implies that Hfq may not contribute to this Dcm-mediated phenotype.

Oxidative stress and osmotic stress were also tested. Oxidative shock was introduced to exponentially grow cells by adding H_2O_2 as described in the Materials and Methods, and cell growth curves were followed for about 3 hours.

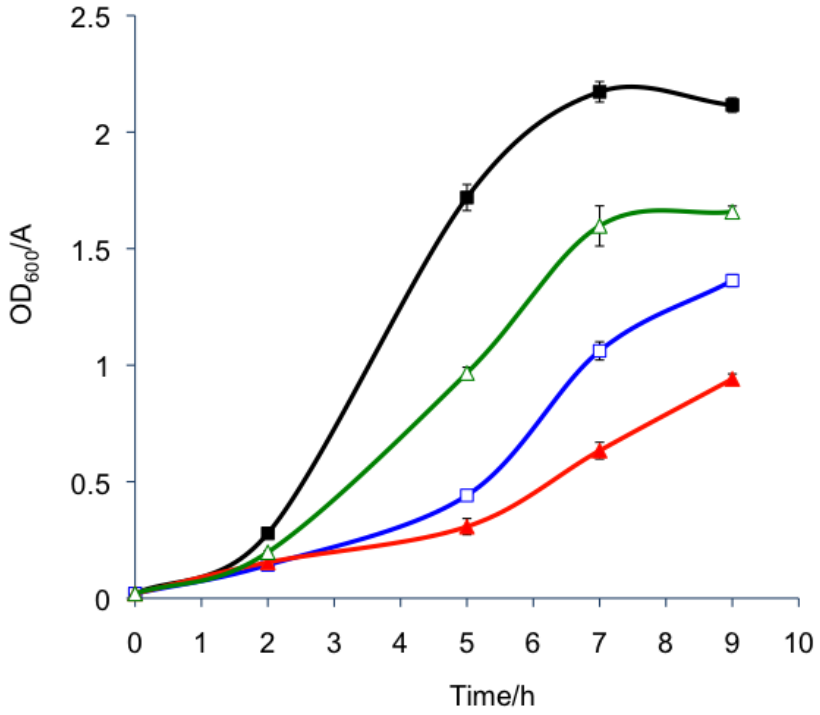


Fig 2.3. Growth curves under nutrient starvation. Growth curves of *E. coli* wt cells (black line with solid squares), *dcm* strains (blue line with open squares), *hfq::kan* stains (green line with open triangles) and the double mutant strain *dcm hfq::kan* (red line with solid triangles) grew in M9 media were measured periodically as indicated. One set of experiments was shown here, and the same experiment was repeated on a different day to confirm the result. Error bars represent ±SEM measured from three independent experiments on the same day.

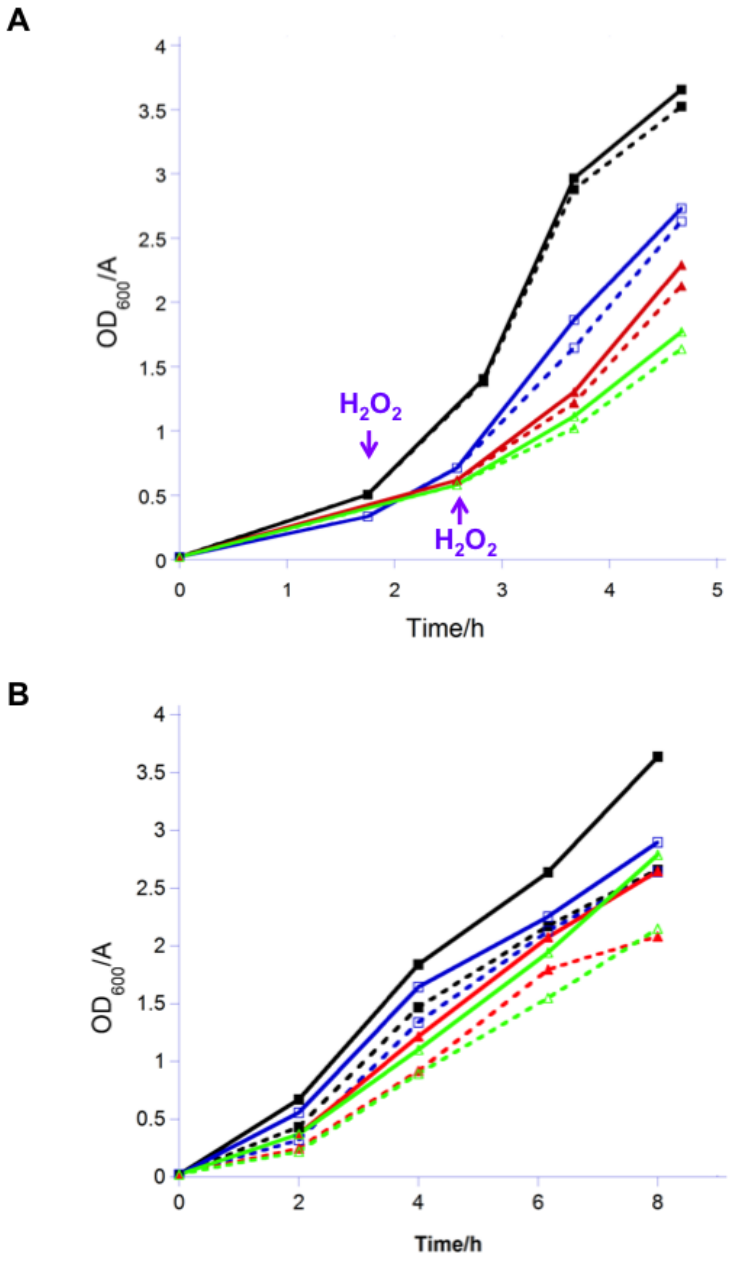


Fig 2.4. Dcm does not affect cell growth under oxidative stress and osmotic stress. Solid lines represent cells grown without stress conditions, and dashed lines represent cells undergo stresses. (A) H_2O_2 was added to cells at the OD_{600}=0.5-0.7 as indicated with arrows on the plot: the first arrow before 2 hour represent the point that H_2O_2 was added to the wt cells, and the second arrow is the point that H_2O_2 was added to the other cell cultures. (B) NaCl was added to cell cultures at time zero. One set of each experiment was shown here, and the same experiment was repeated on different days to confirm the result.

The results showed no significant effect of the oxidative shock on the cell growth rates (Figure 2.4A). Osmotic stress was applied to exponential cells by adding NaCl to the LB media. Similarly, no significant effect was observed as well (Figure 2.4B). Thus, Dcm may not have an obvious role in growth rate in these two stress conditions.

2.2.2 Hfq and Dcm together affect phenotypes under cold stress

Since Hfq did not have an obvious effect on the Dcm-mediated stress response in above conditions tested, I then continued the search under the cold shock. Growth rates were measured in pre-warmed LB media at 30°C (cold stress) and 37°C. The *dcm hfq::kan* double knockout strain grew much faster than the *hfq::kan* strain at 30°C (Figure 2.5). Thus, the slow growth defect in the absence of Hfq was largely recovered by further disruption of Dcm at cold temperature. This phenotype supports the hypothesis that Dcm and Hfq have synergistic effects on cell growth as if the phenotype requires both genes to fulfill their individual roles in the stress response pathway.

Another phenotype examined relates to the formation of minicells under cold stress. Minicells occur due to defects in DNA segregation after replication, resulting in polar cell division such that one of the daughter cells lacks its normal complement of genomic DNA (74-76). To test whether Dcm might be linked to this phenotype, fluorescence microscopy was used to observe cell morphology in stationary cells grown at 30°C in M9 medium, conditions where minicells were previously reported to occur (77). Minicells comprised about 6% of the population

in the *hfq::kan* background (Figure 2.6), similar to previous studies (77), while minicells frequency in *dcm* cells (0.3%) was comparable to wildtype (0.5%). Compared to the *hfq*::*kan* or *dcm* single mutants, the *dcm hfq::kan* mutant produced even more minicells (12%) than the *hfq::kan* mutant (arrows, Figure 2.5D). This study also revealed an additional unexpected phenotype, *dcm hfq::kan* double mutant *E.coli* exhibited marked defects in cell division, including the formations of elongated cells containing diffuse nucleoids (Figure 2.6). The surprising effect of Dcm and Hfq on nucleoid structure may indicate that Dcm and Hfq work together directly or indirectly in regulatory circuits that affect cell division and nucleoid packaging.

Note that the minicells were produced under conditions of both the cold stress and nutrient starvation. Further experiments are needed to elucidate the role of Dcm and Hfq in individual stress conditions. For my study here, the goal of searching for a biologically relevant phenotype controlled by both Dcm and Hfq was achieved in the above experiments.

By comparing phenotypes tested under different stress conditions, several conclusions can be made. First, Dcm has effects on cell growth rate under different stress conditions, especially under nutrient starvation. Second, Dcm and Hfq together have roles in cell growth and cell division under cold stress. Third, Dcm is not involved in all the Hfq-mediated stress conditions, but instead its role is neglected to a subset of these pathways.

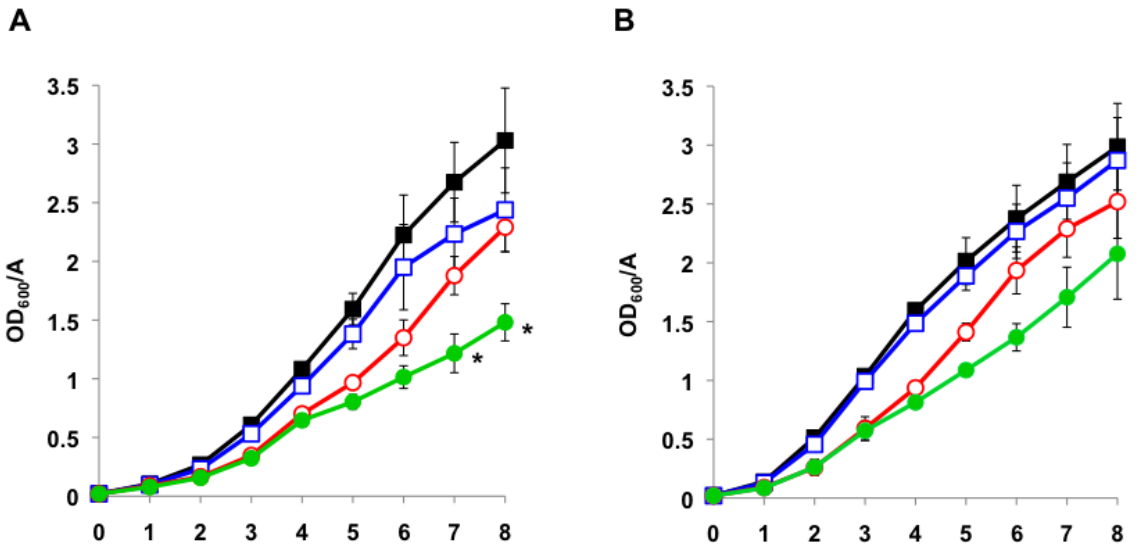


Fig. 2.5. Dcm and Hfq had synergistic phenotypes under cold stress. Growth rates of wild type strain (solid squares and black lines), *dcm* (open squares and blue lines), *hfq* (solid circles and green lines) and double mutant *dcm hfq::kan* (open circles and red lines) strains were measured in pre-warmed LB media with proper antibiotics. OD_{600} values were recorded every hour in (A) 30°C and (B) 37 °C cultures. Error bars represent ±SEM measured from three independent experiments on different days. Significant differences between *hfq* knockout and *dcm hfq::kan* double mutant were labeled (*$p \leq 0.05$) as measured by two-tailed t tests.

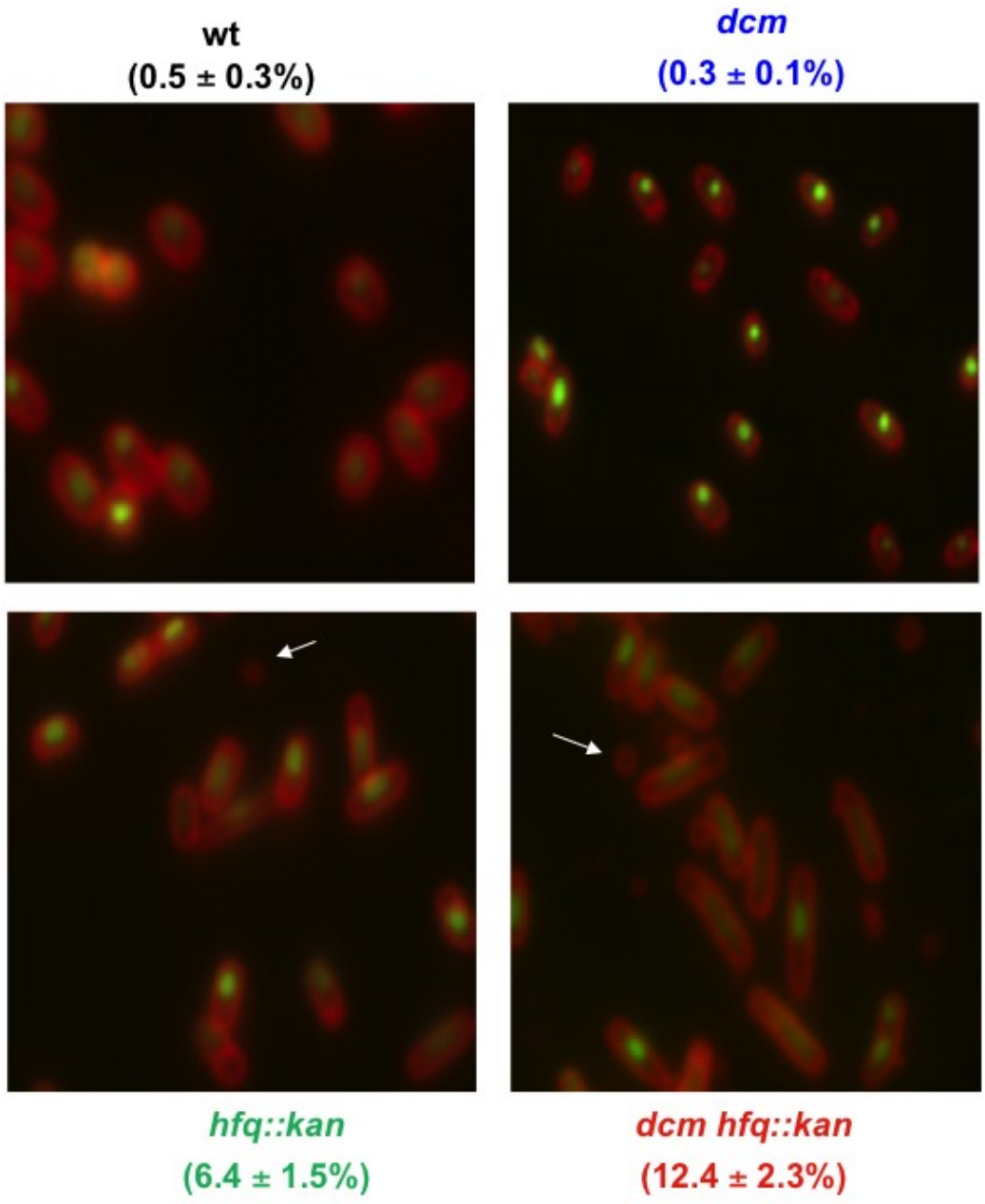


Fig 2.6. Minicell productions under cold stress. Absence of Dcm and Hfq increases minicell production and cell division in stationary-phase cells in M9 medium at 30°C. Cells were stained with FM4-64 (membrane) and DAPI (DNA nucleoids). Representative images from one experiment are shown, and minicells are pointed out by arrows. Percentages of minicells shown in parentheses are averages from three replicate experiments.

