



HAL
open science

Regulation of the *yifK* locus by multi-target small RNA GcvB in *Salmonella*

Qi Yang

► **To cite this version:**

Qi Yang. Regulation of the *yifK* locus by multi-target small RNA GcvB in *Salmonella*. Agricultural sciences. Université Paris Sud - Paris XI, 2013. English. NNT : 2013PA112160 . tel-00954416

HAL Id: tel-00954416

<https://theses.hal.science/tel-00954416>

Submitted on 14 Mar 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE PARIS-SUD

ÉCOLE DOCTORALE : GÈNES GÉNOMES, CELLULES (ED 426)

Laboratoire : “Régulation génique chez Salmonella et ses phages”

Centre de Génétique Moléculaire du CNRS, UPR3404

DISCIPLINE : GÉNÉTIQUE MOLÉCULAIRE

THÈSE DE DOCTORAT

soutenue le 19/09/2013

par

Qi YANG

Regulation of the *yifK* locus by multi-target
small RNA GcvB in Salmonella

Directeur de thèse : Lionello BOSSI Directeur de Recherche, CNRS

Composition du jury :

Président du jury :	Philippe BOULOC	Directeur de Recherche, CNRS, UMR8621, Orsay
Rapporteurs :	Philippe BOULOC	Directeur de Recherche, CNRS, UMR8621, Orsay
	Francis REPOILA	Directeur de Recherche, INRA, Jouy en Josas
Examineurs :	Claudio GUALERZI	Professeur, Université de Camerino, Italie

TABLE OF CONTENTS

ABSTRACT.....	4
RESUMÉ.....	5
I. INTRODUCTION.....	6
I.1. <i>Salmonella</i> : general features.....	6
I.2. Gene Expression.....	7
I.2.1. Transcription.....	8
I.2.1.1. Alternative Sigma Factors.....	9
I.2.1.2. Transcriptional regulators.....	10
I.2.1.3. Transcription termination.....	11
I.3. Translation.....	12
I.3.1. S1 protein.....	14
I.3.2. Translation initiation factors.....	15
I.4. Global transcription regulatory networks: selected examples.....	15
I.4.1. The leucine response regulator Lrp.....	15
I.4.2. The glycine cleavage system.....	17
I.5. Regulation by RNA.....	19
I.5.1. Riboswitches.....	19
I.5.2. CRISPR systems.....	21
I.5.3. Small regulatory RNAs.....	23
I.5.3.1. Protein-targeting sRNAs.....	23
I.5.3.2. RNA-targeting sRNAs.....	24
I.5.3.2a. Cis-encoded sRNAs.....	24
I.5.3.2b. Trans-encoded sRNAs.....	25
I.5.3.3. Hfq: mediator of sRNA activity.....	27
I.5.3.3a. Hfq structure and RNA binding patterns.....	28
I.5.3.3b. Mechanism of action of Hfq.....	30
I.5.3.4 Global sRNA-dependent regulatory networks: selected examples	31
I.5.3.4a. RybB and iron homeostasis.....	31
I.5.3.4b. GcvB and amino acid management.....	33

II. RESULTS.....	36
II.1. Characterization of <i>yifK</i> , a Hfq-regulated locus in <i>Salmonella</i>	36
II.1.1. <i>yifK</i> is repressed by GcvB.....	37
II.1.2. Genetic analysis of GcvB: <i>yifK</i> interaction unveils a translational enhancer in the target sequence.....	40
II.1.3. Separate subdomains of GcvB R1 region pair with <i>yifK</i> and <i>dppA</i>	41
II.1.4. Role of ACA motifs in translation enhancement.....	42
II.2. Manuscript submitted to PloS Genetics.....	44
Table 1.....	60
II.2.1 Supporting information of manuscript.....	72
Table S1.....	75
Table S2.....	77
II.3. An intriguing regulator of <i>yifK</i> expression.....	78
II.4. A specialized mutant ribosome improves initiation from a UUG start codon	83
II.5. Article II.....	86
III. DISCUSSION AND PERSPECTIVES	110
IV. MATERIALS AND METHODS.....	115
Table 2.....	124
Table 3.....	125
Table 4.....	126
V. REFERENCES.....	127

ABSTRACT

GcvB is a conserved 200 nucleotide RNA that downregulates several genes involved in amino acid uptake or biosynthesis in bacteria. The physiological role of GcvB action is not entirely clear, but it is likely aimed at balancing of nutritional resources under fast growth conditions. GcvB inhibits translation of target messenger RNAs by pairing with sequences inside or upstream of ribosome binding sites. In the present study, characterization of a novel GcvB-regulated locus revealed some unique features in the mode of functioning of this regulatory RNA. We found that GcvB represses *yifK* - a highly conserved locus encoding a putative amino acid transporter - by targeting a translational enhancer element. Two ACA motifs within the target sequence are the main determinants of the enhancer activity. Replacing either of these motifs with random triplets caused up to a 10-fold decrease in *yifK* expression regardless of the GcvB allele (deleted or suitably modified to recognize the mutated target). It thus appears that GcvB effectiveness as a regulator results from countering the enhancer activity. When the enhancer is removed, GcvB action no longer constitutes a rate-limiting factor for *yifK* expression. Overall, this study is relevant not only to a better understanding of GcvB function but it also provides insight into an elusive aspect of the translation initiation process.

Besides the GcvB control, the *yifK* locus is regulated at the transcriptional level by the leucine responsive regulator Lrp, and by HdfR (YifA) a poorly known transcriptional regulator, that appears to require the product of the adjacent, divergently oriented gene, *yifE*, for expression or activity. Transcription initiating at the *yifK* promoter extends into the adjacent *argX-hisR-leuT-proM* tRNA operon yielding an unusual primary transcript which both a messenger RNA and a tRNA precursor. This chimeric RNA is rapidly processed by RNase E.

RÉSUMÉ

GcvB est un ARN bactérien conservé de 200 nucléotides, qui régule négativement l'expression de plusieurs gènes impliqués dans l'import et la biosynthèse des acides aminés. Bien que le rôle physiologique de GcvB ne soit pas complètement élucidé, il contribuerait vraisemblablement à équilibrer les ressources nutritionnelles en conditions de croissance rapide. GcvB inhibe la traduction des ARNm cibles en s'appariant avec des séquences à l'intérieur ou en amont du site de liaison du ribosome. Dans cette étude, la caractérisation d'un nouveau locus régulé par GcvB a permis de dévoiler des aspects singuliers du mode de fonctionnement de cet ARN régulateur. Nous avons découvert que GcvB réprime *yifK* - un gène très conservé, codant pour un transporteur d'acides aminés putatif - en ciblant un élément activateur de la traduction sur l'ARNm. Deux motifs ACA dans la séquence cible sont les déterminants principaux de la fonction activatrice. Le remplacement de l'un ou l'autre avec des triplets aléatoires, provoque une diminution de 10 fois du niveau d'expression de *yifK*, quelque soit l'allèle de GcvB (délétion ou changement de séquence permettant la reconnaissance de la cible mutante). Il apparaît ainsi que l'efficacité de GcvB à réguler négativement sa cible serait liée à sa capacité d'antagoniser l'élément activateur. Lorsque l'activateur est éliminé, l'action de GcvB n'est plus un facteur limitant pour l'expression de *yifK*. Dans son ensemble, cette étude apporte une meilleure compréhension de la fonction de GcvB et révèle un nouvel aspect du processus d'initiation de la traduction.

En plus du contrôle par GcvB, le locus *yifK* est régulé au niveau transcriptionnel par Lrp (leucine-responsive regulatory protein) et par HdfR (YifA) un régulateur transcriptionnel peu connu qui requerrait le produit du gène adjacent orienté de façon divergente, *yifE*, pour son expression ou activité. Enfin, la transcription initiée au niveau du promoteur *yifK* s'étend dans l'opéron d'ARNt *argX-hisR-leuT-proM* adjacent, donnant lieu à un transcrit primaire qui est à la fois un l'ARNm et un précurseur des ARNt. Cet ARN chimère est rapidement maturé par l'ARNase E.

I. INTRODUCTION

I.1. *Salmonella*: general features

Salmonella is a Gram-negative, rod shaped Gamma proteobacterium of the *Enterobacteriaceae* family. Almost a century ago, the classification system of various *Salmonella* serotypes based on surface antigens was established by Kauffman and White (Le Minor & Bockemuhl, 1984). The World Health Organization Collaborating Centre for Reference and Research on Salmonella (WHOCC-Salm) has reported in 2007 (supplement data in 2010), that there are 2610 serotypes in the genus *Salmonella* (Patrick A.D. Grimont, 2007, Guibourdenche *et al.*, 2010). All these serotypes can be grouped into two species: *Salmonella enterica* and *Salmonella bongori*. Chromosomal DNA hybridization experiments (Crosa *et al.*, 1973, Le Minor *et al.*, 1982, Le Minor *et al.*, 1986), and multi-locus enzyme electrophoresis (MLEE) (Reeves *et al.*, 1989, Beltran *et al.*, 1988) have led to a further subdivision of *Salmonella enterica* into six subspecies: I (*enterica*), II (*salamae*), IIIa (*arizonae*), IIIb (*diarizonae*), IV (*houtenae*) and V (*indica*).

In order to harmonize the designation and writing of *Salmonella* species and serovars, it has been agreed that species are written in italics following the international convention and the serovar is specified after the species, non-italicized and starting with a capital letter (Brenner *et al.*, 2000, Tindall *et al.*, 2005, Maloy & Hughes, 2007) For example, “*Salmonella enterica* serovar Typhimurium”.

Salmonella is a common and important pathogen found all worldwide (Galanis *et al.*, 2006) that can infect a wide range of hosts, including humans, animals and plants (Gu *et al.*, 2011). Particularly important is the incidence of *Salmonella*-caused foodborne diseases in industrialized countries (Le Hello *et al.*, 2012, Scallan *et al.*, 2011, Majowicz *et al.*, 2010). The pathogenicity of *Salmonella* largely relies on its ability to corrupt the functioning of host cells by secreting virulence factors and effector proteins through specialized type III secretion systems (TTSS). Two distinct type III secretion systems, encoded

by *Salmonella* pathogenicity islands (SPI) 1 and 2 are the most important virulence-related TTSS. The function of secretory system encoded by SPI-1 is to provide ability to invade epithelial cells causing the inflammatory response response (Hapfelmeier *et al.*, 2004). The function of secretory system encoded by SPI-2 allows bacteria to survive and reproduce within epithelial cells and phagocytes (Hensel *et al.*, 1997, Shea *et al.*, 1999, Hensel, 2000).

Phylogenetically, *Salmonella* is related to *Shigella*, *Citrobacter* and *Escherichia* (Groisman & Ochman, 1997). For example, the two laboratory model strains *Escherichia coli* K12 and *Salmonella enterica* serovar Typhimurium (hereafter referred to as *S. Typhimurium*) strain LT2 can be considered highly related at the genomic level (Ochman & Wilson, 1987, Doolittle *et al.*, 1996). The sequence of the genome of *S. Typhimurium* LT2, isolated in the 1940s and used in the first studies on phage-mediated transduction, was completed in 2001 (McClelland *et al.*, 2001). It comprises a circular chromosome of 4,657,432 base-pairs (bp) and the so-called virulence plasmid pSLT (93,939 bp). DNA sequence identity between *S. Typhimurium* LT2 and *E. coli* K12 is around 80%, amino acid similarity of these two organisms reaches to 90% (McClelland *et al.*, 2001).

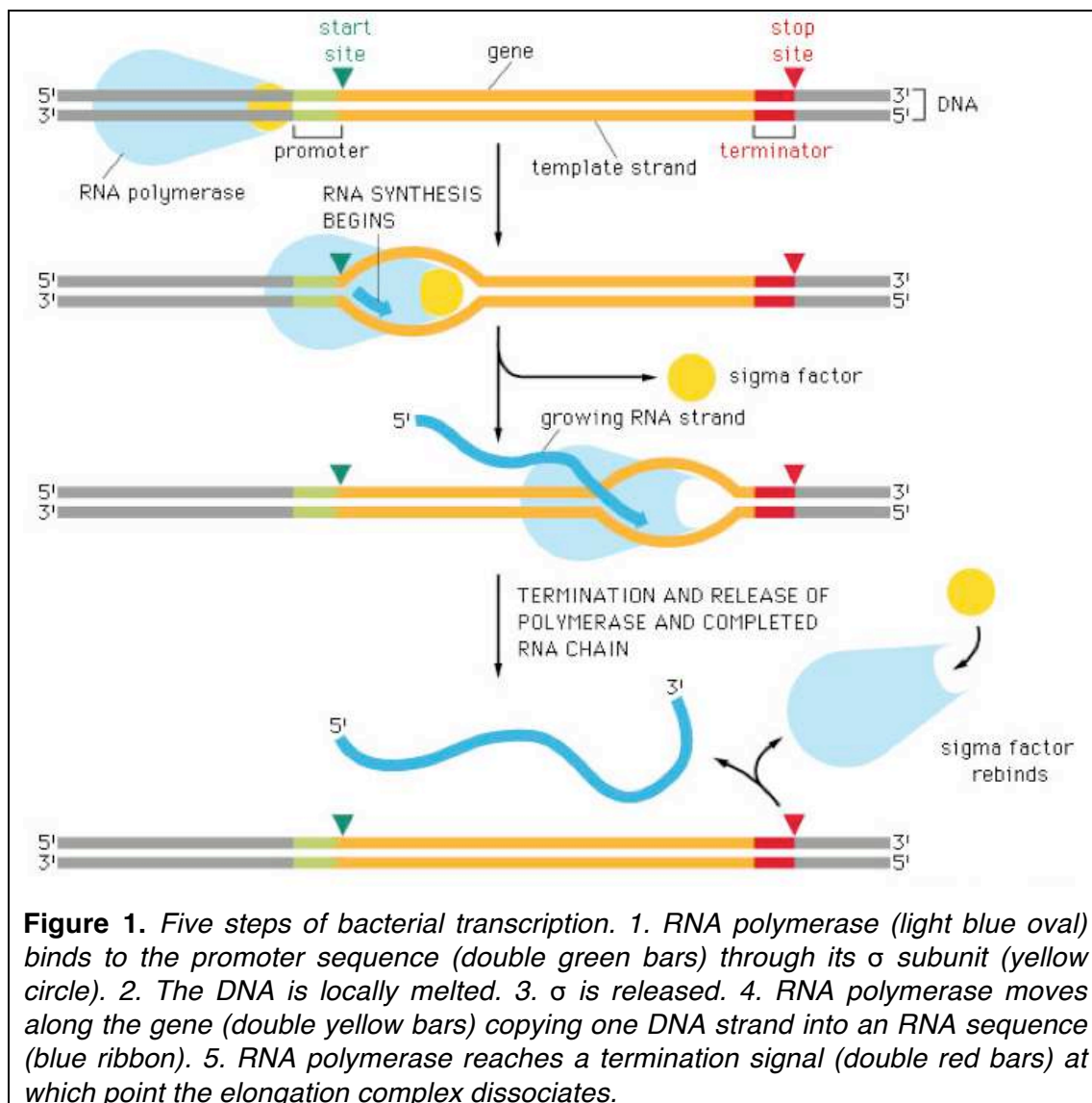
Besides being an important model pathogen, *Salmonella* has constituted a model system for genetic and biochemical analysis since the 1950s (Sanderson *et al.*, 1995). Many tools have been developed and some of the key bacterial cellular mechanisms have been discovered by work in this organism.

I.2. Gene Expression

Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product. These products are often proteins, but in non-protein coding genes, the products also can be ribosomal RNA (rRNA), transfer RNA (tRNA) and small non-coding RNA (sRNA). The passage from gene to protein, comprises two main steps: transcription and translation.

I.2.1. Transcription

Transcription is the process by which a region of the DNA molecule, corresponding to a particular gene, is selected by an enzyme called RNA polymerase to synthesize a complementary single-stranded RNA using one of the strands of the DNA as a template. The process of transcription consists of five discrete stages: promoter recognition, local unwinding, chain initiation, chain elongation and chain termination (Fig. 1).



RNA polymerase binds to double-stranded DNA within a specific base sequence (promoter, 20-40 bases long). After the initial binding step, the open promoter complex is formed by RNA polymerase and unwinding of DNA

occurs. Then, the RNA polymerase selects a transcription start site, which is close to the initial binding site. The first nucleoside triphosphate is added at this site and synthesis begins. RNA polymerase then moves along the DNA in the 5' to 3' direction adding nucleotides to the 3'-OH of the growing RNA chain. When RNA polymerase reaches a termination sequence, both the newly synthesized RNA and the RNA polymerase are released.

Typically, prokaryotic promoters consist of two conserved regions, one localized about 10 base pairs (bp) upstream from the site of initiation (the so-called -10 box), the other at about 35 bp from the start site (-35 box). Also part of the promoter architecture is the spacing between these two elements, 17 bp being considered optimal. These recognition sequences instruct RNA polymerase as to where to start transcribing.

1.2.1.1. Alternative sigma factors

Bacterial RNA polymerase consists of five protein subunits, four of the subunits forming the core enzyme, which catalyzes the polymerization of nucleoside triphosphates into RNA. The fifth subunit, called the σ subunit, is required for promoter binding. Once polymerization begins, the σ subunit dissociates from the core enzyme, when transcription is completed, the core enzyme binds another σ subunit and is then ready to bind to a promoter again. Bacteria makes use of several different σ subunits, also called alternative σ factors, to regulate the coordinate expression of groups of genes (for a review, see (Gruber & Gross, 2003)). Most commonly, polymerase associates with σ^{70} , the most abundant subunit, capable of recognizing the promoters of genes needed to maintain the main physiological processes of the growing cell (housekeeping genes). In addition, and in response to different environmental and internal stimuli, bacteria can use:

σ^{38} (also known as σ^S): stationary phase and starvation/stress factor;

σ^{24} (σ^E): extracytoplasmic stress sigma factor;

σ^{32} (σ^H): heat shock sigma factor;

σ^{54} (σ^N): Nitrogen-limitation sigma factor;

σ^{28} (σ^F): Flagellar sigma factor;

σ^{19} : Ferric citrate sigma factor; regulates the *fec* gene for iron transport; The genes under control of a given σ factor share a specific sequence in their promoter regions that allows recognition of the gene by the corresponding factor. σ will bring along the RNA core polymerase when it binds to the DNA and will promote transcription of the gene. In this way, bacteria can coordinately express groups of genes upon activating a specific σ factor. Binding of σ factors to core RNA polymerase is thought to be mainly dependent on the relative concentration of the different factors. Thus, conditions that lead to increase in the concentration of a given σ factor, will entail the preferential transcription of the genes belonging to its regulon.

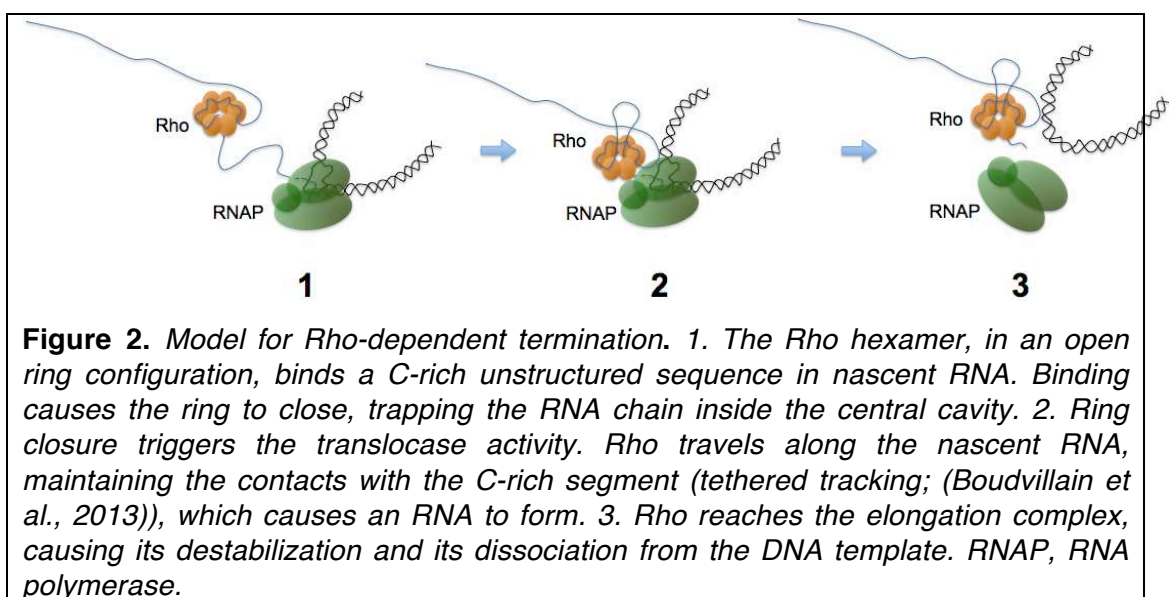
I.2.1.2. Transcriptional regulators

A second main strategy for regulating transcription is the use of specific transcription regulators, that is proteins that can either repress or activate transcription. Bacterial repressors typically bind to an “operator” locus, located downstream from the promoter of target genes (often comprised in the transcript) to prevent transcription of the structural gene(s). Activators facilitate the expression of specific genes by helping RNA polymerase to initiate transcription either through direct binding to the enzyme, thus increasing its affinity for the promoter, and/or by locally modifying the conformation/structure of the DNA to increase transcriptional rate (Lee *et al.*, 2012). There exists over 300 transcription regulators *in E. coli*; some of them work alone and others associate with ligands or auxiliary factors to perform their regulatory action. To illustrate the complexity of transcriptional regulatory networks, the number of transcriptional factors working on a single bacterial promoter can be as high as 30 (Ishihama, 2012). Finally, DNA itself can modulate the transcription rate by its physical properties. Although more seldom than transcription factors, DNA bending and supercoiling can also affect the initiation step of transcription (Bossi & Smith, 1984, Figueroa *et al.*, 1991) reviewed by (Travers & Muskhelishvili, 2005).