2.2.3 Bioinformatic search for potential target genes

To further understand the mechanisms leading to the phenotypes observed, I looked for inspiration from DNA methylation in other biological systems. By analogy to RdDM in eukaryotic transcriptional gene regulation and Dam in bacterial transcriptional gene regulations, we hypothesized that Dcm might regulate transcription by altering DNA cytosine methylation patterns near promoter sequences. It has been observed previously that clusters of 5'-GATC-3' lay in promoters of several genes that regulated by Dam (44,46), and the unmethylated CpG islands locate near promoter regions of genes that undergo regulation by DNMTs in eukaryotes (78). If Dcm functions similarly to DNMTs and Dam in transcriptional gene regulation, then a similar cluster of 5'-CCWGG-3' sites around promoter regions of genes would be seen, and most of the sites would be unmethylated. Thus, a genome-wide search for clusters of 5'-CCWGG-3' sites in *E.coli* was carried out to answer the question of whether the distribution of the CCWGG sites are more abundant near promoters than other genomic regions.

Three groups of data were collected for comparison. The first group was the theoretically random distribution of the number of CCWGG sites in 700bp windows on the genome; the second group was the actual number of the sites in *E.coli* K-12 MG1655 strain in 700bp windows; the last group was the actual number of the sites in between 500bp upstream of translation start site of a gene and 200bp downstream of the translation start site (Figure 2.7A). The reason we selected translation start sites was that more translation start sites were

annotated in the *E.coli* genome whereas most of the transcription start sites were not. Thus, the translation start sites were selected as an approximation of transcription start sites. This approximation, however, ignored the fact that bacteria have the polycistronic mRNA carrying multiple open reading frames. So the result of the bioinformatic search was manually evaluated in the context of operons, or polycistronic units.

The results are surprising. First, the actual 5'-CCWGG-3' sites were not randomly distributed, but instead were significantly more conserved than one would expect for a five base pair sequence (Figure 2.7B). This accumulation of CCWGG sites may indicate that Dcm is important for bacterial survival during evolution. One possibility is related to the role of Dam and Dcm in restriction-modification system. It is also possible that Dcm is significant in the adaption to changing environmental stresses. The restriction-modification system does not require localization of the sites, whereas gene regulation relates to the clusters of 5'-CCWGG-3' sites.

Second, the number of 5'-CCWGG-3' sites was not significantly higher in promoter regions compared to random windows. This finding seems to disfavor our hypothesis that Dcm regulates transcriptional gene regulation as a common strategy like DNMTs. Unlike eukaryotes, bacteria could distribute the gene regulation tasks to two DNA methyltransferases, Dam and Dcm, which could result in decreased number of genes regulated by either Dam and Dcm. However, in light of our findings on Dcm and Hfq associated phenotypes under

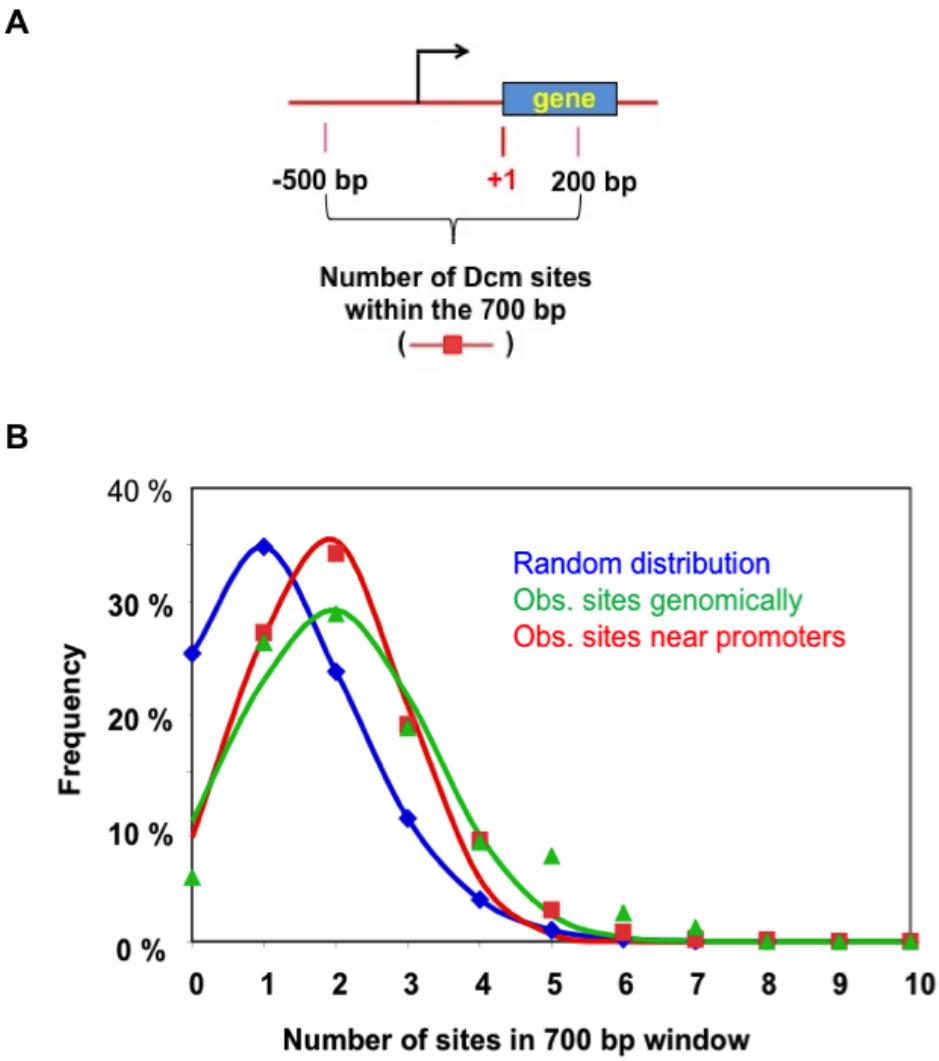


Fig 2.7. Bioinformatic search for the distribution of 5'-CCWGG-3' sites on *E.coli* genome. (A) A 700 bp window near translation start site of a gene was shown as described in main text. (B) Frequencies of the occurrence of 5'-CCWGG-3' sites were plotted against the number of sites in 700 bp windows.

different stress conditions, and the fact that the lack of Dcm did not provoke a growth defect in normal condition (34), we would not anticipate that Dcm affects a large number of genes. A more realistic expectation is that Dcm may be involved in the regulation of a few genes that are related to stress responses.

Thus, we next focused on those genes that were responsible for stress adaptation and contain high number of 5'-CCWGG-3' sites. An interesting observation from our search showed that many stress responsive genes were membrane-binding proteins, such as the MarA, MarB, FtsK and SstT (Figure 2.8). The *marRAB* operon is involved in the response against multiple antibiotic stress by expressing MarA and MarB to form multidrug efflux system, and a protein product MarR is a transcription factor that binds to the promoter to repress the transcription (79). There are two 5'-CCWGG-3' sites lie in the center of the MarR binding sites, but no previous researches has looked for possible Dcm methylation here and subsequent gene regulation (Figure 2.8A). FtsK is an essential protein in cell division, and its transcription can be repressed by the binding of protein LexA (80). The upstream of *ftsK* is the adenine methylation sensitive transcriptional factor Lrp. Clusters of 5'-CCWGG-3' sites locate in its promoter region, but no known proteins bind to this area (Figure 2.8B). SstT is a membrane protein involved in the sodium ion coupled Ser/Thr transportation through cell membranes (81). Small RNA GcvB negatively regulates *sstT* translation (82), and clusters of 5'-CCWGG-3' sites are also in close proximity (Figure 2.8C). All the above observations may indicate a regulatory role of Dcm on the membrane binding proteins under stress conditions.

In summary, by comparing cell growth rate and minicell production, I supported our hypothesis that Dcm has functions in response to stress, and some of its functions are related to Hfq as seen under cold stress. I was encouraged by these phenotypes to carry out a bioinformatic search for potential genes that may be responsible for the Dcm and Hfq functions. That search revealed some interesting potential targets that contain clusters of 5'-CCWGG-3' sites and are involved in stress adaptation. Our phenotype studies on the gene regulatory role of Dcm are in line with other recent researches on Dcm as introduced in Chapter One (51,52). Together, we related Dcm with Hfq in bacterial gene regulation, especially under stress conditions. In the next chapter, I further look into detailed mechanism in such a regulation.

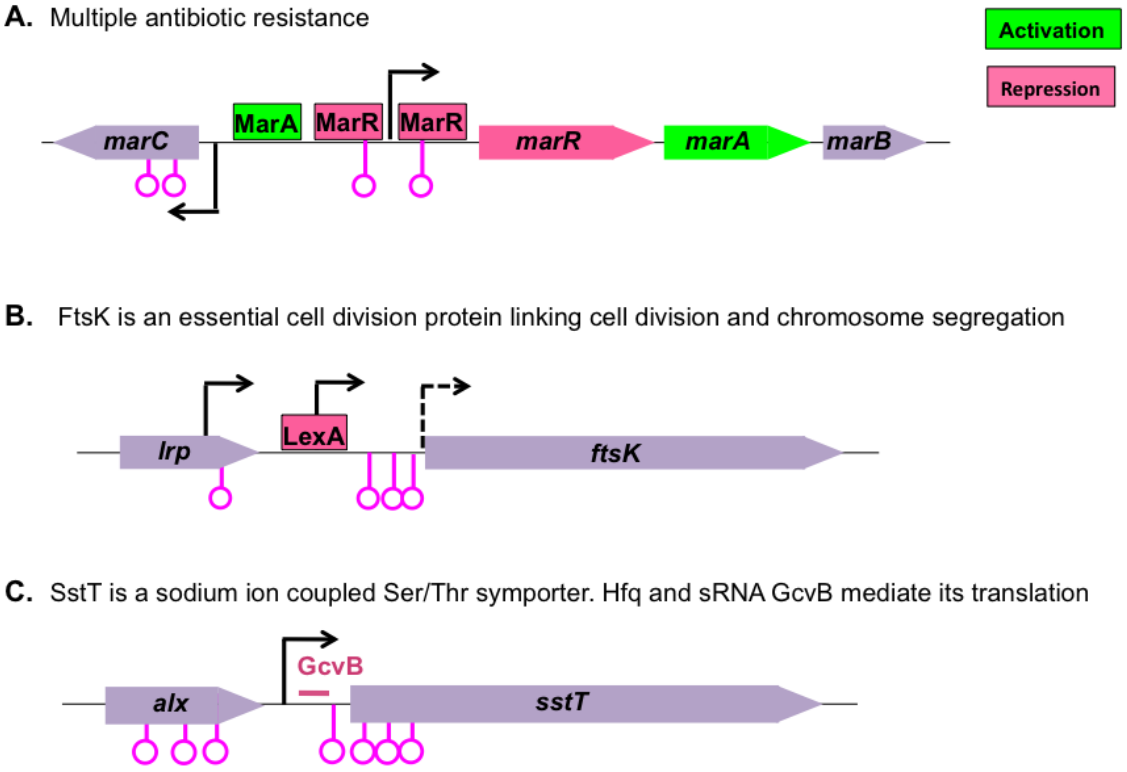


Fig 2.8. Some interesting targets from the bioinformatic search. Pink sticks and balls represent 5'-CCWGG-3' sites. (A) The Mar operon involves in the response to multiple antibiotic stress. (B) FtsK is an essential protein in cell division, and its transcription can be repressed by the binding of protein LexA. The upstream of *ftsK* is the adenine methylation sensitive transcriptional factor Lrp. Clusters of 5'-CCWGG-3' sites lay in its promoter region, but no known proteins bind to this area. (C) Small RNA GcvB negatively regulates *sstT* translation, and clusters of 5'-CCWGG-3' sites are also in close approximation.

2.3 Material and Methods

2.3.1 Genomic knock out of *hfq* and *dsrA*

WT and *dcm* mutant *E.coli* strains were gifts from Dr. Ashok Bhagwat's lab. Isogenic stains DL1 (*dcm hfq::kan*), DL2 (*hfq::kan*), DL7 (*dsrA::cam*) and DL8 (*dcm dsrA::cam*) were constructed by homologous recombination following the manufacture's protocol with minor modifications (Quick & Easy Conditional Knockout Kit, Gene Bridges). Briefly, electro-competent GM30 and GM31 cells were prepared by washing mid-log cell pellets with ice-cold water two to three times. Thermal-control recombinase expression plasmid pRedET (ampicillin resistance) was transformed into both GM30 and GM31 strains, and allowed to undergo replication at 30 °C overnight. Single colonies containing pRedET were selected, and recombinase expression was induced by raising the temperature to 37°C, and by adding L-arabinose to a final concentration of 0.3% - 0.4%. Linear DNA fragments containing either chloramphenicol or kanamycin resistant in FRT cassettes flanked by sequences from either *dsrA* or *hfq* were amplified using primers containing the *dsrA* or *hfq* flanking sequences.

One modification in this step was the use of 1% DMSO during the PCR of chloramphenicol cassette to overcome the difficulty of GC-rich manufacturer suggested primers in PCR. Another modification was that the suggested primer sequences were shortened to remove the GC rich 3'-tails for *hfq::kan* knockout strain construction. The linear PCR products were purified with a PCR clean up kit (Qiagen), and electroporated into above strains containing pRedET. Due to the fact that cells lacking Hfq or DsrA grow much slower after shock, incubation

times at 37 °C had to be adjusted in a strain dependent fashion to allow enough time for recombination.