I.2.1.3. Transcription termination

Two kinds of termination events are well known; one is dependent on the DNA sequence only (intrinsic termination), the other requires the participation of a termination protein factor called Rho (Rho-dependent termination). Most intrinsic terminators (or Rho-independent terminators) contain an inverted repeated sequence, that can base pair with itself once it has been transcribed, to create a stem loop. Usually the hairpin loop is 7-20 bp long and is rich in G:C pairs, making a particularly stable structure that causes polymerase to temporarily pause at this point. Immediately following the stem-loop, there is a stretch of U residues (6 to 8) which form a weak RNA:DNA duplex. This destabilizes the elongation complex and causes its dissociation from the DNA template (reviewed in (Peters *et al.*, 2011)).

Rho-dependent termination is governed by sequences in the nascent RNA. The process leading to termination requires three steps: In the first step, Rho factor, an hexameric protein, binds to C-rich unstructured region in the RNA. This induces a conformational rearrangement that causes the RNA chain to enter in the central cavity of the hexamer (Fig. 2). The rearrangement triggers the RNA-dependent ATPase activity that powers translocation (Richardson, 2003, Boudvillain *et al.*, 2013). In the second step of Rho termination, RNA is translocated 5' to 3' through the central cavity of Rho. In the final step of Rho-dependent termination, an elongation complex that has been halted at a pause site is dissociated by Rho action.

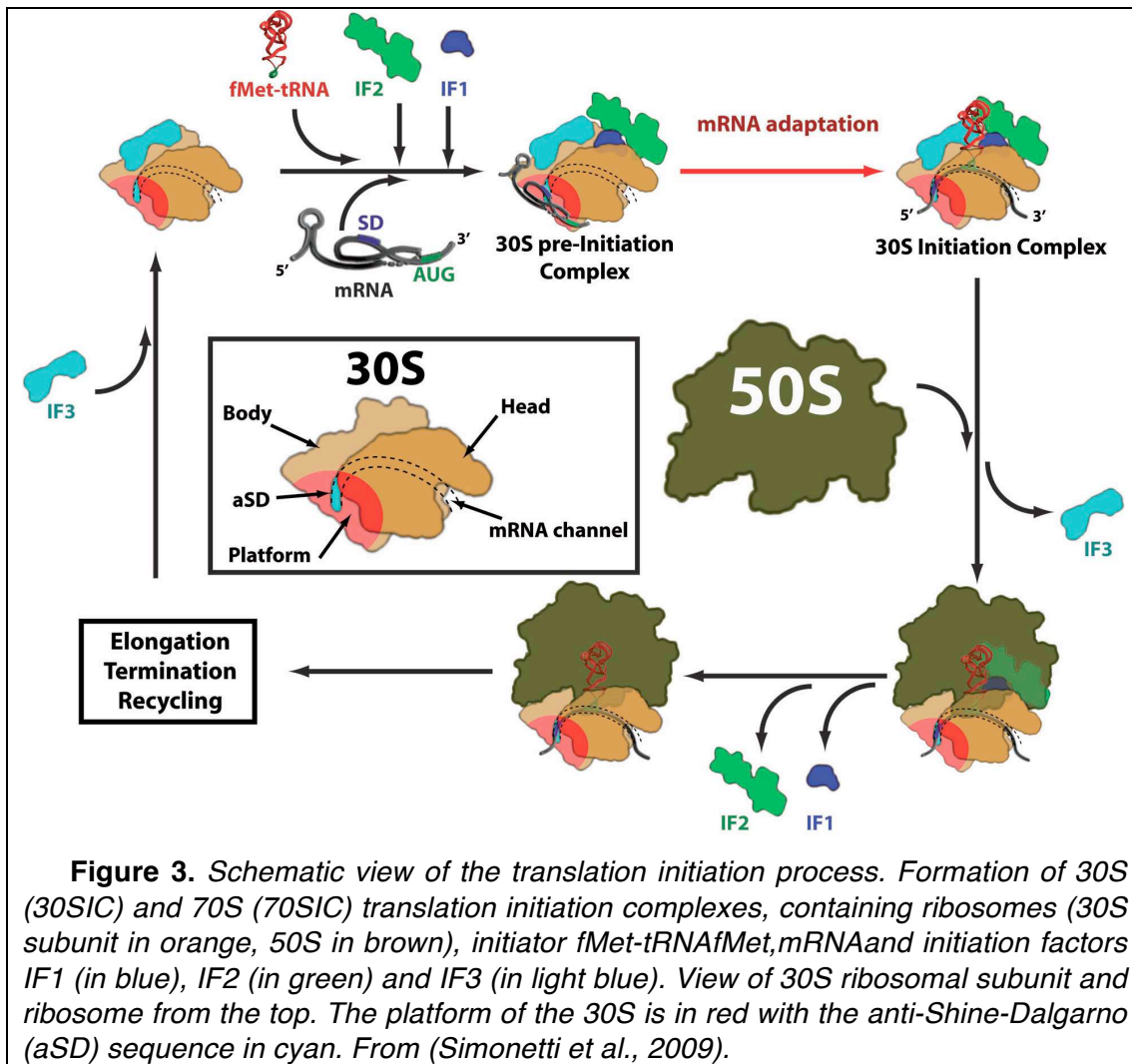


I.3. Translation

Translation is most complexly regulated step of protein synthesis in all three domains (Archaea, Bacteria and Eukarya) of life. The translation cycle can be schematically divided into three parts: initiation, elongation and termination. The translation initiation is the process that brings together mRNA, initiator transfer RNA, and the ribosome (Fig. 3).

In bacteria, translation initiation proceeds through several steps involving the formation of three major intermediary initiation complexes (Simonetti *et al.*, 2009, Myasnikov *et al.*, 2009, Laursen *et al.*, 2005). The first complex starts assembling when initiation factors IF2 and IF3 bind the 30S ribosomal subunit to form an unstable 30S-IF2-IF3 complex. Then there is the arrival of initiation factor IF1, which locks the factors in a stable 30S pre-initiation complex (PIC). At this point, fMet-tRNA^{fMet} will bind. Recent evidence indicates that binding of mRNA is independent of initiation factors and can take place any time during 30S PIC assembly (Milon *et al.*, 2012).

The second step is the formation of the 30S initiation complex through a conformational change which brings the first codon of the mRNA, usually AUG, to physically interact with the anticodon of the initiator tRNA in the 30S peptidyl (P) site. Interaction of this complex with the 50S subunit occurs and IF2 hydrolyzes GTP (Grigoriadou *et al.*, 2007). The 70S initiation complex (70SIC) forms and initiation factors IF1 and IF3 are released. The third step is the release of IF2.GDP with 70S initiation complex now ready for the elongation step of protein synthesis.



In bacteria, AUG is the most common translation initiator codon because of its perfect pairing with the CAU anticodon in fMet-tRNA^{fMet}. However, weaker pairing initiator codons (where two rather than three bases can pair with fMet-tRNA anticodon) also exist in bacteria. For example, 14% of *E. coli* genes use GUG as the start codon and another 3% use UUG; only two *E. coli* genes use AUU as a start codon (Blattner *et al.*, 1997). UUG as a start codon is more common in Gram-positive bacteria and in some bacteriophages (Kunst *et al.*, 1997, Lobočka *et al.*, 2004). Weaker pairing initiation codons can affect translation efficiency. For instance, in *E. coli*, when AUG changed by GUG or UUG, the translation efficiency *in vivo* was reduced. Change to AUU had an even bigger effect than GUG or UUG (Sussman *et al.*, 1996).

In prokaryotes, the mRNA sequence that directs the interaction between mRNA and the ribosome, upstream from start codon, called the Shine–Dalgarno (SD) sequence (Shine & Dalgarno, 1974) plays the main role in initiation. This purine-rich sequence is complementary to, and base pairs with, a sequence near the 3' end of 16S rRNA, the RNA component in the 30S ribosomal subunit (Steitz & Jakes, 1975, Jacob *et al.*, 1987, Brink *et al.*, 1995, Yusupova *et al.*, 2001). The SD sequence (GGAGG) is typically 4 or 5 nucleotides in length, and it is usually positioned 4–8 nt upstream from the start codon. The optimal spacing depends on exactly which bases at the 3' end of 16S rRNA (3'-AUUCCUCCAC5') participate in the interaction (Chen *et al.*, 1994). Many mRNAs contain pyrimidine-rich (AU-rich in *E. coli*) sequences in the region upstream from the SD sequence. These sequences binds to ribosomal protein S1 (Boni *et al.*, 1991, Yusupova *et al.*, 2001, Laursen *et al.*, 2005, Komarova *et al.*, 2005b), A direct interaction between protein S1 and mRNA has been confirmed by cryoelectron microscopy (EM) studies in 2001 (Sengupta *et al.*, 2001).

1.3.1. S1 protein

S1 protein is an mRNA-binding protein associated with the translational machinery (Sorensen *et al.*, 1998). As a ribosomal component, it recognizes and binds to mRNAs affecting translation efficiency during initiation phase (Boni *et al.*, 1991, Zhang & Deutscher, 1992, Nakagawa *et al.*, 2010), this function being relatively independent of the “strength” of SD sequence (Tzareva *et al.*, 1994, Komarova *et al.*, 2002). In contrast, in leaderless mRNA (mRNAs beginning directly with the AUG initiating codon), S1 has proven to be dispensable for translation initiation (Tedin *et al.*, 1997, Moll *et al.*, 2002). Furthermore, because of the interaction between S1 and the AU-rich sequences in mRNA, it is thought that AU-rich sequences act as translational enhancers (Komarova *et al.*, 2005a, O'Connor & Dahlberg, 2001, Hook-Barnard *et al.*, 2007). By its association with these sequences, S1 could also help stabilize mRNA by protecting it against cleavage by RNase E, an

endonuclease with similar sequence recognition patterns (Komarova et al., 2005b) in mRNA.

I.3.2. Translation initiation factors

As described above, three non-ribosomal proteins - the initiation factors (IFs) - play important roles during the translation. IFs are highly conserved and homologous proteins are present in all three domains of life (Kyrpides & Woese, 1998, Sorensen *et al.*, 2001).

IF1, encoded by the *infA* gene, is the smallest of the initiation factors (Pon *et al.*, 1979). IF1 stimulates IF2 and IF3 activities by increasing their binding efficiency to the 30S subunit (Wintermeyer & Gualerzi, 1983, Pon & Gualerzi, 1984) and can occupy the aminoacyl (A) side of ribosome to prevent tRNA binding (Moazed *et al.*, 1995).

IF2 is the product of the *infB* gene and largest of the initiation factors. It belongs to the family of the GTP-GDP binding protein (GTPase) (Bourne *et al.*, 1991). The presence of IF2 on the 30S subunit promotes the binding of initiator tRNA (Canonaco *et al.*, 1986, Gualerzi & Pon, 1990).

IF3 is encoded by the *infC* gene (Sacerdot *et al.*, 1982). Several roles have been suggested for IF3 in translation initiation. For example, IF3 stimulates the rapid formation of codon-anticodon interaction at the ribosomal P-site (Gualerzi *et al.*, 1977, Wintermeyer & Gualerzi, 1983) and it is involved in the adjustment of the mRNA from the standby site to the decoding P-site of the 30S ribosomal subunit (La Teana *et al.*, 1995). IF3 prevents the formation of 70S ribosome by binding to the 30S subunit (Grunberg-Manago *et al.*, 1975, Sacerdot *et al.*, 1996).