2.3.2 Growth curve measurements

Glycerol cell stocks were streaked onto LB agar plates with proper antibiotics. Triplicates were done by selecting three colonies from each strain on fresh plates, and then grew in LB media with antibiotics at 37 ºC overnight. Cell cultures were diluted into fresh 37 ºC LB media with 100 μg/mL streptomycin to grow to exponential phase, and then diluted into pre-warmed (30 °C, 37 °C and 42 °C) 10 ml media to OD_{600} = 0.02. Cultures were then grown at 30 °C, 37 °C and 42 °C respectively. OD_{600} values were measured every hour using UV/VIS spectrometer (Perkin Elmer MBA 2000). Average OD_{600} values were calculated from triplicate cultures under the same condition. Biological replicates of growth curve experiments were repeated three times.

M9 media was prepared as in previous studies (77). Briefly, 15 g agar was dissolved in ddH_2O. Freshly prepared 5 X M9 stock (9 g Na_2HPO_4, 4.5 g KH_2PO_4, 0.75 g NaCl, 1.5 g NH_4Cl) dissolved in 300 ml ddH_2O. The final M9 media also has the following reagents: 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.2% glucose, 20 μg/ml theronine, leucine and histidine in L forms, and 1μg/ml thiamine. All reagent solutions were autoclaved before mixing, except the glucose solutions which were filter sterilized.

To introduce oxidative stress, 60 μM final concentration of H_2O_2 was added to cell cultures growing in LB media about OD_{600}=0.5~0.7. For osmotic shock, final concentration of 0.5 M NaCl was added to cell cultures in LB media.

2.3.3 Fluorescence microscopy

Cells were inoculated from glycerol stocks to 5 mL pre-warmed 30 °C M9 medium, and were incubated for 50 hours at 30 °C with shaking at 250 rpm (stationary phase). 500 µL of cultures were centrifuged at 21,000 x *g* for 10 min at room temperature. Pellets were re-suspended in 50 µL M9 medium containing 25 µg/mL FM4-64 and 0.2 µg/mL DAPI, which stain membranes and DNA, respectively, and incubated for 15 min at room temperature in the dark. 0.5 µL of stained cells were applied to a clean slide and viewed using a Nikon E-400 epifluorescence phase-contrast microscope fitted with DAPI (Ex 320-400 nm, Em 435-485 nm) or Tx-Red (Ex 532-587 nm, Em 608-683 nm) filter cubes. Images were captured using a CoolSNAP *HQ*^{2} CCD camera (Photometrics) and processed with NIS-Elements AR imaging software (Nikon). Between 1,200 and 3,300 cells and minicells were counted for each strain. The experiment was performed three times with similar results.

2.3.4 Bioinformatic search

The bioinformatic search was based on the sequence of *E.coli K-12 MG1655* obtained from the database of National Center for Biotechnology Information. Three groups of data were collected. The first group was the random distribution of the number of CCWGG sites in 700bp windows on the genome. This was theoretically calculated based on the assumption that the occurrences of the bases A, G, T and C were equal. The second group was the actual number of the sites in *E.coli* K-12 MG1655 strain in any 700bp window. One million bases were tested to represent the whole bacterial genome. The last group was the

actual number of the sites in between 500bp upstream of translation start site of a gene and 200bp downstream of the translation start site. Forward strand of the genome was evaluated to represent the whole genome.

CHAPTER THREE

DCM, HFQ AND DSRA REGULATE RPOS EXPRESSION IN A BACTIERAL RNA-DIRECTED DNA METHYLATION PATHWAY

3.1 Introduction

The phenotype survey and genome-wide bioinformatic search in the previous chapter provided us general framework for how Dcm functions in bacterial gene regulation, and suggested several interesting target genes for further analysis. Next, I chose one of the interesting targets, RpoS, to provide the proof of concept about the relationship between Dcm and Hfq/sRNA in gene regulation.

Clusters of 5'-CCWGG-3' sites at the promoter region of *rpoS* provide a prerequisite for a potential Dcm-mediated transcriptional regulation (Figure 3.1). The primary promoter, *rpoS*p, lies within the upstream gene *nlpD,* and has been reported to be stress-responsive (83). Basal levels of *rpoS* mRNA can also be co-transcribed from the stress-insensitive *nlpD* promoters (*nlpD*p1 and *nlpD*p2) as a bicistronic message (84). A group of Dcm methylation sites are present within the *rpoS* promoter region. Within the 1.3 kilobase region encompassing the *rpoS* promoters and sequence coding for 5'-UTR, there are seven possible Dcm methylation sites (Figure 3.1). Given a random distribution of nucleotides, one would expect roughly 3 ± 1 CCWGG sites within a DNA fragment of this size and this cluster of Dcm methylation sites is statistically significant

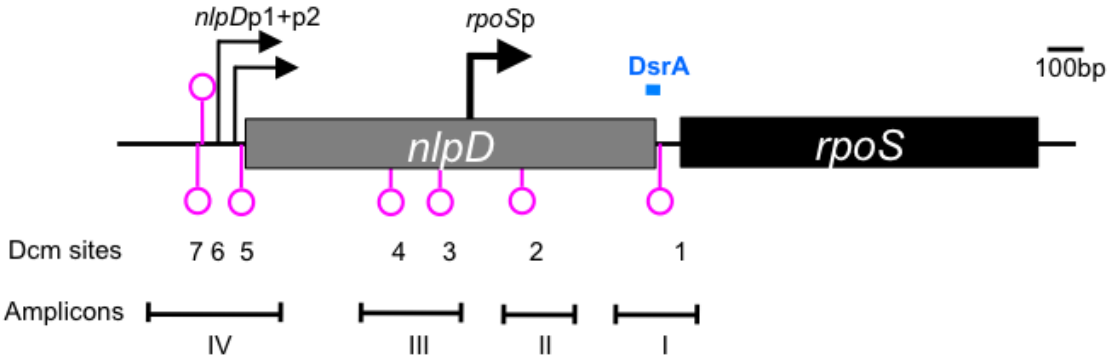


Fig. 3.1. An unusual cluster of Dcm methylation sites is found near *rpoS* promoters. The *rpoS* mRNAs are transcribed from two *nlpD* promoters and the *rpoS* main promoter. Dcm methylation sites (sticks and balls) within this promoter region are numbered sequentially for reference purposes. Regions tested in qPCR assays are labeled as amplicons. The DNA sequence that encodes the site at which DsrA binds the *rpoS* mRNA is labeled.

(***p $\leqslant$ 0.001). This pattern was recently commented upon by Kahramanoglou et al. (52).

RpoS is an important global gene regulator during stress conditions, which is related to the many stress conditions I tested in Chapter Two. RpoS is the second most abundant sigma factor of the RNA polymerase (RNAP), and it recognizes additional promoters for RNAP under stress conditions, so the additional proteins can help bacteria survive through stresses. It is the key regulator governing bacteria as they enter the stationary phase (26). It also recognizes genes that must be expressed during numerous stress responses, including cold shock, osmotic shock, oxidative stress and nutrient limitation (85-87).

Regulation of RpoS by sRNAs and Hfq is well known, and it is one of the best-studied examples of bacterial sRNA regulation. Thus, it would be quite interesting if another layer of regulation were uncovered for this system. Given the critical regulatory role of RpoS, it is not surprising that its own expression is highly regulated through a variety of mechanisms including both transcriptional and post-transcriptional controls, as well as targeted protein degradation (26,88). Post-transcriptional control of *rpoS* by several small non-coding RNAs (ncRNAs) has been widely studied and is considered one of the primary determinants of cellular protein levels under stress conditions. Its translation can be activated by sRNAs like DsrA, RprA and ArcZ during cold stress (89,90), cell surface stress or osmotic stress (91), and oxidative stress (92), respectively. The long 5'-untranslated region (5'-UTR) of the *rpoS* mRNA folds into an inhibitory structure,

repressing its translation at 37 °C (93,94). Upon cold adaptation (growth at or below 30 °C), DsrA base pairs with a segment of the *rpoS* 5'-UTR (90,95), causing it to refold and making the ribosome binding site accessible for translation initiation. RNase III cleavage within the *rpoS* 5'-UTR is also altered by base pairing with DsrA under cold stress, further enhancing translation (96).

Collectively, given the above reasons, I then looked into the potential functions of Dcm and Hfq on the *rpoS* expression during cold stress.

3.2 Results and Discussion

3.2.1 Dcm regulates the *rpoS* mRNA levels under cold stress

Since cold adaptation in *E. coli* involves translational regulation of *rpoS* through the action of Hfq and DsrA, we tested the hypothesis that this regulation may also require the recruitment of Dcm to this locus by examining *rpoS* transcript levels by RT-qPCR. During exponential growth at 37 °C, *dcm* had no effect on *rpoS* mRNA levels (Table 3.1). Upon entry into stationary phase, however, the *dcm* mutant showed 3.7-fold increased expression relative to wt. This compares favorably with the 3.5-fold overexpression of *rpoS* mRNA recently reported for a comparable *dcm* strain from a genome-wide analysis of *E. coli* gene expression during transition to stationary phase (52). As expected, cells lacking Hfq or DsrA did not significantly affect *rpoS* mRNA levels at 37 °C in exponential phase, nor did the deletion of *dcm*. However, Dcm significantly impacted *rpoS* mRNA levels produced in response to cold stress during exponential growth. The *dcm* strain showed a 2.7-fold increase in *rpoS* mRNA

levels relative to wt (Table 3.1) illustrating that Dcm represses *rpoS* during 30 °C growth. Interestingly, the *dcm hfq::kan* double mutant strain returned the mRNA level to near the wild-type level. We interpret these phenotypes as indicating that Hfq and Dcm function together (directly or indirectly) in a pathway that controls the *rpoS* mRNA levels, and thus the deletion of both genes shuts down the regulatory pathway altogether. This finding correlates with the phenotype that the *dcm hfq::kan* double mutant ameliorates the slow growth rate of *hfq::kan* cells under cold stress (Figure 2.5A). The quenching effect in the *dcm hfq::kan* strain, however, was not observed in the *dcm dsrA::cam* strain (Table 3.1). The differences between those two types of double mutant strains (*$p \leqslant 0.05$) may potentially be explained by the presence of other sRNAs that also regulate *rpoS* and may affect this behavior in the absence of DsrA.

Previous studies have shown that during low-temperature growth, Hfq and DsrA stabilize *rpoS* mRNA and slow its rate of degradation while acting to promote RpoS translation (15,93). It was therefore surprising that *rpoS* mRNA levels were enhanced more than 4-fold in the *dsrA* and *hfq* mutants, indicating that their presence reduced *rpoS* mRNA levels. To eliminate the possibility that the observed differences resulted from strain variations, the activation of an RpoS-GFP protein fusion by DsrA was confirmed in our system (Figure 3.2). Therefore, the observation that Hfq and DsrA repressed *rpoS* mRNA levels emphasizes that Hfq and DsrA have functions at the transcriptional level in addition to altering mRNA stability and translatability. This observation is consistent with research indicating the Hfq effect on transcription (97). A possible

Table 3.1. RpoS gene expression as measured by RT-qPCR under different growth conditions.

Stains in comparison to the wildtype	37 °C exponential phase	30 °C exponential phase	37 °C stationary phase	30 °C stationary phase
dcm/wt	0.9 ± 0.2 [a]	2.7 ± 0.4*	3.7 ± 1.0 [a]	1.2 ± 0.4
hfq/wt	0.5 ± 0.2	4.1 ± 0.8**	1.2 ± 0.1	2.8 ± 0.9*
dcm hfq/wt	0.8 ± 0.2	1.8 ± 0.5	4.4 ± 2.5	0.9 ± 0.0
dsrA/wt	0.8 ± 0.3	4.6 ± 1.2**	0.3 ± 0.1*	4.9 ± 2.0*
dcm dsrA/wt	0.6 ± 0.2	7.1 ± 1.6**	2.4 ± 0.1**	4.7 ± 2.0*

Data are reported relative to the wildtype strain grown under comparable conditions. (*$p \leqslant 0.05$, ** $p \leqslant 0.01$)

[a] Data are consistent with previous studies that assessed *rpoS* mRNA levels during the transition from exponential to stationary phase in a comparable *dcm* strain (52).

explanation for our data that more *rpoS* mRNA was produced in the absence of Hfq or DsrA is somehow related to a feedback loop. If cells lack Hfq or DsrA, then they can not make enough RpoS proteins to fight against cold stress, hence, bacteria may need to activate all their synthetic pathways, especially transcription here when translational machinery was not function normally, to compensate the lack of RpoS.

Another notable observation from our data is the mRNA level differences in different cell growth phases. The absolute *rpoS* mRNA levels (as opposed to the relative levels we have been discussing up to this point) were substantially lower in stationary phase than in exponential phase. This trend is opposite to that of protein expression. The delay of protein synthesis in exponential phase may provide a reasonable time frame for post-transcriptional gene regulation by Hfq and sRNAs, which may subsequently impact further transcription. Therefore, the difference between RpoS protein and mRNA levels suggests a level of regulation linking Hfq/DsrA-mediated RpoS translational activation and on-going transcription of the gene.