I.4. Global transcription regulatory networks: selected examples

I.4.1. The leucine response regulator Lrp

Lrp is a major bacterial global transcriptional regulator controlling at least 10% of all *E.coli* genes, according to transcriptomic data (Tani *et al.*, 2002). In most instances, its regulatory action (negative or positive) operates at the level of transcription of genes involved in amino acid metabolism and

transport. However, some genes participating in other cellular functions, particularly fimbrial operons (Baek *et al.*, 2011), and pilus assembly (van der Woude *et al.*, 1992) are also part of the regulon.

Lrp belongs to a family of regulators that is widely distributed among bacteria and Archaea. Lrp is a relatively small protein (18.8 kD in *E. coli*) that can exist in the cell as a dimer but also further associate in higher order structures, especially octamers (Chen & Calvo, 2002) and/or hexadodecamers. (de los Rios & Perona, 2007). Alternation between the two latter forms is thought to correspond to the autotrophic versus heterotrophic mode, octamer formation being stimulated by leucine (see below; (Kawashima *et al.*, 2008)). Although there is no strict consensus sequence for Lrp binding, a modified “Systematic Evolution of Ligands by Exponential Enrichment” (SELEX) procedure has identified a sequence YAGHAWATTWTDCTR, where Y = C or T, H = not G, W = A or T, D = not C, and R = A or G (Cui *et al.*, 1995).

The biological significance of the multimeric structure of Lrp is underlined by the fact that many promoters controlled by Lrp have multiple binding sites. The best characterized member of the Lrp family is the 164 amino acids Lrp protein from *E. coli*. Like most transcriptional regulators it has a DNA binding domain and a ligand binding site. The amino acid leucine is such a ligand, which modulates the action of Lrp. Since L-leucine is the most abundant amino acid in proteins (9% occurrence for leucine, followed by 7.5% for alanine), it has been speculated that the bacterial cell could use this amino acid as a reporter for the availability of proteins, peptides and amino acids (Brinkman *et al.*, 2003). Upon binding to leucine, the efficiency of Lrp action, either negative or positive, is further stimulated or reduced (reviewed in (Brinkman *et al.*, 2003)). Lrp, together with the regulator of asparagine synthase C gene product, AsnC, are called feast/famine regulatory proteins, by virtue of their implication in regulation of metabolic pathways as a function of amino acid and nitrogen bases availability in the environment (for a review, see (Kawashima *et al.*, 2008)).

Lrp can also function independently of leucine (as in its self-control mechanism) or it can associate with other regulators (CRP, H-NS or IHF). Because of Lrp's relatively abundance (3000 Lrp dimers per *E. coli* cell; (Willins *et al.*, 1991)), and since it appears to be involved in DNA condensation and DNA bending (Wang & Calvo, 1993), it has been proposed that it can also function as non-specific DNA organizer, depending on nutritional availability and phase of growth.

In *Salmonella enterica*, the Lrp protein is 99% identical to that of *E. coli* at the aminoacid level and controls also several virulence genes. The expression of virulence plasmide *spv* operon (Marshall *et al.*, 1999) and conjugal transfer gene *traJ* (Camacho & Casadesus, 2005) are under Lrp control. In addition, virulence regulators *hilA*, *invF*, and *ssrA* have also been shown to be controlled by Lrp at the transcriptional level (Baek *et al.*, 2009) .

It has been shown that constitutive expression of the *lrp* gene attenuates *Salmonella* virulence, whereas deletion of the gene enhances *Salmonella* invasion of cultured cell lines and virulence on BALB/C mice (Baek *et al.*, 2009). Because of the global nature of Lrp action and the complexity of the network, the precise role of the protein in several of the described mechanisms remains to be fully understood.

1.4.2. The glycine cleavage system

De novo biosynthesis of several cellular constituents, including nucleotides and amino acids, requires one-carbon units supplied directly as formyl tetrahydrofolate (CH₂-H₄ folate). There are two pathways leading to CH₂-H₄ folate production. The major pathway involves the enzyme serine-hydroxymethyl-transferase (the product of the *glyA* gene) which converts serine to glycine plus CH₂-H₄ folate (Stauffer, 1987). In bacteria growing in a glucose-supplemented minimal medium – conditions where the requirement for one-carbon units is the highest – as much as 15% of the carbon from glucose is thought to be channeled to the formation of serine and glycine (Calvo & Matthews, 1994). A second route for one-carbon unit production involves the oxidative cleavage of glycine to NH₃ and CO₂ and the transfer of

a C-1 methylene unit to folate. This reaction, which is conserved in all domains of life, is carried out by the four enzymes of the glycine cleavage system: the P, T, L and H proteins. Since the serine-hydroxymethyltransferase reaction can run in both directions, the glycine cleavage pathway can also be a source of serine as CH₂-H₄ folate can donate a carbon to glycine to form serine (Jourdan & Stauffer, 1999, Lin *et al.*, 1992).

In *E. coli* and *S. Typhimurium*, the genes encoding the P,T and H components of the glycine cleavage systems are organized in an operon (*gcvTHP*) while the L protein is encoded by a gene at a separate chromosomal location. Transcription of *gcvTHP* is activated in the presence of glycine (Wilson, 1993 #1039) and repressed by purines (Wilson *et al.*, 1993). Key mediator of both responses is the GcvA protein, a LysR-type regulator that binds to three sites in the region upstream from the *gcvTHP* promoter. While GcvA alone is required for glycine-dependent activation, purine-induced repression requires both GcvA and a second regulatory protein, GcvR (Ghrist & Stauffer, 1995). Stauffer and coworkers have proposed that GcvA homocomplexes function as activators while GcvA-GcvR heterocomplexes repress *gcvTHP* transcription (Ghrist & Stauffer, 1995, Jourdan & Stauffer, 1999). In addition to these specific regulators, global regulatory protein Lrp is required for both activation and repression of the *gcvTHP* operon (Lin *et al.*, 1992, Stauffer & Stauffer, 1998, Stauffer & Stauffer, 1999). Lrp is thought to play primarily a structural role by bending the DNA and allowing GcvA to make contact with separate sites on the DNA or with GcvR (Fig. 4).

This action (independent of leucine) is essential for GcvA-mediated activation of *gcvTHP* as this operon is essentially uninducible in a strain devoid of Lrp (Lin *et al.*, 1992, Stauffer & Stauffer, 1998). Since Lrp levels are significantly reduced in cells growing in rich medium, the *gcvTHP* operon is expected to be expressed at low levels under such conditions, consistent with the reduced demand from one-carbon units in rich nutritional environments (Calvo & Matthews, 1994).

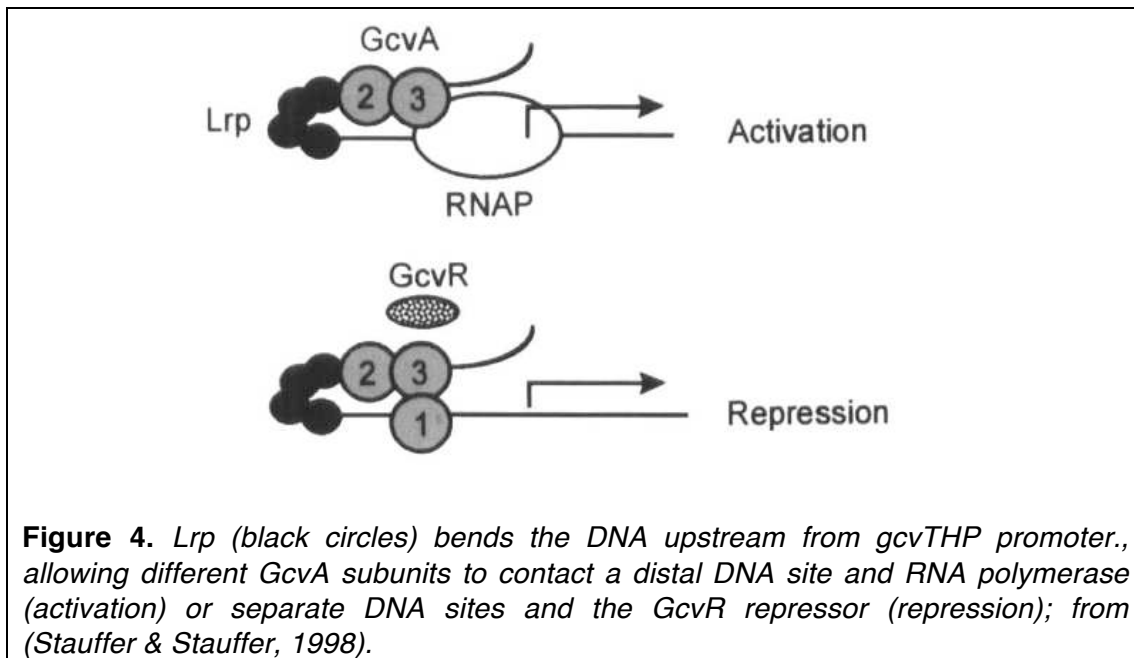


Figure 4. *Lrp* (black circles) bends the DNA upstream from *gcvTHP* promoter., allowing different *GcvA* subunits to contact a distal DNA site and RNA polymerase (activation) or separate DNA sites and the *GcvR* repressor (repression); from (Stauffer & Stauffer, 1998).

I.5. Regulation by RNA

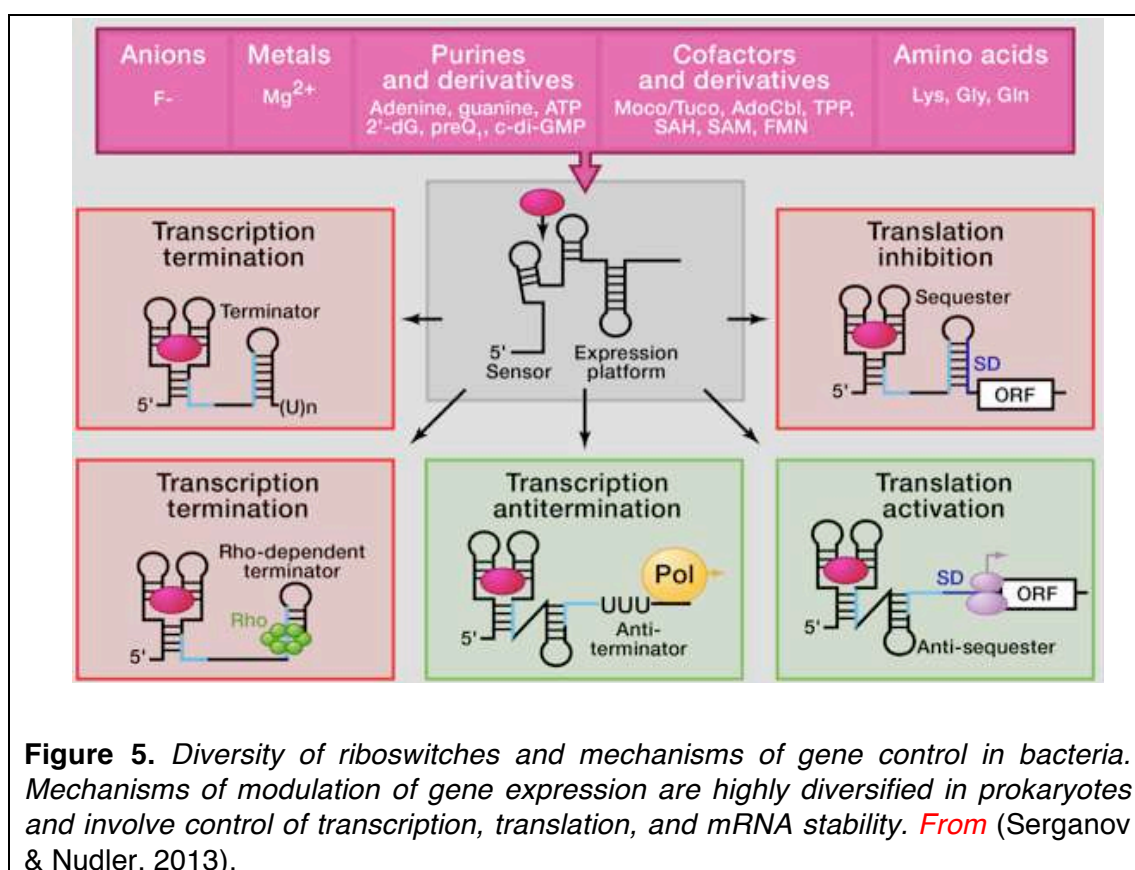
For a long time, the only function of RNA molecules in the cell was thought to be that of carrying the genetic information from gene to protein. In the recent years, experimental evidence has accumulated indicating the existence of different types of untranslated RNA molecules, which can act as regulators and affect a large variety of processes. These include sensing and responding to the availability of nutrients, interfering with bacteriophage infection, participating in responses to changes in environmental conditions and modifying bacterial virulence. Most RNA regulators function at the transcriptional and/or post-transcriptional level. In bacteria, three types of RNA regulators have been found.