As the low temperature cells transition from exponential to stationary phase, the repression of Dcm on *rpoS* mRNA went away (Table 3.1). Similar trends were observed in the *hfq::kan*, *dcm hfq::kan* and *dcm dsrA::cam* strains. This result implies that Hfq and DsrA might be involved in the change from exponential to stationary under cold stress in a manner more significant than just the translational control of *rpoS*. A recent study showed that *rpoS* promoter methylation is altered during the transition to the stationary phase and the

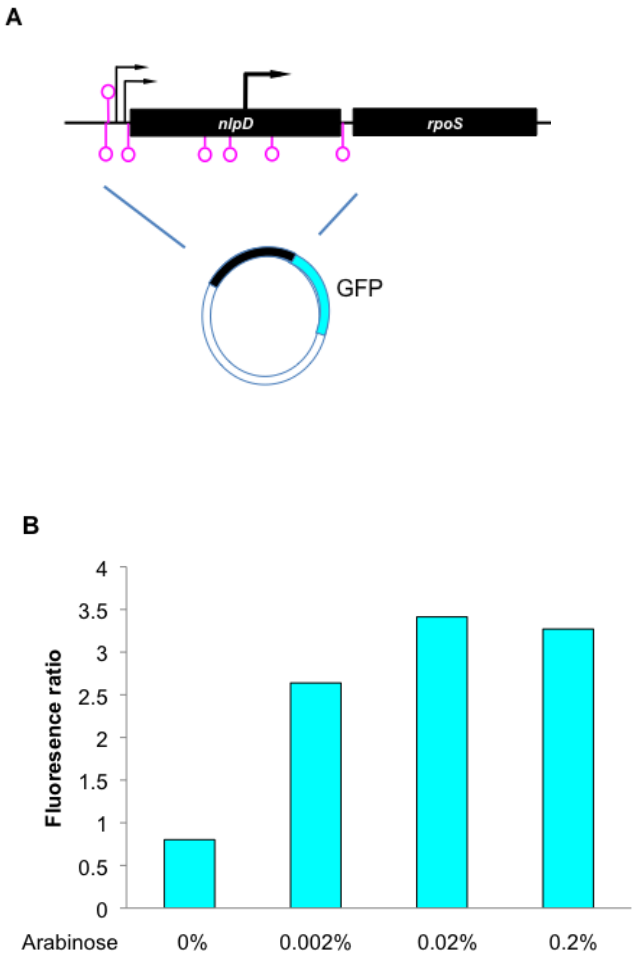


Fig. 3.2. DsrA activates the RpoS translation. (A) DNA sequence from *nlpD* promoters to 69nt downstream of RpoS translation start site was cloned into chloramphenicol resistant single-copy pBAC vector with emGFP. The reporter plasmid was verified by sequencing. (B) DsrA was complemented back into DL9 (*dsrA::cam*) strain under a P_{ara}-controlled expression of DsrA (pNM13) as a function of increasing arabinose concentration to mid-log phase. Empty vector pNM12 were tested in the same condition. GFP fluorescent intensity ratios of pNM13/pNM12 were plotted and normalized based on total protein concentrations.

authors proposed that this may result in differential gene expression (52). The *rpoS* expression data from our mutants presented above support this interpretation and link it to the sRNA biology of Hfq and DsrA through changes in DNA methylation. Therefore, we looked directly for the effect of Hfq on site-specific methylation patterns within the *rpoS* locus.

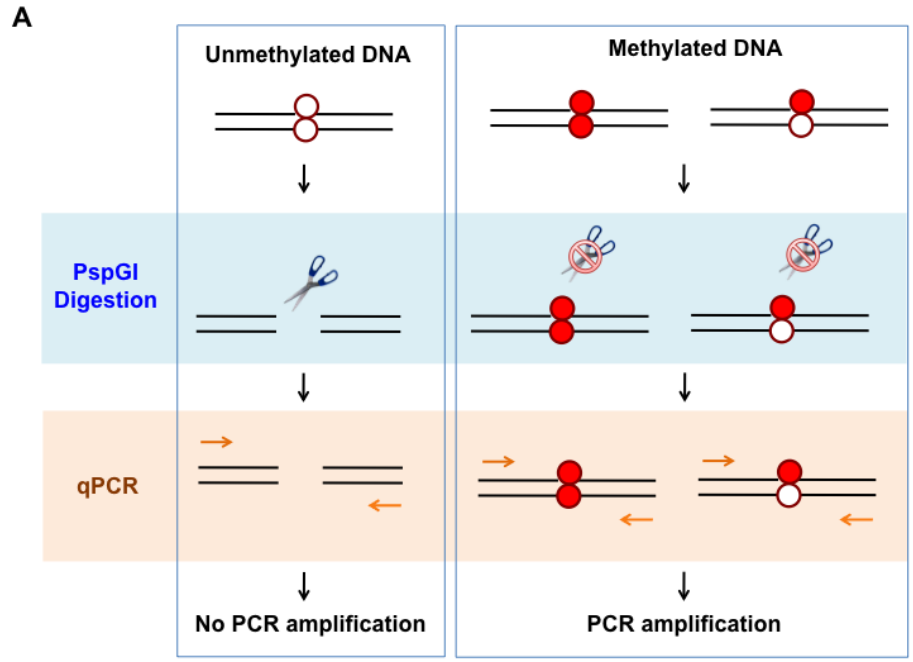
Taken together, the RT-qPCR data showed that *rpoS* is a target gene for Dcm at the mRNA level, and validated the result from the bioinformatic search. The next step was to test if the cluster of the 5'-CCWGG-3' sites near the *rpoS* promoter region are sensitive to cold stress as well.

3.2.2 Hfq stimulates Dcm methylation near the *rpoS* promoter region

To test the hypothesis that Dcm methylation within the *rpoS* promoter region changes upon cold stress, cytosine methylation levels near the *rpoS* promoter region were measured using CHOP-PCR based on the Dcm-sensitive restriction enzyme *Psp*GI (98). In this experiment, unmethylated DNA is cleaved and hence PCR amplification across the CCWGG site(s) is prevented. Methylated DNA is protected from cleavage and this amplifies normally (Figure 3.3A). Hence, the methylation load at a site can be measured. A randomly selected genomic sequence containing no CCWGG site was used as an internal control. The method was validated first on *dcm* strains and genomic DNA digested with *Bst*NI, an isochizomer of *Psp*GI whose activity is not blocked by Dcm methylation (Figure 3.3B).

Interestingly, Hfq selectively modulated cytosine methylation levels at specific genomic sites under cold stress (Figure 3.4). At 37 °C, the absence of Hfq had no significant effect on the methylation pattern near the *rpoS* promoters in either exponential or stationary phase. However, under cold stress, changes in methylation levels became obvious. Dramatic change in methylation levels were apparent in the stationary phase, where methylation loads were only about half that of wild-type cells at sites 1- 4, clustering around the main *rpoS* promoter. Methylation levels at sites 5 - 7 were affected by neither *hfq* nor *dsrA* (Figure 3.4D), consistent with the stress-insensitive activity of the *nlpD* promoters. Without an active cytosine demethylation pathway in bacteria, the decreased methylation level in the absence of Hfq observed above is most likely to be a result from the Hfq effect on altered Dcm function. These data support our hypothesis that Hfq affects cytosine methylation specifically at the *rpoS* promoter regions during cold stress.

Hfq did not just alter the Dcm methylation levels in response to cold shock, but it was also in charge of the methylation when cells enter the stationary phase. In general, the cytosine methylation levels did not saturate in exponential phase, and almost all sites were saturated in stationary phase (Figure 3.4). This change in methylation is consistent with recent bisulfite sequencing studies on the whole *E. coli* genome (52). Our data and their data both show that the cytosine methylation level eventually saturates in wild type stationary phase cells. This observation is very similar to the fact that ummethylated CpG island are often



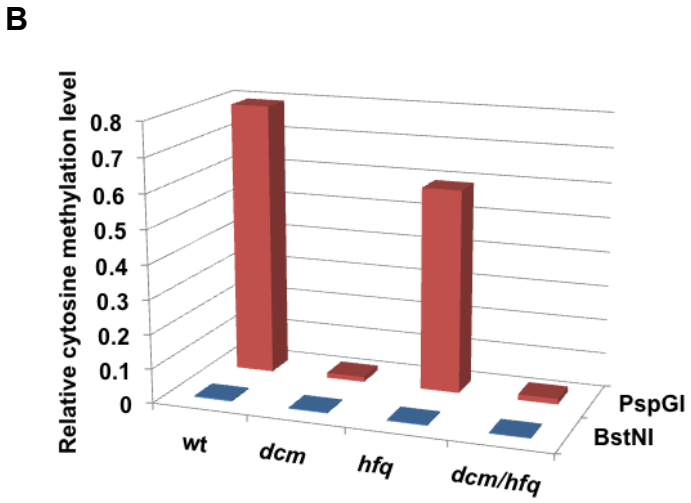


Fig. 3.3. Method for genomic cytosine methylation level quantifications. (A) Principles of *Psp*GI digestion on genomic DNA following by quantitative PCR. Unmethylated 5'-CCWGG-3' sites (open cycles) will be digested by *Psp*GI. Methylated sites (red solid cycles) cannot be digested by *Psp*GI; hence the DNA can be used as template in the following qPCR measurements. (B) Genomic DNA samples were treated with Dcm methylation sensitive restriction enzyme *Psp*GI (red bars) and Dcm methylation insensitive *Bst*NI (blue bars). Cytosine methylation levels shown here were measured in amplicon III.

found near promoters. In the *hfq* knockout strains, however, similar increases were not observed at low temperature for sites 1 - 4. Thus, it suggests that the low methylation levels in *hfq* strains may play an important regulatory role at transcriptional level and that Hfq is required to stimulate cytosine methylation at specific loci as cells enter stationary phase.

Methylation levels in this region were also measured in *dsrA::cam* mutant strains. Slight decreases in methylation were apparent under several conditions but the magnitudes of the changes were subtle and may not be biologically significant. Compared to *hfq::kan* cells, *dsrA::cam* had less influence on methylation at site 1 in stationary phase under cold stress (*$p \leqslant 0.05$), so Hfq appears to have a greater impact than the cognate sRNA species. Alternatively, Hfq may have functions separate from its role in sRNA:mRNA pairing, such as recruiting other proteins to the mRNA (99). I will approach the role of DsrA from a different angle later.

Noticeably, for sites 1-4, there was a trend that DNA cytosine methylation levels were slightly decreased in both *hfq::kan* and *dsrA::cam* strains, with site 1 being the most affected (Figure 3.4A-C). This result might be a coincidence, but it may also indicate a spread of DNA cytosine methylation from site 1 to adjacent sites. We should note that it is also possible that a general decrease in DNA cytosine methylation levels could be explained by a mechanism in which Hfq and DsrA control Dcm expression in vivo. Thus, two separate questions need to be clarified regarding these results.

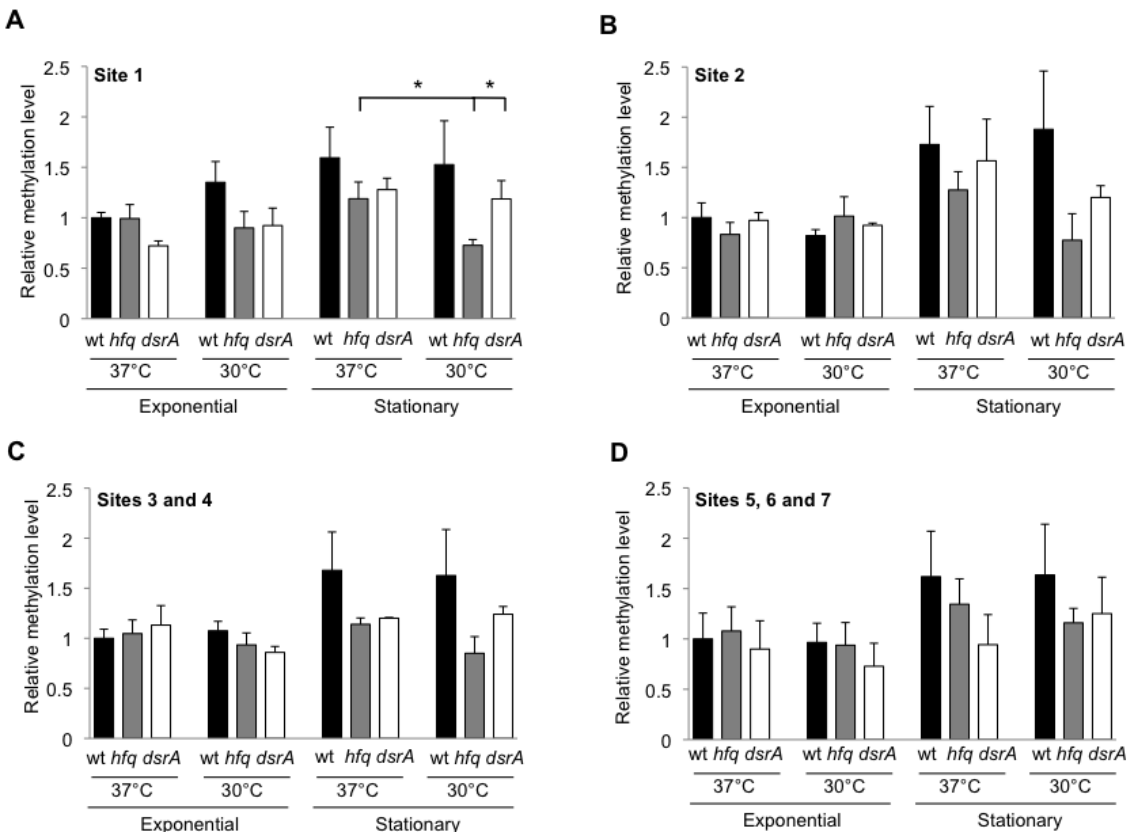


Fig. 3.4. Hfq Stimulates Dcm methylation on *rpoS* promoters specifically under cold stress. Genomic DNAs from wild type (black bars), *hfq::kan* (gray bars) and *dsrA::cam* (open bars) cells were extracted and digested by Dcm methylation sensitive restriction enzyme *Psp*GI. Methylation levels were quantified by real-time PCR on (A) site 1, (B) site 2, (C) sites 3 and 4, and (D) sites 5, 6 and 7.

The first question is whether the Hfq/ DsrA effects near the *rpoS* promoter are specific or whether it represents a more ubiquitous genome-wide phenomenon. The unchanged methylation levels at site 5-7 gave us a hint that the Hfq/DsrA effects may just be specific on site 1-4, but to avoid the possibility that those two sites may have special relationships (for example, methylation on site 1-4 inhibits the methylation on site 5-7), another genomic site far away from RpoS was selected and tested under the same condition. Measurements of methylation at the unrelated genomic site near *rpoD* showed that Hfq and DsrA did not influent this site (Figure 3.5A). This result supports a specific role of Hfq/DsrA on the Dcm methylation at *rpoS* promoter region.

The second question is whether Hfq/DsrA affects cellular concentration of Dcm. To probe this question, a high-copy plasmid was transformed into GM30, GM31, DL1, DL2, DL7 and DL8 strains to see if there was enough Dcm in these bacteria to methylate the additional DNA. After growing inside bacteria, the plasmids were then isolated and digested by *Psp*GI. DNA fragments were then run on an agarose gel. The result did not show an obvious defect caused by the deletion of *hfq* or *dsrA* (Figure 3.5B). Thus, the decreased cytosine methylation level near the rpoS main promoter does not likely result Hfq and DsrA effecting the cellular concentration of Dcm.

Taken together, these data show that Hfq specifically regulates Dcm activity near the *rpoS* main promoter under cold stress. DsrA may have the same effect on Dcm as Hfq, but investigations on the role of DsrA were needed to draw a conclusion.

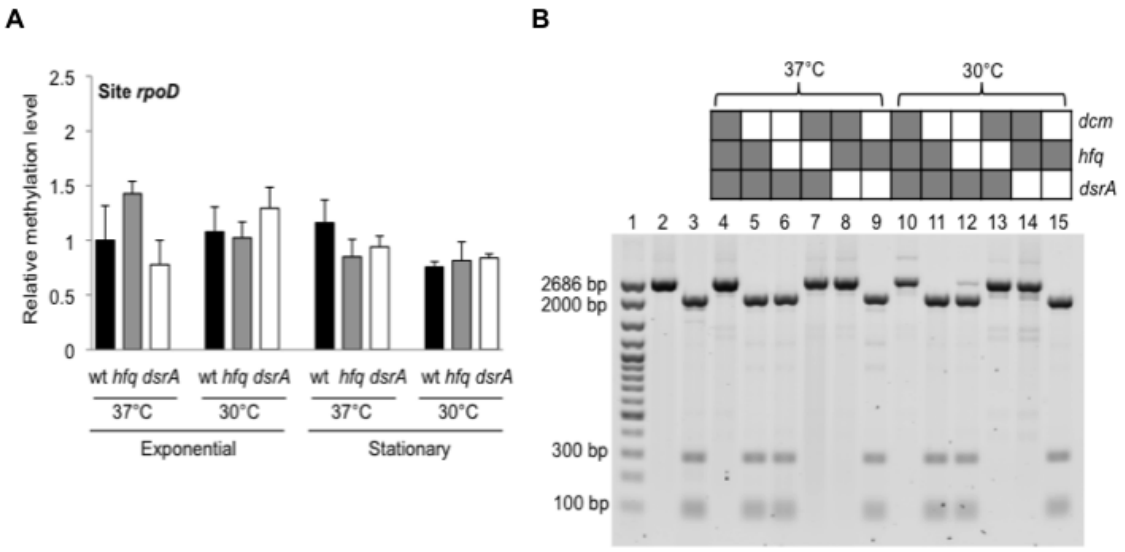


Fig. 3.5. Dcm concentration in general is not affected by Hfq and DsrA. (A) Methylation levels were measured at a CCWGG site near *rpoD*. Data in each panel were normalized relative to methylation levels observed in the 37 °C exponential phase wild type cells. (B) High-copy plasmids pUC19 were extracted from the stationary cells. Genotypes of strains are indicated on the top of the gel: a gray square represents the presence of the gene, and a blank square represents the deletion of the gene. Plasmids were linearized by *Hin*dIII (Lane 2), and then treated with methylation insensitive restriction enzyme *Bst*NI (Lane 3) or methylation sensitive *Psp*GI (Lane 4 - Lane 15).

3.2.3 Overexpressed DsrA promotes Dcm methylation at a specific site near *rpoS* promoters

While the data above implicated Hfq as a determinant of methylation levels and patterns near the *rpoS* gene during cold stress, it was still unclear whether the effect was dependent on an sRNA as well. In the absence of DsrA, methylation levels slightly decreased relative to the wild type strains (Figure 3.4), but the results were ambiguous due to the small magnitude of the changes. RpoS regulation is sufficiently complex that other sRNAs also regulating *rpoS* expression might also influence the methylation patterns, and the *hfq* knockout would have altered all of those pathways simultaneously.

To confirm the role of DsrA on DNA methylation, we therefore went to the other extreme, overexpressing DsrA instead of knocking it out. Previous studies have shown that overexpressing DsrA can partially overcome the requirement of Hfq on RpoS translation, albeit with a lower efficiency relative to cells containing Hfq (100). If DsrA stimulates Dcm methylation, then overexpressed DsrA in *hfq* cells should increase the cytosine methylation. Plasmid pNM13 (90) was used to allow arabinose-induced overexpression of DsrA in *hfq* knockout strains. A 13-fold increase in DsrA was observed upon addition of 0.02% arabinose in this system (Figure 3.6A), which is sufficient to allow paring between DsrA and *rpoS* mRNA spontaneously, even in the absence of Hfq. As expected, overexpression of DsrA in *hfq::kan* cells triggered methylation at site 1, and complemented the Dcm methylation defects caused by the deletion of *hfq* back to the same level as

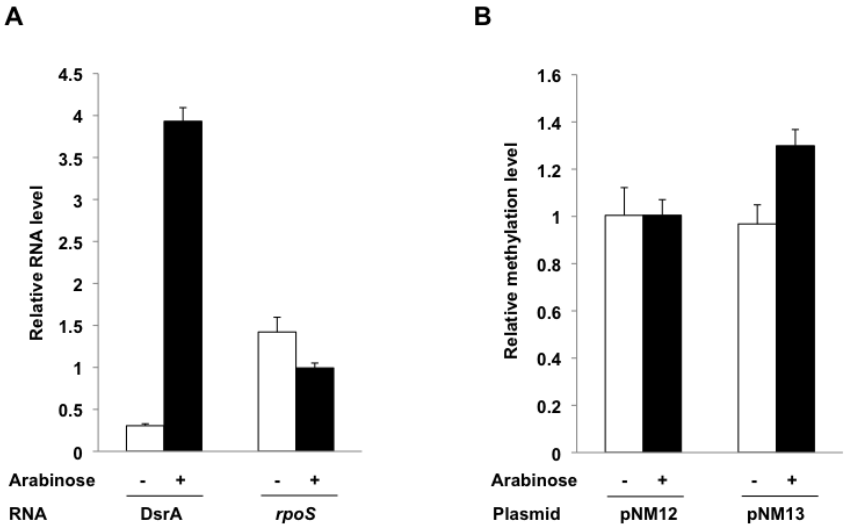


Fig. 3.6. Overexpressed DsrA increases cytosine methylation at *rpoS* promoter. (A) Relative RNA levels for DsrA and *rpoS* mRNA were quantified by RT-qPCR. Cells containing P_{ara}-controlled DsrA plasmid grew, in the absence (open bars) and presence (solid bars) of 0.02% arabinose in *hfq::kan* knockout cells at 30°C to the mid-log phase. (B) Relative DNA cytosine methylation levels at site 1 with empty vector (pNM12) and overexpressed exogenous DsrA in pNM13 were measured as in Figure 3.4. Error bars represent SEM.

in wild type cells (Figure 3.4 and 3.6B). The data confirmed a positive role of DsrA on the cytosine methylation within the *rpoS* operon.

3.2.4 Hfq stimulates Dcm methylation activity on dsDNA in vitro

One possible mechanism for the enhanced methylation near an sRNA/Hfq binding site is that Hfq somehow activates Dcm, non-specifically increasing its methylation rates. To determine if Hfq stimulates Dcm, in vitro methylation assays were performed with purified recombinant Dcm in the absence of sRNAs. Dcm activity was measured by using ^{14}C-SAM, and monitoring the ^{14}C-methyl group incorporation into a 39-mer dsDNA containing a 5'-CCWGG-3' site. The initial rate of Dcm methylation increased about 2.4-fold in the presence of Hfq (Figure 3.7A,B). Assays were followed for up to 25 hours to determine if the effect was associated with the rate or the extent of methylation. Total methylation in the presence of Hfq was still about 3-fold higher than when Hfq was absent. Thus, Hfq did not just stimulate the Dcm methylation rate, but also enhanced the extent of methylation in vitro. The results were consistent with our in vivo data on the decreased methylation level at *rpoS* promoter sites in *hfq* strains (Figure 3.4).

Recruitment of Dcm to the *rpoS* promoter region could function in a transcription-dependent manner similar to that proposed for eukaryotic RdDM pathways (55). In such a model, DsrA binding to single-strand DNA during transcription might induce methylation on the RNA·DNA hybrids. Alternatively, it could result from the sRNA binding to 5'-UTR of *rpoS* mRNA, recruiting the methyltransferase to this locus, and methylating any adjacent dsDNA within

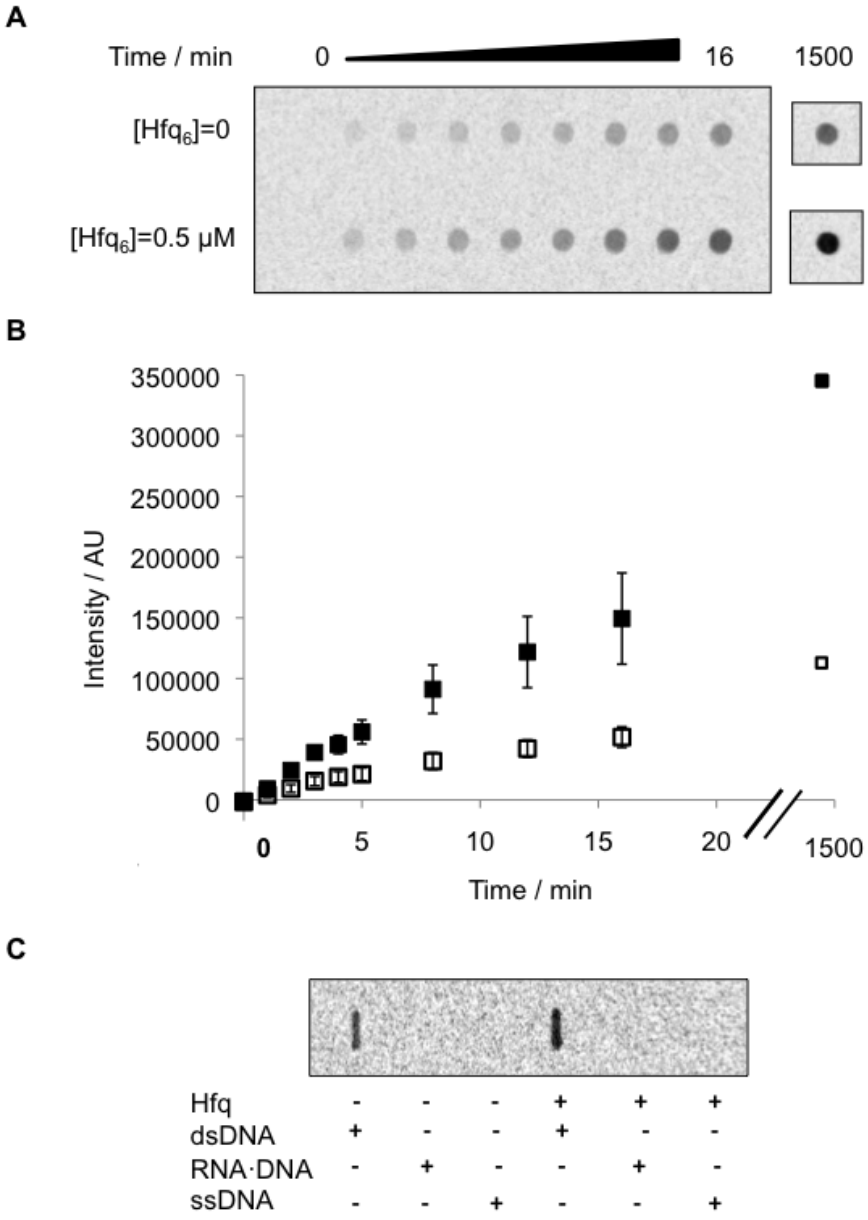


Fig. 3.7. Hfq stimulates Dcm methylation activity. (A) Dcm activity was measured by quantifying the transfer rate of radioactive methyl groups from ^{14}C-SAM onto 39-mer dsDNAs containing a 5'-CCWGG-3' site. Dot-blot was used to quantify the Dcm activity with (solid squares) or without (open squares) Hfq hexamer. Error bars the represent standard errors of the mean from triplicate experiments. (B) Quantitative plot for the result in (A) from triplicate experiments. (C) Dcm substrate selectivity was tested with dsDNA, DNA·RNA hybrid and ssDNA on slot-blot apparatus. One representative experiment of three is shown.

reach. A third possibility is that Dcm might methylate ssDNA generated transiently during transcription. Since Dcm is the only cytosine methyltransferase in *E. coli* K-12 strain MG1655, these possibilities can be tested in vitro by assessing the substrate selectivity of Dcm. In vitro methylation assays were performed using dsDNAs, RNA·DNA hybrids and ssDNAs all containing three potential cytosine methylation sites. The results showed that Dcm methylates only dsDNA, and neither the RNA·DNA hybrid nor the ssDNA were viable substrate. The addition of Hfq to the reactions did not affect substrate selectivity in any way (Figure 3.7C). Based on these results, we favor the hypothesis that Hfq facilities DsrA binding to the 5'-UTR in *rpoS* mRNA, and then helps recruit Dcm to the *rpoS* locus to enhance methylation.

Taken together, our results show that Hfq simulates the Dcm methylation activity on dsDNA, without altering Dcm specificity on its recognition sequences. This positive role of Hfq protein and its physical presence with certain mRNAs may then cause the recruitment of Dcm to specific genomic loci in vivo.

3.2.5 Dcm and Hfq may not bind directly in vivo

To test if Dcm and Hfq bind directly *in vivo*, co-immunoprecipitation assays were performed for Hfq and C-terminal V5-tagged Dcm. *In vitro* transcribed DsrA RNA and purified His-tagged Hfq were supplemented to some samples as indicated. The Dcm-V5 bands above 50kDa were present in the input and supernatant in the Co-IP experiment, but were not present in the product on the protein A beads (Figure 3.8). Hfq monomer around 15kDa and several Hfq

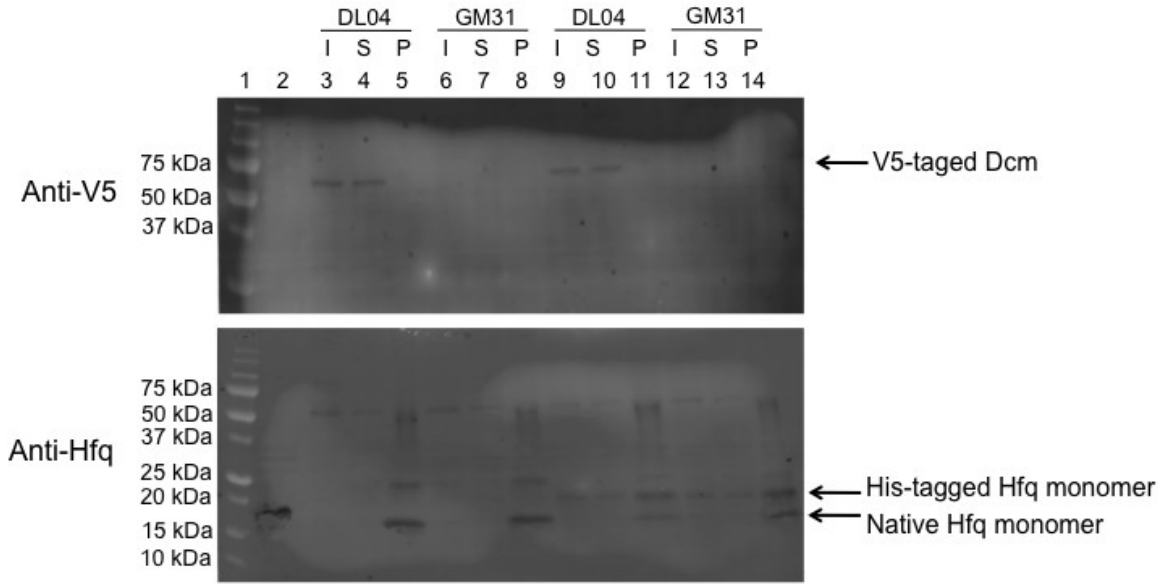


Fig. 3.8. Co-immunoprecipitation assay shows no direct binding between Hfq and Dcm. 37 °C overnight cultures of GM31 (*dcm*) and pDcm-V5 in GM31 strain (DL04) were incubated at 30 °C and induced with 0.02% L-arabinose for 3 hours. Lane 2 was purified His-tagged Hfq. In lane 9-14, in vitro transcribed and folded DsrA and purified His-tagged Hfq were added. I, S and P represent Input of the Co-IP, Supernatant after incubating input with protein A beads and Products that bind to beads, respectively. Rabbit anti-Hfq antibody was bound on protein A beads, and mouse anti-V5 antibody was applied on the western blot and visualized by Dylight 549 Anti-mouse secondary antibody (top figure). The same blot was striped and incubated with anti-Hfq followed by anti-rabbit dye and fluorescein labeled anti-rabbit secondary antibody incubation (bottom figure).

multimers (heavier bands and smear) indicated the successful IP of Hfq. One possibility is that the V5-tag on Dcm interferes with the Hfq binding surface. This data, together with the above in vitro methylation assays, imply that Hfq may bind to the DNA rather than the Dcm, and this binding might make the DNA a better substrate for the Dcm methylation.