I.5.1. Riboswitches

A riboswitch is a non-coding RNA element located in the 5' untranslated region (5'-UTR) of mRNA that regulates mRNA expression by directly sensing the presence of small metabolites. The mechanism of "riboswitching" was first discovered in 2002 and experimentally validated by a number of examples (Nahvi *et al.*, 2002, Mironov *et al.*, 2002, Winkler *et al.*, 2002b, Winkler *et al.*, 2002a). These data also confirmed the earlier idea that expression of some mRNAs might be directly related with concentration

changes in small metabolites (Gold *et al.*, 1997, Gelfand *et al.*, 1999, Miranda-Rios *et al.*, 2001).

Riboswitches comprise two domains: the aptamer (also called ligand-binding domains or sensor) and the expression platform. The sequence and structure of the aptamer component is highly conserved among different bacteria. Riboswitches can sense multifarious signals such as coenzymes (Nahvi *et al.*, 2002), metal ions (Cromie *et al.*, 2006) amino acids, nucleobases and their derivatives (Serganov & Nudler, 2013) including negatively charged fluoride anions (Baker *et al.*, 2012). After metabolite binding to the aptamer moiety, the conformation of the platform changes, leading to modulation of downstream events. Such events include transcription termination, antitermination, translation activation or activation (Fig. 5).



I.5.2. CRISPR systems

Clustered regularly interspaced short palindromic repeat (CRISPR) are sequences composed of direct repeats that are separated by similarly sized non-repetitive spacers. In bacteria the CRISPR RNAs assemble with CRISPR-associated proteins (Cas) to large surveillance complexes that interfere with foreign nucleic acids such as bacteriophage and plasmid DNA.

The first report of what is now known as a CRISPR was made by Ishino and his colleagues in 1987 (Ishino *et al.*, 1987). Further characterization of the locus revealed the existence of 14 repeats of 29 base pairs (bp) that were interspersed by 32–33 bp non-repeated spacer sequences (Nakata *et al.*, 1989). In 2005, G. Vergnaud and his coworkers identified the origin of the spacers in CRISPR elements in *Yersinia pestis* (Pourcel *et al.*, 2005). In recent years, computational analyses have revealed the presence of CRISPRs in almost 40% of the sequenced bacterial genomes (Kunin *et al.*, 2007).

The function of CRISPR systems is provide a defensive response against foreign nucleic acids in bacteria (Fig. 6). In CRISPR systems, foreign DNA is integrated into the CRISPR locus, and the host spacer sequence replaced by invading genetic elements (protospacers). Then the CRISPR loci are transcribed and yield a long primary transcript which is processed by a Cas or by RNase III family nucleases (Carte *et al.*, 2008, Haurwitz *et al.*, 2010, Deltcheva *et al.*, 2011, Sashital *et al.*, 2011) to form a library of short CRISPR-derived RNAs (crRNAs). Each crRNA contain a sequence complementary to a foreign invading nucleic acid (Fig. 6). After processing, the crRNA associates with one or more Cas proteins to form a surveillance complex that targets and destroys invading genetic material (Brouns *et al.*, 2008, Hale *et al.*, 2009, Garneau *et al.*, 2010, Wiedenheft *et al.*, 2011a, Wiedenheft *et al.*, 2011b).

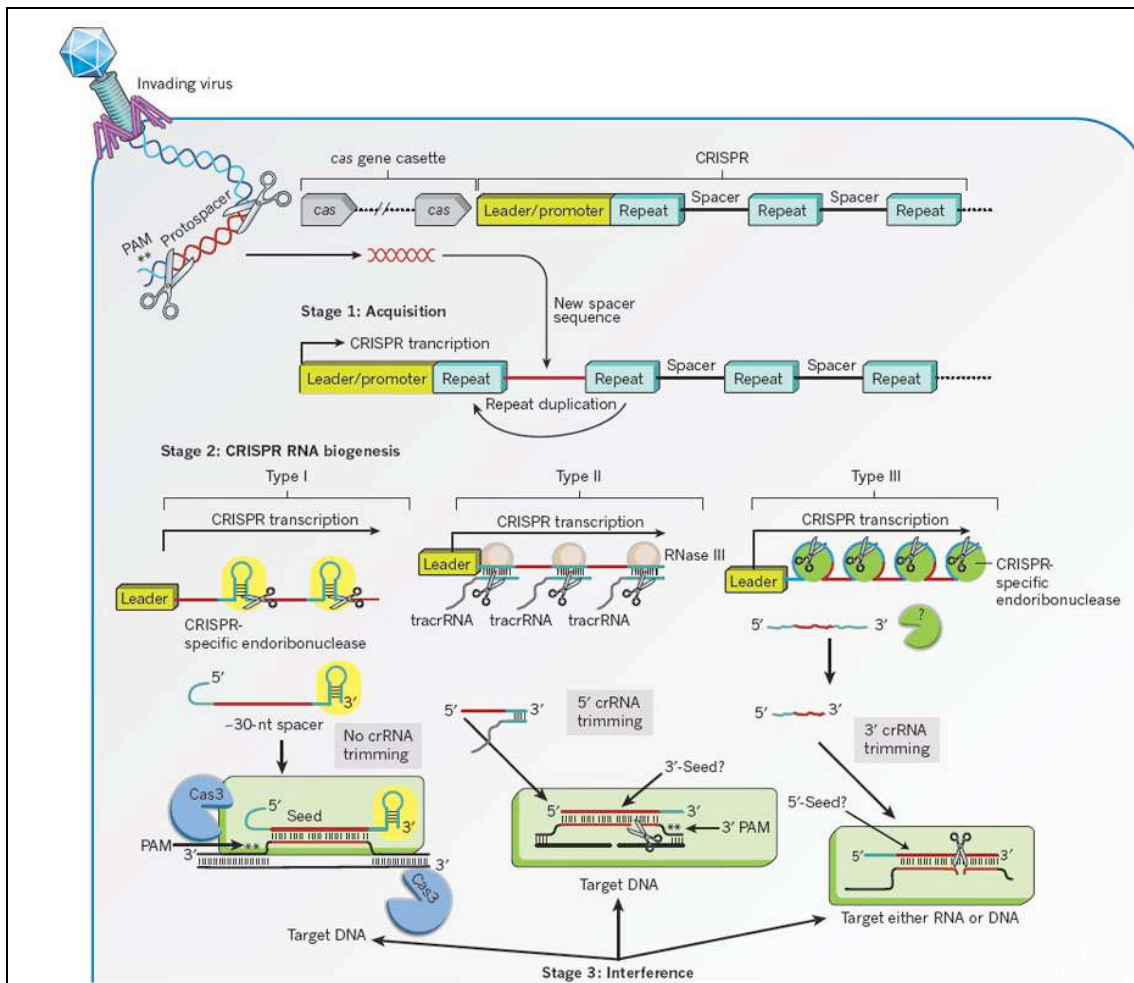


Figure 6. A diverse set of CRISPR-associated (*cas*) genes (grey arrows) encode proteins required for new spacer sequence acquisition (Stage 1), CRISPR RNA biogenesis (Stage 2) and target interference (Stage 3). Each CRISPR locus consists of a series of direct repeats separated by unique spacer sequences acquired from invading genetic elements (protospacers). Protospacers are flanked by a short motif called the protospacer adjacent motif (PAM) that is located on the 5' (type I) or 3' (type II) side in foreign DNA. Long CRISPR transcripts are processed into short crRNAs by distinct mechanisms. In type I and III systems, a CRISPR-specific endoribonuclease (yellow ovals and green circles, respectively) cleaves 8 nucleotides upstream of each spacer sequence. In type III systems, the repeat sequence on the 3' end of the crRNA is trimmed by an unknown mechanism (green pacman, right). In type II systems, a trans-acting antisense RNA (*tracrRNA*) with complementarity to the CRISPR RNA repeat sequence forms an RNA duplex that is recognized and cleaved by cellular RNase III (brown ovals). This cleavage intermediate is further processed at the 5' end resulting in a mature, approximately 40-nucleotide crRNA with an approximately 20-nucleotide 3'-handle. In each system, the mature crRNA associates with one or more Cas proteins to form a surveillance complex (green rectangles). Type I systems encode a Cas3 nuclease (blue pacman), which may be recruited to the surveillance complex following target binding. A short high-affinity binding site called a seed-sequence has been identified in some type I systems^{27, 60}, and genetic experiments suggest that type II systems have a seed sequence located at the 3' end of the crRNA spacer sequence. From: (Wiedenheft et al., 2012).

I.5.3. Small regulatory RNAs

Small RNAs (sRNAs) comprise an important class of regulators in all kingdoms of life. They participate in the cell responses to stress, changes in environmental conditions and nutrient availability. In doing so, sRNAs play a key role in cell adaptation (Repoila *et al.*, 2003).