In summary, my research furthers our understanding of DNA methylation and Hfq/sRNA-modulated gene regulations. Dam has been previously shown to affect gene regulation in bacteria (41,52,101). In those cases, changes in methylation patterns lead to gene expression. Unclear, however, was the manner in which those methylation marks were selectively placed at appropriate genomic loci. Our results show for the first time a mechanism in bacteria that site-specifically alters methylation levels with a direct impact on expression of the associated gene. Our data is in complete agreement with recent studies in which *rpoS* was identified as one of the genes whose promoter cytosine methylation changes correlates with the mRNA level changes (52). Here, a bacterial sRNA together with its RNA chaperone, Hfq, can recruit Dcm methylation to a genetic locus to where the sRNA would pair with the nascent transcript. Thus, Dcm recruitment directly impacts new transcription and thus acts as part of a feedback loop to negatively regulate new transcription once the translational activation has begun to take effect.

This result has several important implications for bacterial gene regulation. First, it shows that DNA methylation levels are critical for bacterial gene regulation and that these can be altered in response to environmental conditions.

Because bacterial sRNAs are typically made in response to environmental stresses, sRNA-directed DNA methylation provides a level of bacterial epigenetic regulation that has not previously been observed. By placing long-lived methylation marks at specific DNA loci in a kinetically controlled manner, expression of those loci can be uniquely regulated over many generations of cell division, long after the stress has disappeared or the sRNAs have been degraded. Eventually these methylation marks can be lost by dilution or by global methylation of the genome during stationary phase, but it provides the means over modest time scales to alter expression of specific genetic circuits with a memory of a recent stress.

Second, our findings imply that sRNA modulation of mRNA translation is not just a post-transcriptional event as it is often envisioned, but rather it may occur co-transcriptionally, in a manner more akin to riboswitches (102). The ability to recruit Dcm to specific loci within the *rpoS* operon would be impossible if DsrA pairs with the RpoS mRNA after transcription terminates.

The relationship between DNA cytosine methylation patterns and transcription may not be as clear as in eukaryotic gene silencing mechanisms. In bacteria, DNA methylation can induce positive or negative regulation of expression as shown by the Low Group by altering the manner and/or location that transcription factors or other regulatory proteins bind the methylated DNA (41). Further studies are needed to draw a conclusion on this regulation. It will also be interesting to investigate whether the slight preference for an extended

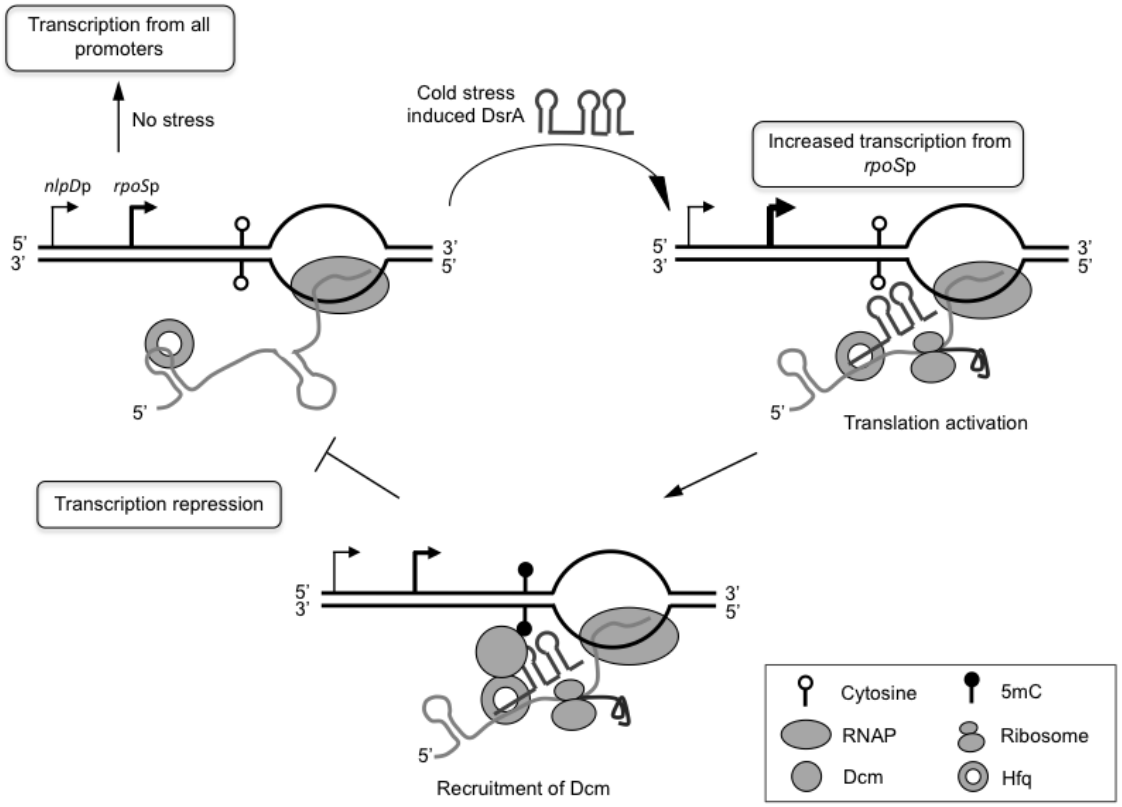


Fig. 3.9. Proposed model of bacterial RdDM. The *rpoS* mRNA can be transcribed from all its promoters in the absence of stress. 5'-UTR of the mRNA may bind to Hfq protein and wait for sRNA regulation. Upon induction of cold stress, DsrA transcription is largely activated, and DsrA binds to 5'-UTR of the *rpoS* mRNA with the help of Hfq. This leads to translational activation, which in turn recruits Dcm to this genetic locus and leads to rapid DNA methylation at specific sites within the *rpoS* promoter region. The recruitment of the Dcm protein or the methylated cytosine nucleotides results in the repression of transcription of the gene under cold stress.

CCCWGG site found in recent studies is related to the sRNA-mediated gene regulation (52).

Dcm methylation at site 1 may also be tied to *rpoS* transcription, since thses methylation sites are relatively distant from the primary *rpoS* promoter. Recent studies using high throughput sequencing to discover new transcription start sites (TSS) in *E. coli* suggest several potential new TSS for *rpoS*, which lie close to site 1 (103). Unfortunately, other studies focused on understanding the *rpoS* TSS did not show these TSS (83), indicating that some of these might be artifacts resulting from RNase cleavage of the transcripts (15). More work will be required to deconvolute the implications of methylation patterns on such data and the potential for alternative sites for transcription initiation.

In summary, our data suggest a bacterial system akin to the eukaryotic RNA-directed DNA methylation seen in transcriptional gene silencing (Figure 3.9). Here, Hfq facilitates DsrA and *rpoS* 5'-UTR pairing, which in turn recruits Dcm to this genetic locus leading to rapid DNA methylation during the transition from exponential phase to stationary phase. This methylation then may be read by transcription factors, which alter further transcription of the gene. More extensive genome wide studies under stress conditions (or simulated stress by induction of sRNAs) will be required to understand how widespread this type of regulation is in *E. coli*. Recent studies have shown a high level of overlap between transcriptional networks and sRNAs and Hfq regulated translational networks (14,104). Our model of a Dcm, Hfq and DsrA regulated feedback loop provides a potential mechanism for cross talk between these transcriptional and

translational regulation further integrating bacterial responses to environmental stress.

3.3 Material and Methods

3.3.1 Reverse transcription and real-time PCR

All synthetic oligonucleotides are listed Table 3.2. Total RNAs were extracted using TriZol reagent as per the manufacturer's protocol (Invitrogen). RNA concentration and quality were measured by UV/VIS spectrophotometer. RNase-free DNase I was used to clean up 1 µg total RNAs in a 50 µL reaction for each sample according to the manufacturer's protocols (Ambion). DNase I reactions were stopped by addition of 5 µL inhibitor suspension, and the clear supernatant was collected after centrifugation. For each sample, 9 µL of the supernatant was directly used in the reverse transcription step with 50 units of Superscript III reverse transcriptase (Invitrogen) and 100 ng random hexamers (Applied Biosystems) in a 20 µL reaction. RT reactions were performed based on the manufacturer's protocols, including incubation at 25 °C for 5 min, followed by 55 °C for one hour and finally 70 °C for 10 min. RNA templates were then hydrolyzed by adding 150 mM KOH·Tris and heating to 90 °C for 10 minutes. Reactions were neutralized by addition of 150 mM HCl to a final pH value of 8.0. Quantitative PCR (Applied Biosystems Fast 7500) reactions were run in a 10 µL reaction with 2 µL H_2O, 1 µL of 10 µM gene specific primer mixtures, 3 µL of 100-fold diluted RT reaction and 5 µL of master mixture (Applied Biosystems Fast SYBR Green Mix). 16S rRNA was used as internal control for all the strains.

Table 3.2 oligonucleotides used in my study.

Oligonucleotide*	Experiment	Sequence
rpoS 1F	qPCR amplicon I	CGT TCT CAT CAA ATT CCG CAT C
rpoS 1R	qPCR amplicon I	GAA CCA GTT CAA CAC GCT TG
rpoS 2F	qPCR amplicon II	GCT GCC TGC GAT ATC AAT CC
rpoS 2R	qPCR amplicon II	CGA ACA ACA AGC CAA CTG C
rpoS 3F	qPCR amplicon III	CAA CGG TGG AAT TTT GTG CAG G
rpoS 3RR	qPCR amplicon III	AAC CGT CAG TAT GGG AAC ATT C
rpoS 4F	qPCR amplicon IV	TGC CGG TGG ATT TGA AGT G
rpoS 4RR	qPCR amplicon IV	GGG ATT CTC GTC TTA CCC GTA G
rpoD q1F	qPCR for an *rpoD* promoter region	ATT ACC GCG CAG TGT AGG A
rpoD q1R	qPCR for an *rpoD* promoter region	AGG GAC TTT TAT AAG GGT GAG GA
fhlA qF	qPCR internal control for methylation detection	TGG CAC GGG TAA AGA GCT GA
fhlA qR	qPCR internal control for methylation detection	AGG GAG CTT TTA TCC GCC AGT T
16S qF	qPCR internal control for RT-qPCR	TGT CGT CAG CTC GTG TTG TGA A
16S qR	qPCR internal control for RT-qPCR	GCA CTT TAT GAG GTC CGC TTG CT
DsrA qF	qPCR for DsrA	AAC ACA TCA GAT TTC CTG GT
DsrA qR	qPCR for DsrA	GCT TAA GCA AGA AGC ACT TAA A
mdh qF	qPCR for Mdh mRNA	AAT AAC CGG CAC TTC AAC TT
mdh qR	qPCR for Mdh mRNA	CAC GCT GGA TAT CAT TCG TT
DsrA KO gm30 R	comfirmation for *dsrA::cam*	TAT TCA TGA CTT CAG CGT CTC TG
DsrA upstream F	comfirmation for *dsrA::cam*	ACT CCT CTT ACC AGG ATG ATT CTC
FRT Upper Oligo for Hfq	*hfq::kan* knockout	AAA GGT TCA AAG TAC AAA TAA GCA TAT AAG GAA AAG AGA GAA ATT AAC CCT CAC TAA AGG
FRT Lower for Hfq	*hfq::kan* knockout	CAG GAT CGC TGG CTC CCC

v2		GTG TAA AAA AAC AGC CCG AAA CCT AAT ACG ACT CAC TAT AGG
DsrA knockout F	*dsrA::cam* knockout	ATA TGG CGA ATA TTT TCT TGT CAG CGA AAA AAA TTG CGG ATA AGG TGA TGA ATT AAC CCT CAC TAA AGG GCG
DsrA knockout R	*dsrA::cam* knockout	TAT TCA TGA CTT CAG CGT CTC TGA AGT GAA TCG TTG AAT GCA CAA TAA AAT AAT ACG ACT CAC TAT AGG GCT C

PCR conditions consisted of denaturation at 95 °C for 20 seconds, followed by 40 cycles of 10 seconds at 95 °C and 30 seconds at 60 °C. Melting curves were measured automatically after real-time PCR, or the PCR amplified products were checked on agarose gels to confirm the quality of real-time PCR products. Control reactions lacking template or RT were also run to ensure the quality of specific real-time PCR results. Target gene levels were calculated using software ABI 1.4, and analyzed by the $\Delta\Delta C_t$ method based on equation 1.

$$x = 2^{[\Delta C_{t(control)} - \Delta C_{t(sample)}]} \quad \text{(Eq 1)}$$

3.3.2 Methylation sensitive restriction digestion and real-time PCR

Cells were harvested by centrifugation at mid-log phase and stationary phase. The *hfq* strains containing pNM13 or pNM12 were inoculated into pre-warmed 30 °C LB media with 100 µg/mL ampicillin to an OD_{600} = 0.04. After one hour shaking (250 rpm) at 30 °C, DsrA over-expression was induced by addition of 0.02% L-(+)-arabinose (Sigma). Cells were harvested with OD_{600} values around 0.5.

Genomic DNA was isolated using Generation Capture Columns (Qiagen). The restriction enzyme *Psp*GI (0.83 U per ng genomic DNA) was used to digest genomic DNA at 75 °C overnight. A random genomic sequence within the *fhlA* gene that does not contain any 5'-CCWGG-3' site was used as an internal control. Real-time PCR (ABI Fast 7500) reactions were run and data were analyzed as described above. Methylation levels on individual sites (M) were calculated based on equation (2) where D_{obs} is the qPCR amplicon detected

value across single or multiple methylation sites, and n is the number of methylation sites within each qPCR amplicon.

$$M = D_{obs}^{\ 1/n} \quad \text{(Eq 2)}$$

3.3.3 Construction of pDCM-V5 plasmid and pRpoS-GFP constructs

pDcm-V5 was constructed by inserting the *dcm* gene sequence followed by 3'-end V5 tag into pBAD24 vector. BamHI and EcoRI were used to digest both vector pBAD24 and PCR amplified Dcm sequence following manufacturer's protocols (NEB). T4 DNA ligase was used for ligation at room temperature for 1 hour following manufactory protocol (NEB), and 20 ng DNA from the ligation reaction was directly electrophoresed into DL1 (*dcm hfq::kan*) competent cells. The plasmid sequence was verified by DNA sequencing following manufactory protocol (Beckman coulter).

Plasmid pRpoS-GFP was constructed by inserting *rpoS* gene with the full promoter region into the pBacEmGH plasmid. Arabinose inducible promoter of the parent plasmid was deleted to avoid the complication in later experiments because the DsrA containing plasmid is arabinose inducible. NsiI and NheI were used to digest both the vector and PCR amplified Dcm sequence following manufactory protocols (NEB). T4 DNA ligase was used for ligation at room temperature for 1 hour following manufactory protocol (NEB), and 20 ng DNA from the ligation reaction was directly electrophoresed into competent cells. The plasmid sequence was verified by DNA sequencing following manufactory protocol (Beckman coulter). One technical note worth mention here: since this is a single-copy plasmid with about 10 kb long, it is hard to get enough plasmid as

template for sequencing. So as Qiagen protocol recommended with such a plasmid, I use 25 mL cell culture, instead of 1~5 mL cell culture, per mini column to isolate the plasmid.

3.3.4 Purification of Dcm protein

Dcm was purified according to the protocols described in previous studies (105) with modifications. Briefly, a total of 300 mL pBAD24-DcmV5 in DL1 culture (30 μg/ml Kan and 100 μg/ml Amp) was grown to an OD_{600} between 0.3~ 0.4, and then the protein production was induced by freshly prepared L-(+)-arabinose (Sigma) to the final concentration of 0.02%. Cell pellets were harvested after overnight incubation by centrifuging at 5,000 rpm/15 min/4 °C. Pellets were resuspended in 10 mL ice-cold buffer A (10 mM potassium phosphate, pH 7.0, 1 mM EDTA, and 10 mM 2-mercaptoethanol) plus 0.1 M NaCl, and 1/4 pill of EDTA-free complete protease inhibitor cocktail (Roche). Cells were cracked on ice by sonication for 5 cycles of 30 seconds with 2 min rest in between. The first purification step was carried out with pre-equilibrium cellulose phosphate column P-11 (Whatman) at 4 °C. Filtered cell supernatant was applied to the column. Column was washed with two-column volume (CV) of Buffer A containing increasing NaCl concentration (0.1 M, 0.2 M, 0.4 M, 0.6 M), respectively. Protein fractions were eluted with two CV of buffer A plus 0.8 M NaCl. Fractions were analyzed by Bio-Rad protein assay and SDS-PAGE.

Hiload superdex200 sizing exclusive column (GE healthcare) was used to further purify Dcm with buffer A plus 0.4 M NaCl at 4 °C. Fractions from Sizing exclusive column were concentrated with 30k NMWL Ultra-15 Centrifugal Filter

Devices (Pierce) and dialyzed against storage buffer (45% glycerol solution 10 mM potassium phosphate, pH 7.0, 10 mM 2-mercaptoethanol) twice overnight. The final Dcm concentration was measured by UV/Vis spectrometer at 280 nm.

3.3.5 In vitro methylation assays and filter binding assays

In vitro methylation assays A 39-mer dsDNA containing a single 5'-CCWGG-3' site was annealed in 100 mM Tris·HCl, pH 8.0, 150 mM NaCl at 95°C for 2 min, and then chilled at 37 °C for at least 5 min. DTT (0.4 mM) and $MgCl_2$ (10 mM) were added to the annealed duplex, and the solution was kept at room temperature until use. Methylation reactions were incubated at 37°C in 100 mM Tris·HCl, pH 8.0, 150 mM NaCl, 0.4 mM DTT, 10 mM $MgCl_2$ with 25 nM Dcm and 0.01 µCi Adenosyl-L-methionine, S-[methyl-^{14}C] (Perkin Elmer) in the presence or absence of 0.5 µM Hfq hexamer.

At each time point, 20 µL aliquots were removed and applied to pre-equilibrated Hy-BOND nucleic acid binding membrane (Amersham Biosciences) on a dot-blot device (Bio-Rad), chased by 100 µL of reaction buffer (without ^{14}C-SAM or enzymes). At the end of the time course, membranes were removed from the dot blot device and washed in reaction buffer for 10 min, and dried with suction for 10 min. Membranes were imaged using a phosphorimager cassette and a Typhoon9210 scanner (Amersham Biosciences). The intensities of radioactive bands were quantified using ImageQuant 5.1. Assays assessing Dcm specificity were performed in a similar manner but using appropriate su*Bst*rates.

3.3.6 Co-immunoprecipitation assays

Overnight cell cultures were shocked at 30 °C for 3 hours with freshly prepared L-(+)-arabinose (Sigma) to the final concentration of 0.02%. 10 mL cells were harvested by centrifugation, and re-suspend in 4 ml cold cell lysis buffer (50 mM Tris·HCl, pH 8.0, 50 mM NaCl, 1 mM $MgCl_2$ and 0.5 mM DTT). Cells were cracked by sonication in the presence of EDTA-free complete protease inhibitor cocktail (Roche). The supernatants were collected by centrifuge at 10,000 rpm for 30 minutes at 4 °C. 800μL of each lysate sample and 1 μL Anti-Hfq polyclonal antibody were incubated at 4 °C overnight with rotation. 40 μL Dynabeads protein A (Invitrogen) were pelleted and added to each cell lysate sample, and incubated at 4 °C for another 3 hours. Beads were then washed with washing buffer (PBS and 0.02% Tween20) four times. 40 μL SDS loading buffer were added to each reactions, and heated at 95 °C for 2 hours. Supernatants were loaded on 10-20% SDS-PAGE (Bio-rad). Proteins were transferred at 300 mA, 4 °C for 90 minutes to Immuno-Blot PVDF membrane (Bio-Rad). Membranes were blocked with 5% dry milk in washing buffer for an hour at room temperature with shaking. Monoclonal anti-V5 antibody (Sigma) and anti-Hfq antibody were used for western blot at room temperature for 1 hour. After washing with PBS/Tween20 for four times, 1:25,000 Dylight 549 Anti-mouse secondary antibody (Thermo Scientific) and 1:4,000 fluorescein labeled anti-rabbit secondary antibody (Abcam) were incubated with western blot for an hour, separately or together. After washing with PBS/Tween20 for six times, green laser 532 nm/ 580 nm filter was used to visualize anti-mouse secondary antibodies for V5-tagged Dcm

detection, and green laser 532 nm /526 nm filter was used to detect FITC labeled anti-rabbit secondary antibodies for Hfq signal with Typhoon scanner.

DsrA RNA used in Co-IP assays was *in vitro* transcribed from *Dra*I digested pBAU10301 with final concentration of 30 ng/μL T7 polymerase in reaction buffer (40 mM Tris·HCl, pH 8.0, 10 mM MgCl2, 10mM DTT, 2 mM sperimidine, 0.1% Triton X-100, 1.25 mM of each GTP, CTP, ATP and UTP) at 37 °C for 5 hours. DsrA was purified by denaturing PAGE and ethanol precipitation.

3.3.7 GFP fluorescence assays

Assays were carried out based on previous protocol in our lab with some modifications (106). The modifications are as following. Instead of shocking cell culture with high percentage of arabinose, I only used optimized 0.02% arabinose. I also normalized all the data using total amount of protein concentrations in each sample.

REFERENCES

1. Ramachandran, P.V. and Ignacimuthu, S. (2012) RNA interference as a plausible anticancer therapeutic tool. *Asian Pacific journal of cancer prevention : APJCP*, **13**, 2445-2452.
2. Baer, C., Claus, R. and Plass, C. (2013) Genome-wide epigenetic regulation of miRNAs in cancer. *Cancer Res*, **73**, 473-477.
3. Motavaf, M., Safari, S. and Alavian, S.M. (2012) Therapeutic potential of RNA interference: a new molecular approach to antiviral treatment for hepatitis C. *Journal of viral hepatitis*, **19**, 757-765.
4. Lee, J.T. (2012) Epigenetic regulation by long noncoding RNAs. *Science*, **338**, 1435-1439.
5. Siomi, H. and Siomi, M.C. (2009) On the road to reading the RNA-interference code. *Nature*, **457**, 396-404.
6. Zaratiegui, M., Irvine, D.V. and Martienssen, R.A. (2007) Noncoding RNAs and gene silencing. *Cell*, **128**, 763-776.
7. Mello, C.C. and Conte, D., Jr. (2004) Revealing the world of RNA interference. *Nature*, **431**, 338-342.
8. Montgomery, M.K., Xu, S. and Fire, A. (1998) RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans. *Proc Natl Acad Sci U S A*, **95**, 15502-15507.
9. Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806-811.
10. Zamore, P.D., Tuschl, T., Sharp, P.A. and Bartel, D.P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, **101**, 25-33.
11. Bernstein, E., Caudy, A.A., Hammond, S.M. and Hannon, G.J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363-366.
12. Hannon, G.J. (2002) RNA interference. *Nature*, **418**, 244-251.
13. Filipowicz, W., Jaskiewicz, L., Kolb, F.A. and Pillai, R.S. (2005) Post-transcriptional gene silencing by siRNAs and miRNAs. *Curr Opin Struct Biol*, **15**, 331-341.
14. Storz, G., Vogel, J. and Wassarman, K.M. (2011) Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Mol Cell*, **43**, 880-891.
15. McCullen, C.A., Benhammou, J.N., Majdalani, N. and Gottesman, S. (2010) Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: pairing increases translation and protects rpoS mRNA from degradation. *J Bacteriol*, **192**, 5559-5571.
16. Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G. and Gottesman, S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol*, **50**, 1111-1124.
17. Gottesman, S., McCullen, C.A., Guillier, M., Vanderpool, C.K., Majdalani, N., Benhammou, J., Thompson, K.M., FitzGerald, P.C., Sowa, N.A. and

FitzGerald, D.J. (2006) Small RNA regulators and the bacterial response to stress. *Cold Spring Harb Symp Quant Biol*, **71**, 1-11.
18. Sauer, E. (2013) Structure and RNA-binding properties of the bacterial LSm protein Hfq. *RNA Biol*, **10**.
19. Gottesman, S. (2004) The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annu Rev Microbiol*, **58**, 303-328.
20. Vogel, J. and Luisi, B.F. (2011) Hfq and its constellation of RNA. *Nat Rev Microbiol*, **9**, 578-589.
21. Soper, T.J. and Woodson, S.A. (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA*, **14**, 1907-1917.
22. Mikulecky, P.J., Kaw, M.K., Brescia, C.C., Takach, J.C., Sledjeski, D.D. and Feig, A.L. (2004) Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nat Struct Mol Biol*, **11**, 1206-1214.
23. Lease, R.A. and Woodson, S.A. (2004) Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA. *J Mol Biol*, **344**, 1211-1223.
24. Romilly, C., Caldelari, I., Parmentier, D., Lioliou, E., Romby, P. and Fechter, P. (2012) Current knowledge on regulatory RNAs and their machineries in Staphylococcus aureus. *RNA Biol*, **9**, 402-413.
25. Gottesman, S. and Storz, G. (2010) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol*.
26. Battesti, A., Majdalani, N. and Gottesman, S. (2011) The RpoS-mediated general stress response in Escherichia coli. *Annu Rev Microbiol*, **65**, 189-213.
27. Castel, S.E. and Martienssen, R.A. (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat Rev Genet*, **14**, 100-112.
28. Geier, G.E. and Modrich, P. (1979) Recognition sequence of the dam methylase of Escherichia coli K12 and mode of cleavage of Dpn I endonuclease. *J Biol Chem*, **254**, 1408-1413.
29. Hattman, S. (2005) DNA-[adenine] methylation in lower eukaryotes. *Biochemistry (Mosc)*, **70**, 550-558.
30. May, M.S. and Hattaman, S. (1975) Deoxyribonucleic acid-cytosine methylation by host- and plasmid-controlled enzymes. *J Bacteriol*, **122**, 129-138.
31. Hanck, T., Schmidt, S. and Fritz, H.J. (1993) Sequence-specific and mechanism-based crosslinking of Dcm DNA cytosine-C5 methyltransferase of E. coli K-12 to synthetic oligonucleotides containing 5-fluoro-2'-deoxycytidine. *Nucleic Acids Res*, **21**, 303-309.
32. Turek-Plewa, J. and Jagodzinski, P.P. (2005) The role of mammalian DNA methyltransferases in the regulation of gene expression. *Cellular & molecular biology letters*, **10**, 631-647.
33. Sturm, R.A. and Yaciuk, P. (1989) DNA cleavage by restriction endonuclease PflMI is inhibited in recognition sites modified by dcm methylation. *Nucleic Acids Res*, **17**, 3615.