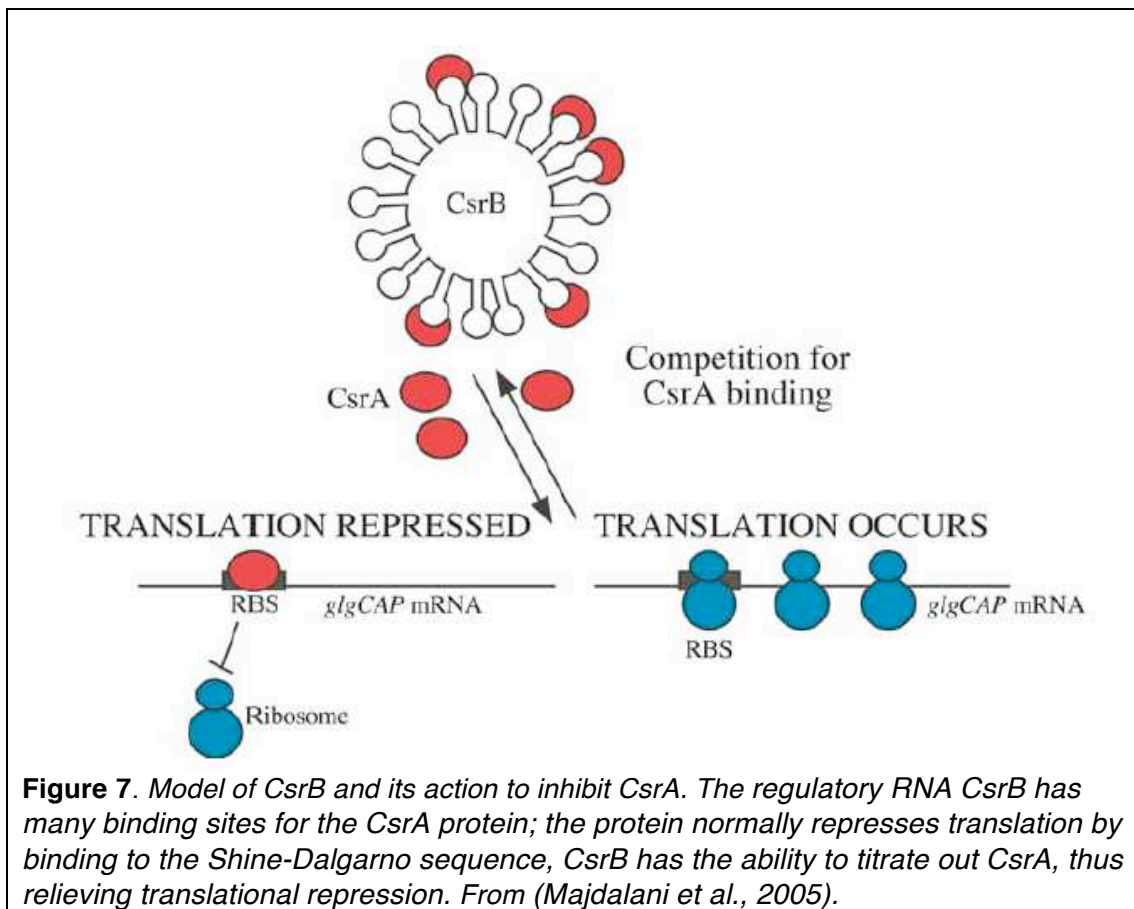
The first sRNA capable of regulating gene expression was discovered in the 1980s (Mizuno *et al.*, 1984). For some time this sRNA, MicF, remained an isolated example. But starting in the early 2000s, a number of different *in silico* and experimental strategies converged to uncover more than 70 sRNAs by 2005 (Vogel, 2005 #956}. In 2011 the number of sRNA has reached to 100 (Gottesman & Storz, 2011a). Based on the size of enterobacterial genomes, it has been estimated that the number of sRNA genes may be in the range of 200–300 (Vogel & Wagner, 2007). There appears to be a core of conserved sRNAs in different related pathogens, such as *E.coli*, *S.Typhimurium* and *S. flexneri* (Hershberg *et al.*, 2003).

Regulatory sRNAs can be divided into two classes depending on whether they target a protein or an RNA.

I.5.3.1. Protein-targeting sRNAs

They can be further divided into two groups. In the first group, the sRNA itself has the essential biological activity or contributes functions to a protein. This is the case, for example, of M1, the RNA component of Ribonuclease P, (RNase P) (Esakova & Krasilnikov, 2010) and of the 4.5S RNA component of the signal recognition particle (Ribes *et al.*, 1990). In contrast, sRNAs of the second group can regulate the activities of their target proteins by mimicking the structures of other nucleic acids. For example, CsrB targets CsrA (Carbon Storage Regulator) protein, a global regulator of carbon metabolism which represses translation of a number of mRNAs by binding to their Shine-Dalgarno sequences (Romeo, 1998). CsrB sRNA contains multiple copies of a sequence mimicking the CsrA binding site. When the level of CsrB

is high, this sRNA can compete out CsrA from target mRNA, thus relieving the translational repression (Fig. 7)



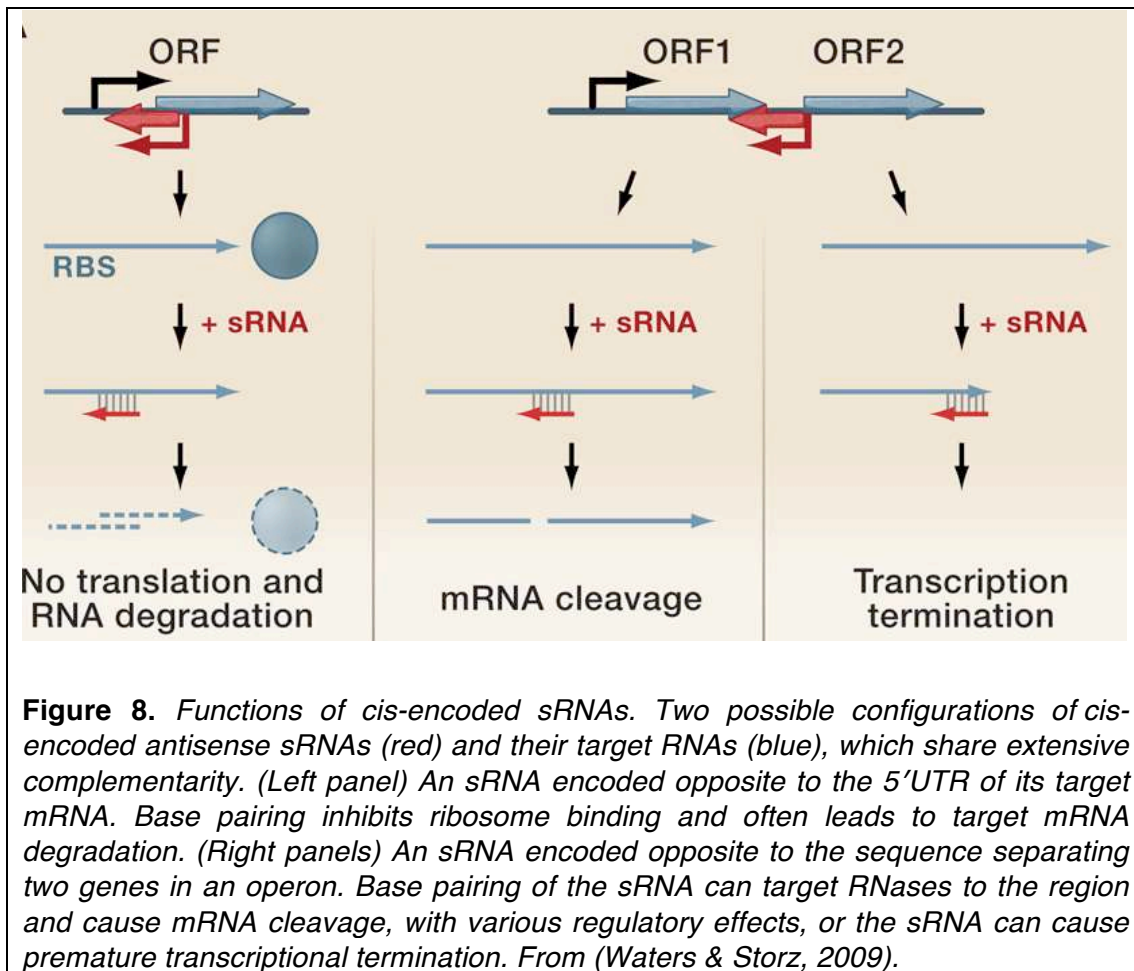
I.5.3.2. RNA-targeting sRNAs

Undoubtedly, the largest and most extensively studied class of sRNAs acts through base pairing with complementary sequence in target messenger RNAs (mRNAs), affecting translation and/or mRNA stability (Masse *et al.*, 2003, Caron *et al.*, 2010). RNA-targeting sRNAs can be divided into two groups depending on the positions of their genes relative to the target genes.

I.5.3.2a. Cis-encoded sRNAs

These sRNAs that are encoded at the same locus of the target RNA but in opposite strand and, as a result, share extended regions of sequence complementarity with their target. Cis-encoded sRNAs have been mostly found in plasmids, phages and transposons (Brantl, 2002) and only a few in

the bacterial chromosome (Kawano *et al.*, 2005). They can function by several mechanisms depending of the position of the pairing region. They can inhibit translation initiation, promote transcription termination or induce RNA degradation (Fig. 8) (Brantl, 2002, Waters & Storz, 2009).



I.5.3.2b. Trans-encoded sRNAs

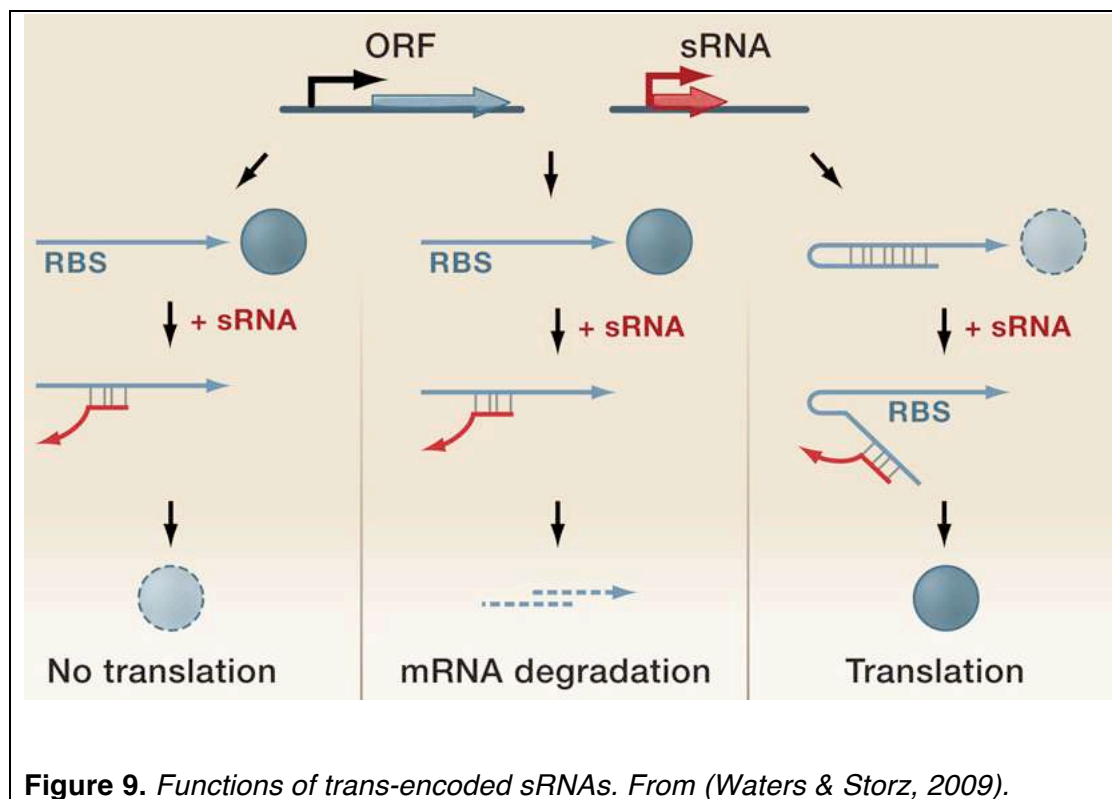
The largest class of regulatory sRNAs comprises molecules that are encoded at separate location from their target genes. They are characterized by a limited and often imperfect sequence complementarity with their mRNA targets and by the fact that they often have multiple targets (Gottesman, 2005, Prevost *et al.*, 2007). Pairing normally involves at least 6–8 contiguous nucleotides, which in multi-target sRNAs are generally located at the same position of the molecule (Balbontin *et al.*, 2010, Bouvier *et al.*, 2008). On the mRNA side, the pairing residues are usually part of the 5' UTR or of the initial

portion of the coding sequence. An additional relevant feature of trans-encoded sRNA is their requirement for RNA chaperon protein Hfq for stability and function (Valentin-Hansen *et al.*, 2004, Vogel & Luisi, 2011). The role of Hfq in sRNA function is described in more detail in section 1.5.3.3.