34. Palmer, B.R. and Marinus, M.G. (1994) The dam and dcm strains of Escherichia coli--a review. *Gene*, **143**, 1-12.
35. He, X.J., Chen, T. and Zhu, J.K. (2011) Regulation and function of DNA methylation in plants and animals. *Cell Res*, **21**, 442-465.
36. Zhu, J.K. (2009) Active DNA demethylation mediated by DNA glycosylases. *Annu Rev Genet*, **43**, 143-166.
37. Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930-935.
38. Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C. and Zhang, Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*, **333**, 1300-1303.
39. Auclair, G. and Weber, M. (2012) Mechanisms of DNA methylation and demethylation in mammals. *Biochimie*, **94**, 2202-2211.
40. Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D. *et al.* (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell*, **146**, 67-79.
41. Casadesus, J. and Low, D. (2006) Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev*, **70**, 830-856.
42. 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.
43. Seshasayee, A.S. (2007) An assessment of the role of DNA adenine methyltransferase on gene expression regulation in E coli. *PLoS One*, **2**, e273.
44. Braaten, B.A., Nou, X., Kaltenbach, L.S. and Low, D.A. (1994) Methylation patterns in pap regulatory DNA control pyelonephritis-associated pili phase variation in E. coli. *Cell*, **76**, 577-588.
45. Broadbent, S.E., Davies, M.R. and van der Woude, M.W. (2010) Phase variation controls expression of Salmonella lipopolysaccharide modification genes by a DNA methylation-dependent mechanism. *Mol Microbiol*, **77**, 337-353.
46. Haagmans, W. and van der Woude, M. (2000) Phase variation of Ag43 in Escherichia coli: Dam-dependent methylation abrogates OxyR binding and OxyR-mediated repression of transcription. *Mol Microbiol*, **35**, 877-887.
47. Blyn, L.B., Braaten, B.A. and Low, D.A. (1990) Regulation of pap pilin phase variation by a mechanism involving differential dam methylation states. *EMBO J*, **9**, 4045-4054.
48. Weyand, N.J. and Low, D.A. (2000) Regulation of Pap phase variation. Lrp is sufficient for the establishment of the phase off pap DNA methylation pattern and repression of pap transcription in vitro. *J Biol Chem*, **275**, 3192-3200.

49. Reik, W. and Walter, J. (2001) Genomic imprinting: parental influence on the genome. *Nat Rev Genet*, **2**, 21-32.
50. Feinberg, A.P., Ohlsson, R. and Henikoff, S. (2006) The epigenetic progenitor origin of human cancer. *Nat Rev Genet*, **7**, 21-33.
51. Militello, K.T., Simon, R.D., Qureshi, M., Maines, R., Van Horne, M.L., Hennick, S.M., Jayakar, S.K. and Pounder, S. (2012) Conservation of Dcm-mediated cytosine DNA methylation in Escherichia coli. *FEMS Microbiol Lett*, **328**, 78-85.
52. Kahramanoglou, C., Prieto, A.I., Khedkar, S., Haase, B., Gupta, A., Benes, V., Fraser, G.M., Luscombe, N.M. and Seshasayee, A.S. (2012) Genomics of DNA cytosine methylation in Escherichia coli reveals its role in stationary phase transcription. *Nature communications*, **3**, 886.
53. Morris, K.V., Chan, S.W., Jacobsen, S.E. and Looney, D.J. (2004) Small interfering RNA-induced transcriptional gene silencing in human cells. *Science*, **305**, 1289-1292.
54. Wassenegger, M., Heimes, S., Riedel, L. and Sanger, H.L. (1994) RNA-directed de novo methylation of genomic sequences in plants. *Cell*, **76**, 567-576.
55. Zhang, H. and Zhu, J.K. (2011) RNA-directed DNA methylation. *Curr Opin Plant Biol*, **14**, 142-147.
56. Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E. *et al.* (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell*, **126**, 1189-1201.
57. Herr, A.J., Jensen, M.B., Dalmay, T. and Baulcombe, D.C. (2005) RNA polymerase IV directs silencing of endogenous DNA. *Science*, **308**, 118-120.
58. Kanno, T., Huettel, B., Mette, M.F., Aufsatz, W., Jaligot, E., Daxinger, L., Kreil, D.P., Matzke, M. and Matzke, A.J. (2005) Atypical RNA polymerase subunits required for RNA-directed DNA methylation. *Nat Genet*, **37**, 761-765.
59. Yu, B., Yang, Z., Li, J., Minakhina, S., Yang, M., Padgett, R.W., Steward, R. and Chen, X. (2005) Methylation as a crucial step in plant microRNA biogenesis. *Science*, **307**, 932-935.
60. Yang, Z., Ebright, Y.W., Yu, B. and Chen, X. (2006) HEN1 recognizes 21-24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. *Nucleic Acids Res*, **34**, 667-675.
61. Li, C.F., Pontes, O., El-Shami, M., Henderson, I.R., Bernatavichute, Y.V., Chan, S.W., Lagrange, T., Pikaard, C.S. and Jacobsen, S.E. (2006) An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. *Cell*, **126**, 93-106.
62. Zheng, X., Zhu, J., Kapoor, A. and Zhu, J.K. (2007) Role of Arabidopsis AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. *EMBO J*, **26**, 1691-1701.

63. Wierzbicki, A.T., Haag, J.R. and Pikaard, C.S. (2008) Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. *Cell*, **135**, 635-648.
64. Wierzbicki, A.T., Ream, T.S., Haag, J.R. and Pikaard, C.S. (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet*, **41**, 630-634.
65. Buhler, M., Verdel, A. and Moazed, D. (2006) Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell*, **125**, 873-886.
66. Han, J., Kim, D. and Morris, K.V. (2007) Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. *Proc Natl Acad Sci U S A*, **104**, 12422-12427.
67. Zheng, B., Wang, Z., Li, S., Yu, B., Liu, J.Y. and Chen, X. (2009) Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in Arabidopsis. *Genes Dev*, **23**, 2850-2860.
68. Hussein, R. and Lim, H.N. (2011) Disruption of small RNA signaling caused by competition for Hfq. *Proc Natl Acad Sci U S A*, **108**, 1110-1115.
69. Gottesman, S. (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet*, **21**, 399-404.
70. Wilusz, C.J. and Wilusz, J. (2005) Eukaryotic Lsm proteins: lessons from bacteria. *Nat Struct Mol Biol*, **12**, 1031-1036.
71. Marinus, M.G. (1973) Location of DNA methylation genes on the Escherichia coli K-12 genetic map. *Mol Gen Genet*, **127**, 47-55.
72. Pulvermacher, S.C., Stauffer, L.T. and Stauffer, G.V. (2009) Role of the Escherichia coli Hfq protein in GcvB regulation of oppA and dppA mRNAs. *Microbiology*, **155**, 115-123.
73. Chen, S., Zhang, A., Blyn, L.B. and Storz, G. (2004) MicC, a second small-RNA regulator of Omp protein expression in Escherichia coli. *J Bacteriol*, **186**, 6689-6697.
74. Fralick, J.A., Fisher, W.D. and Adler, H.I. (1969) Polyuridylic acid-directed phenylalanine incorporation in minicell extracts. *J Bacteriol*, **99**, 621-622.
75. Adler, H.I., Fisher, W.D., Cohen, A. and Hardigree, A.A. (1967) MINIATURE escherichia coli CELLS DEFICIENT IN DNA. *Proc Natl Acad Sci U S A*, **57**, 321-326.
76. de Boer, P.A., Crossley, R.E. and Rothfield, L.I. (1989) A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli. *Cell*, **56**, 641-649.
77. Takada, A., Wachi, M. and Nagai, K. (1999) Negative regulatory role of the Escherichia coli hfq gene in cell division. *Biochem Biophys Res Commun*, **266**, 579-583.
78. Caiafa, P. and Zampieri, M. (2005) DNA methylation and chromatin structure: the puzzling CpG islands. *J Cell Biochem*, **94**, 257-265.

79. Grkovic, S., Brown, M.H. and Skurray, R.A. (2002) Regulation of bacterial drug export systems. *Microbiol Mol Biol Rev*, **66**, 671-701, table of contents.
80. Lewis, L.K., Jenkins, M.E. and Mount, D.W. (1992) Isolation of DNA damage-inducible promoters in Escherichia coli: regulation of polB (dinA), dinG, and dinH by LexA repressor. *J Bacteriol*, **174**, 3377-3385.
81. Ogawa, W., Kim, Y.M., Mizushima, T. and Tsuchiya, T. (1998) Cloning and expression of the gene for the Na+-coupled serine transporter from Escherichia coli and characteristics of the transporter. *J Bacteriol*, **180**, 6749-6752.
82. Pulvermacher, S.C., Stauffer, L.T. and Stauffer, G.V. (2009) The small RNA GcvB regulates sstT mRNA expression in Escherichia coli. *J Bacteriol*, **191**, 238-248.
83. Lange, R., Fischer, D. and Hengge-Aronis, R. (1995) Identification of transcriptional start sites and the role of ppGpp in the expression of rpoS, the structural gene for the sigma S subunit of RNA polymerase in Escherichia coli. *J Bacteriol*, **177**, 4676-4680.
84. Lange, R. and Hengge-Aronis, R. (1994) The nlpD gene is located in an operon with rpoS on the Escherichia coli chromosome and encodes a novel lipoprotein with a potential function in cell wall formation. *Mol Microbiol*, **13**, 733-743.
85. White-Ziegler, C.A., Um, S., Perez, N.M., Berns, A.L., Malhowski, A.J. and Young, S. (2008) Low temperature (23 degrees C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in Escherichia coli K-12. *Microbiology*, **154**, 148-166.
86. Typas, A., Becker, G. and Hengge, R. (2007) The molecular basis of selective promoter activation by the sigmaS subunit of RNA polymerase. *Mol Microbiol*, **63**, 1296-1306.
87. Hengge-Aronis, R. (1993) Survival of hunger and stress: the role of rpoS in early stationary phase gene regulation in E. coli. *Cell*, **72**, 165-168.
88. Hengge, R. (2009) Proteolysis of sigmaS (RpoS) and the general stress response in Escherichia coli. *Res Microbiol*, **160**, 667-676.
89. Sledjeski, D.D., Gupta, A. and Gottesman, S. (1996) The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in Escherichia coli. *EMBO J*, **15**, 3993-4000.
90. Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. and Gottesman, S. (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A*, **95**, 12462-12467.
91. Majdalani, N., Chen, S., Murrow, J., St John, K. and Gottesman, S. (2001) Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol Microbiol*, **39**, 1382-1394.
92. Mandin, P. and Gottesman, S. (2010) Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J*, **29**, 3094-3107.

93. Lease, R.A. and Belfort, M. (2000) A trans-acting RNA as a control switch in Escherichia coli: DsrA modulates function by forming alternative structures. *Proc Natl Acad Sci U S A*, **97**, 9919-9924.
94. Brown, L. and Elliott, T. (1997) Mutations that increase expression of the rpoS gene and decrease its dependence on hfq function in Salmonella typhimurium. *J Bacteriol*, **179**, 656-662.
95. Brescia, C.C., Mikulecky, P.J., Feig, A.L. and Sledjeski, D.D. (2003) Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure. *RNA*, **9**, 33-43.
96. Resch, A., Afonyushkin, T., Lombo, T.B., McDowall, K.J., Blasi, U. and Kaberdin, V.R. (2008) Translational activation by the noncoding RNA DsrA involves alternative RNase III processing in the rpoS 5'-leader. *RNA*, **14**, 454-459.
97. Le Derout, J., Boni, I.V., Regnier, P. and Hajnsdorf, E. (2010) Hfq affects mRNA levels independently of degradation. *BMC Mol Biol*, **11**, 17.
98. Earley, K.W., Pontvianne, F., Wierzbicki, A.T., Blevins, T., Tucker, S., Costa-Nunes, P., Pontes, O. and Pikaard, C.S. (2010) Mechanisms of HDA6-mediated rRNA gene silencing: suppression of intergenic Pol II transcription and differential effects on maintenance versus siRNA-directed cytosine methylation. *Genes & development*, **24**, 1119-1132.
99. Argaman, L., Elgrably-Weiss, M., Hershko, T., Vogel, J. and Altuvia, S. (2012) RelA protein stimulates the activity of RyhB small RNA by acting on RNA-binding protein Hfq. *Proc Natl Acad Sci U S A*.
100. Soper, T., Mandin, P., Majdalani, N., Gottesman, S. and Woodson, S.A. (2010) Positive regulation by small RNAs and the role of Hfq. *Proc Natl Acad Sci U S A*, **107**, 9602-9607.
101. Low, D.A. and Casadesus, J. (2008) Clocks and switches: bacterial gene regulation by DNA adenine methylation. *Curr Opin Microbiol*, **11**, 106-112.
102. Wickiser, J.K., Winkler, W.C., Breaker, R.R. and Crothers, D.M. (2005) The speed of RNA transcription and metabolite binding kinetics operate an FMN riboswitch. *Mol Cell*, **18**, 49-60.
103. Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juarez, K., Contreras-Moreira, B. *et al.* (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in E. coli. *PLoS One*, **4**, e7526.
104. Wu, T.H., Chang, I.Y., Chu, L.C., Huang, H.C. and Ng, W.V. (2010) Modularity of Escherichia coli sRNA regulation revealed by sRNA-target and protein network analysis. *BMC Bioinformatics*, **11 Suppl 7**, S11.
105. Wyszynski, M.W., Gabbara, S., Kubareva, E.A., Romanova, E.A., Oretskaya, T.S., Gromova, E.S., Shabarova, Z.A. and Bhagwat, A.S. (1993) The cysteine conserved among DNA cytosine methylases is required for methyl transfer, but not for specific DNA binding. *Nucleic Acids Res*, **21**, 295-301.
106. Salim, N.N., Faner, M.A., Philip, J.A. and Feig, A.L. (2012) Requirement of upstream Hfq-binding (ARN)x elements in glmS and the Hfq C-terminal

region for GlmS upregulation by sRNAs GlmZ and GlmY. *Nucleic Acids Res*, **40**, 8021-8032.

ABSTRACT

INVESTIGATION OF BACTERIAL RNA-DIRECTED DNA METHYLATION VIA DCM AND HFQ

by

DANDAN LI

May 2013

Advisor: Dr Andrew Feig

Major: Chemistry (Biochemistry)

Degree: Master of Science

Bacterial small RNAs and the RNA chaperone Hfq play crucial roles in post-transcriptional gene regulation, often as parts of stress-response pathways, but little is known about their roles in regulation of gene transcription. A recent report showed that changes in methylation patterns caused by DNA cytosine methyltransferase (Dcm) were linked to gene regulation occurring during the transition to stationary phase. Here, we show that Dcm involves in the stress responses under nutrient starvation and cold stress. Dcm and Hfq together mediate gene expression under cold stress. Hfq promotes Dcm-catalyzed cytosine methylation at specific sites near the *rpoS* promoter, which is consistent with the genome-wide analysis and linking known stress response pathways to altered methylation. Overexpressing DsrA, an sRNA induced at low temperature to regulate genes required for cold adaptation, stimulates this DNA methylation behavior, showing that the regulation is sRNA-dependent. This represents the first example of an RNA-directed DNA methylation mechanism in bacteria responsible for modulating gene expression.

AUTOBIOGRAPHICAL STATEMENT

Dandan Li

Educational background

Master of Science (Sep 2007-Apr 2013): Advisor: Dr Andrew Feig, Department of Chemistry, Wayne State University, Detroit, Michigan. Thesis title Investigation of Bacterial RNA Directed DNA Methylation via Dcm and Hfq.

Bachelor of Science (Sep 2003-Jun 2007): Fujian Normal University, Fuzhou, Fujian, P.R.China.