Trans-encoded sRNAs can act as either positive or negative regulators (Fig. 9). In the case of positive regulation, the most common mechanism involves the sRNA activating translation by preventing the formation of the inhibitory secondary structure in the mRNA (Majdalani *et al.*, 2005, Prevost *et al.*, 2007). Annealing of the sRNA to the mRNA changes the mRNA structure, freeing the ribosome binding site (Shine-Dalgarno sequence) so that the ribosome can access the translation initiation region (Hammer & Bassler, 2007, Urban & Vogel, 2008). It was proposed that an additional mechanism by which sRNA can activate gene expression is by increasing the stability of the mRNA (Opdyke *et al.*, 2004, Ramirez-Pena *et al.*, 2010). This mechanism remained hypothetical until recently, when the group of Jörg Vogel showed that the small RNA SgrS activates the synthesis of YigL (a phosphatase involved in the detoxification of phosphosugars) by stabilizing a decay intermediate of the bicistronic *pldB-yigL* mRNA (Papenfert *et al.*, 2013). Still, some aspects of this mechanism remain elusive; in particular the surprising finding that RNase E, an enzyme that normally degrades mRNAs, is required for the activation step (Papenfert *et al.*, 2013).

Negative regulation is by far the most outcome of sRNA activity. By base pairing with sequences in the 5'UTR of the mRNA, the sRNA occludes the ribosome binding site, thus inhibiting translation initiation (Fig. 9). No longer shielded by translating ribosomes, the mRNA becomes exposed to degradation by RNase E (Masse *et al.*, 2003, Morita *et al.*, 2005). Furthermore, in some cases, mRNA degradation is not just the consequence of translation inhibition but the formation of the sRNA:mRNA duplex directly recruits RNase E (Pfeiffer *et al.*, 2009). Interestingly, there are some sRNAs that inhibit translation through base pairing far upstream of the initiating AUG, for example: GcvB and RyhB and RybB (Sharma *et al.*, 2007, Vecerek *et al.*, 2007, Balbontin *et al.*, 2010). It was proposed that sRNA can act at such

distant sites by sequestering a ribosome standby site (Darfeuille *et al.*, 2007), or through direct RNase E recruitment (Wagner, 2009).



I.5.3.3. Hfq: mediator of sRNA activity

In enteric bacteria, Hfq is a protein that participates in mRNA expression and stability and, as a chaperone, in sRNA-mediated regulation. Hfq was first identified as a host factor for RNA phage Q β replication *in vitro* (also known as host factor 1) (Franze de Fernandez *et al.*, 1968, Franze de Fernandez *et al.*, 1972, Miranda *et al.*, 1997, Su *et al.*, 1997). Phylogenetic analyses show that the *hfq* gene (or a related gene) is present in approximately half of all sequenced Gram-positive and Gram-negative genomes, with some bacteria harboring more than one *hfq*-like gene (Sun *et al.*, 2002, Valentin-Hansen *et al.*, 2004). Expression of these genes and functionality of their products remains to be demonstrated in most cases. Proteins of the Hfq family range from 70 to 110 amino acids and are generally organized in a homohexameric structure (Franze de Fernandez *et al.*, 1972, Moller *et al.*, 2002, Zhang *et al.*, 2002). *E. coli* contains between 50,000 to 60,000 Hfq monomers (about 10,000 hexamers), the majority of which (80%

to 90%) are found in the cytoplasmic fraction in association with ribosomes (Vasil'eva lu & Garber, 2002). A significant amount is also located in the close proximity to the cytoplasmic membrane (Diestra *et al.*, 2009, Vassilieva *et al.*, 2002).

A main role of Hfq is to associate with trans-encoded sRNA and participate in the mechanism of post-transcriptional regulation. Thus, the protein is a direct regulator of genes under sRNA control, which include genes involved in stress responses (Muffler *et al.*, 1997), virulence (Sittka *et al.*, 2007, Ding *et al.*, 2004) and quorum-sensing regulation (Meibom *et al.*, 2009, Lenz *et al.*, 2004). Furthermore Hfq plays a key role in the cellular response to phosphosugar toxicity and low iron levels (Fantappie *et al.*, 2009, Gorke & Vogel, 2008, Vanderpool, 2007, Masse & Gottesman, 2002). Hfq also functions in cell envelope homeostasis through its role in the activity of MicA and RybB. These two sRNAs coordinately downregulate the expression of outer membrane proteins in stationary phase and under stress conditions (Bossi *et al.*, 2008, Figueroa-Bossi *et al.*, 2006a, Papenfort *et al.*, 2006, Rasmussen *et al.*, 2005).

1.5.3.3a. Hfq structure and RNA binding patterns

Hfq is composed of 6 identical 11.2 kDa subunits. The stable, ring-shaped structure of the hexamer was initially visualized by electron microscopy (Zhang *et al.*, 2002). and subsequently resolved by x-ray crystallography, with or without synthetic RNA bound to it (Schumacher *et al.*, 2002). The ring, about 70 Å in diameter, has a positively charged central pore on one of its two faces (Fig. 10). Studies with synthetic oligomers suggest that the two surfaces of the protein have different binding specificities: sequences rich in U residues bind preferentially to the proximal face, whereas A-containing sequences bind to the distal face (Link *et al.*, 2009, Mikulecky *et al.*, 2004). It should be mentioned that even though the global architecture of the hexamer is conserved among species, some differences exist, in particular in the charge distribution. For instance, *E. coli* Hfq has a positive electrostatic surface for the trough that connects the proximal and distal faces,

which is in contrast to the same area on the *S. aureus* Hfq, which shows a negative electrostatic surface (Fig. 10).

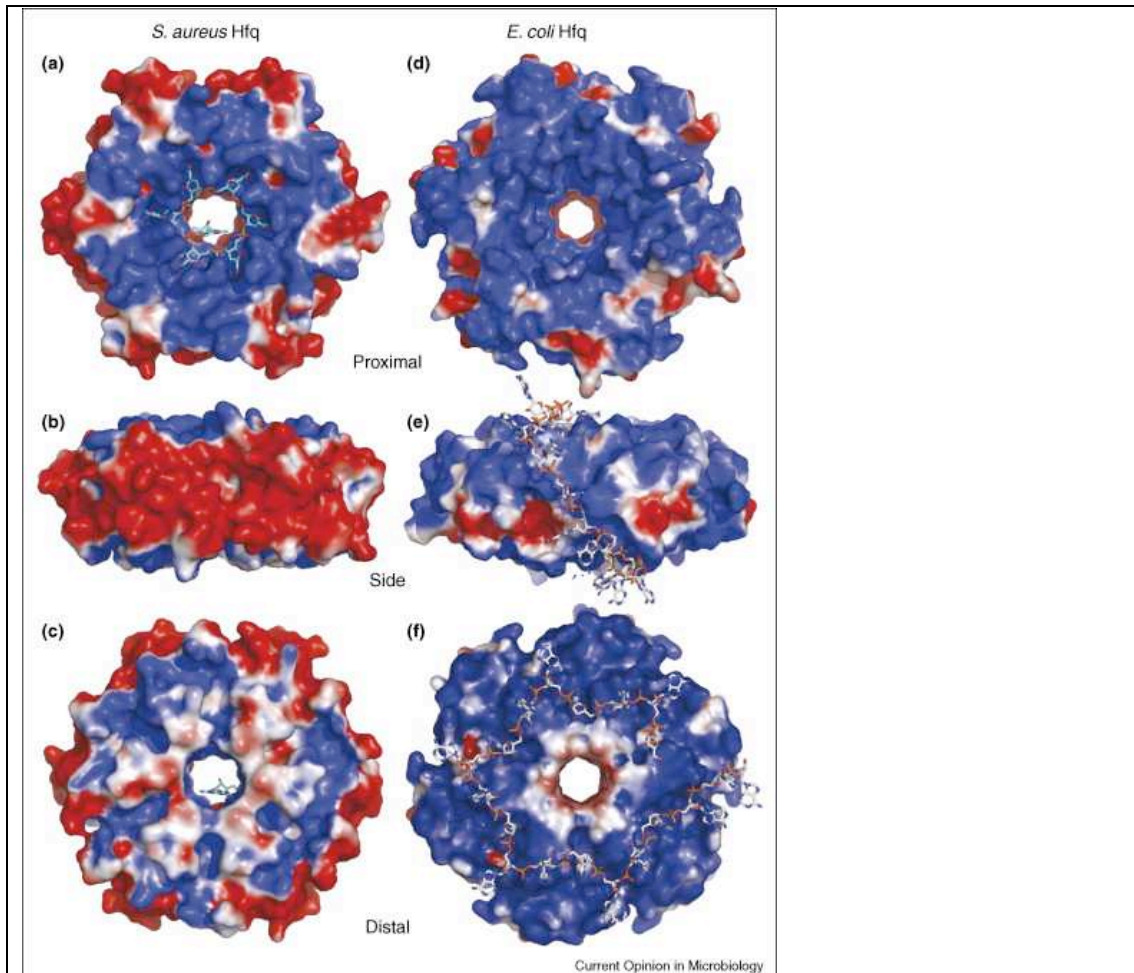


Figure 10. Electrostatic potential energy surfaces of the known and proposed RNA binding sites of the *S. aureus* and *E. coli* Hfq proteins. (a–c) Views of the proximal side and distal electrostatic surfaces of *S. aureus* Hfq, respectively. (d–f) Views of the proximal, side and distal electrostatic surfaces of *E. coli* Hfq, respectively (blue is electropositive and red is electronegative). The side view of *E. coli* Hfq includes a plausible RNA binding cleft that would enable A27 to bind to both the Proximal and Distal Sites. The view of the Distal Site of the *E. coli* Hfq shows a possible binding site for A18 (i.e. a poly(A) tail). The bound RNAs are shown as solid sticks with carbon, nitrogen, oxygen and phosphorous atoms colored white, blue, red and orange, respectively. The electrostatic potential energy surfaces were created by PyMol and the APBS plug-in (Delano Scientific LLC, Palo Alto, CA). From (Brennan & Link, 2007).

The overall architecture of Hfq is strongly reminiscent of that of Sm proteins that participate in many different RNA-processing reactions in eukaryotes (Schumacher et al., 2002, Arluison *et al.*, 2002, Achsel *et al.*,

1999, Salgado-Garrido *et al.*, 1999, Tharun *et al.*, 2000). Further *in vitro* experiments show that besides poly(A) tails, Hfq can bind poly(A-R-N) triplets, such as (AGG)₈, (AGC)₈, and the shorter (A-R-N)₄ AACACAAGAAG (Link *et al.*, 2009). Binding takes place on the distal face of Hfq (Sun & Wartell, 2006, Mikulecky *et al.*, 2004, Brennan & Link, 2007). In contrast, the proximal face of the protein is thought to be involved in the binding of sRNAs. This occurs at the level of the U-rich 3' ends resulting from Rho-independent transcription termination in sRNA genes (Sauer & Weichenrieder, 2011, Otaka *et al.*, 2011). Work from our laboratory showed that Hfq has a strong binding affinity for the sequence AAUAA found in regulatory sRNA ChiX (Figueroa-Bossi *et al.*, 2009) in the 5' UTR of the *ompC* mRNA (Balbontin *et al.*, 2010) and in the *yifK* mRNA studied here (see Results section). Furthermore Balbontin and coworkers showed that Hfq can bind to a short stretch of Us located internally in the RybB sRNA (Balbontin *et al.*, 2010).

I.5.3.3b. Mechanism of action of Hfq

The main role of Hfq in sRNA regulation is to stimulate the base-pair interaction between sRNAs and mRNAs (Fig. 11). Hfq was shown to increase the rate of sRNA-mRNA binding (Schumacher *et al.*, 2002, Kawamoto *et al.*, 2006) and to remodel RNA secondary structures (Geissmann & Touati, 2004, Moller *et al.*, 2002). For example Hfq binding to the 5' end of *rpoS* mRNA - to the sequence (AAN)₄ - was shown to facilitate pairing between *rpoS* mRNA and DsrA sRNA (Soper & Woodson, 2008). The (AAN)₄ sequence is also required *in vivo* for Hfq-dependent regulation of *rpoS* translation by *rpoS* activating sRNAs in *E. coli* (Soper *et al.*, 2010).

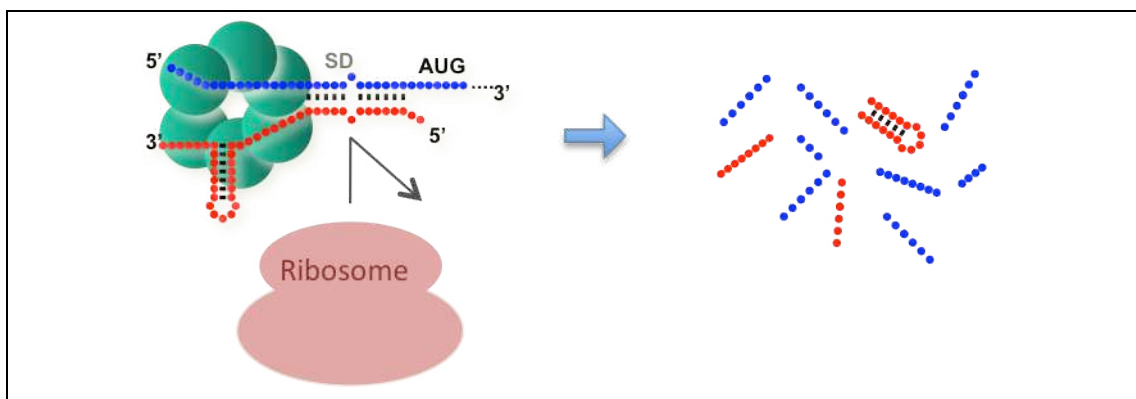


Figure 11. Model for Hfq action. Hfq protect sRNA against degradation, stimulates sRNA:mRNA pairing and in some systems stimulates the degradation of the sRNA:mRNA complex (see text for details).

In vivo, the loss of Hfq leads to a rapid degradation of the majority of trans-encoded sRNA, indicating that Hfq binding protects these sRNA against the action of RNases, mostly RNaseE (Zhang *et al.*, 2003, Gottesman, 2004). This effect is opposite to that exerted on most sRNA-regulated mRNA, which are destabilized upon Hfq binding. A study by the Aiba group showed that the Hfq/sRNA complex can recruit RNase E and stimulate mRNA cleavage following sRNA pairing (Morita *et al.*, 2005). Furthermore, Hfq can affect mRNA stability directly by stimulating poly(A) adenylation (Fig. 12) (Hajnsdorf & Regnier, 2000, Mohanty *et al.*, 2004) and by binding to poly(A) tails (Folichon *et al.*, 2003). These latter activities are somewhat contradictory since stimulation of poly(A) adenylation is expected to stimulate mRNA degradation whereas poly(A) binding was shown to prevent degradation (Folichon *et al.*, 2003). To explain the paradox, Folichon and coworkers proposed that Hfq prevents poly(A)-dependent degradation of mRNAs whose stability is mostly controlled by RNase E and favors poly(A)-dependent degradation of structured mRNA lacking RNase E sites (Folichon *et al.*, 2003).

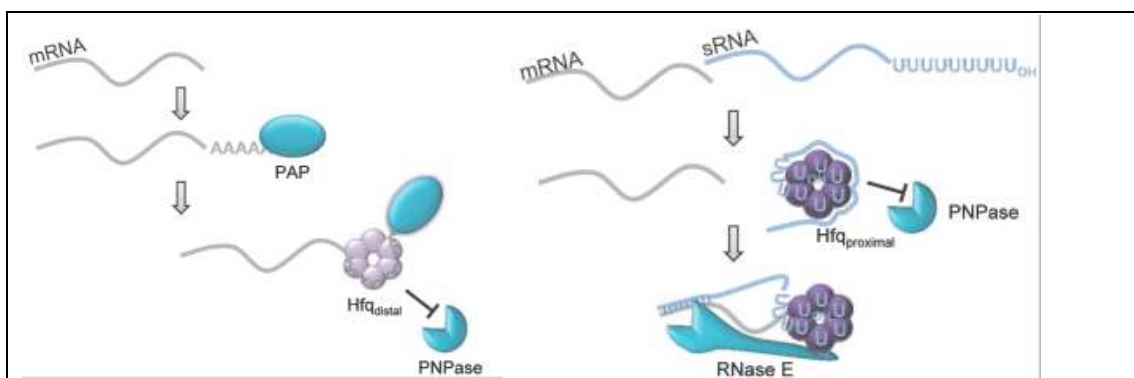


Figure 12. Model for Hfq bind 3' end and activity RNA decay. From (Wilusz & Wilusz, 2013).

I.5.3.4 Global sRNA-dependent regulatory networks: selected examples

I.5.3.4a. RybB and iron homeostasis

Thanks to the ability to regulate multiple targets, trans-encoded sRNAs can mediate global regulatory responses. A paradigm for this type of response is provided by the RyhB sRNA, which controls iron homeostasis in *E. coli* and *Salmonella*. Iron is one of most important metal ions for bacteria, as it is used as a cofactor in many enzymes, such as those involved in the tricarboxylic acid (TCA) cycle, respiration and DNA synthesis. However, iron is also potentially toxic since it can react with oxygen to form free radicals that can damage cell components. Therefore, the concentration of free ferrous ion in all cells is tightly regulated. In *E. coli* and other bacteria, this regulation involves two factors: the Fur protein and RyhB sRNA. Fur is an iron-binding protein, that when bound to iron, represses transcription of genes involved in iron uptake (Hantke, 2001). Among the genes repressed by Fur is the *ryhB* gene (Masse & Gottesman, 2002). Like the other members of the Fur regulon, the *ryhB* gene is activated (derepressed) when iron becomes limiting. Upon accumulating, RybB downregulates the synthesis of a number of iron-utilization and iron-storage proteins by pairing with sequences near the translation initiation sites of their mRNAs and inducing their degradation (Fig. 13) (Masse & Gottesman, 2002). In doing so, the sRNA limits iron consumption when the metal is scarce. As iron levels return to normal, *ryhB* repression by Fur is restored. Therefore, indirectly Fur upregulates all of the RybB-repressed genes, as many as 18 of which have been identified in *E. coli* (Masse *et al.*, 2005).

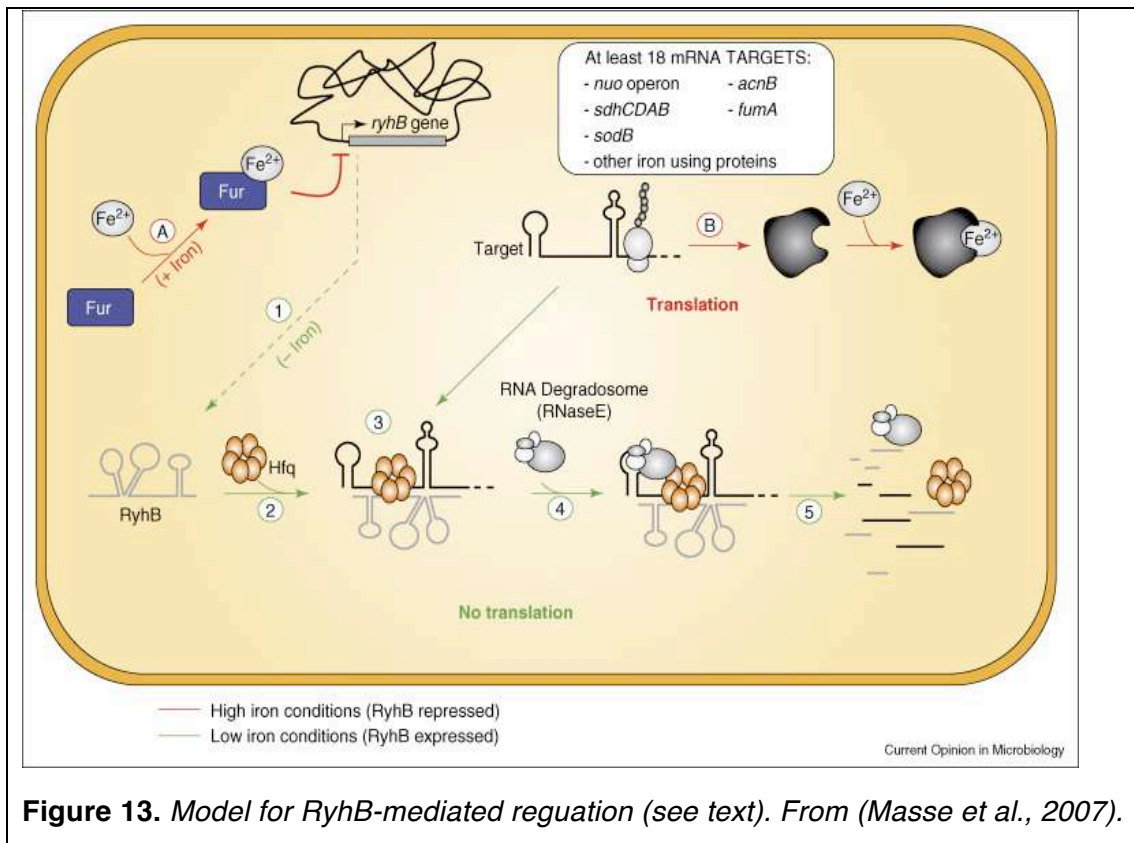


Figure 13. Model for RyhB-mediated regulation (see text). From (Masse *et al.*, 2007).

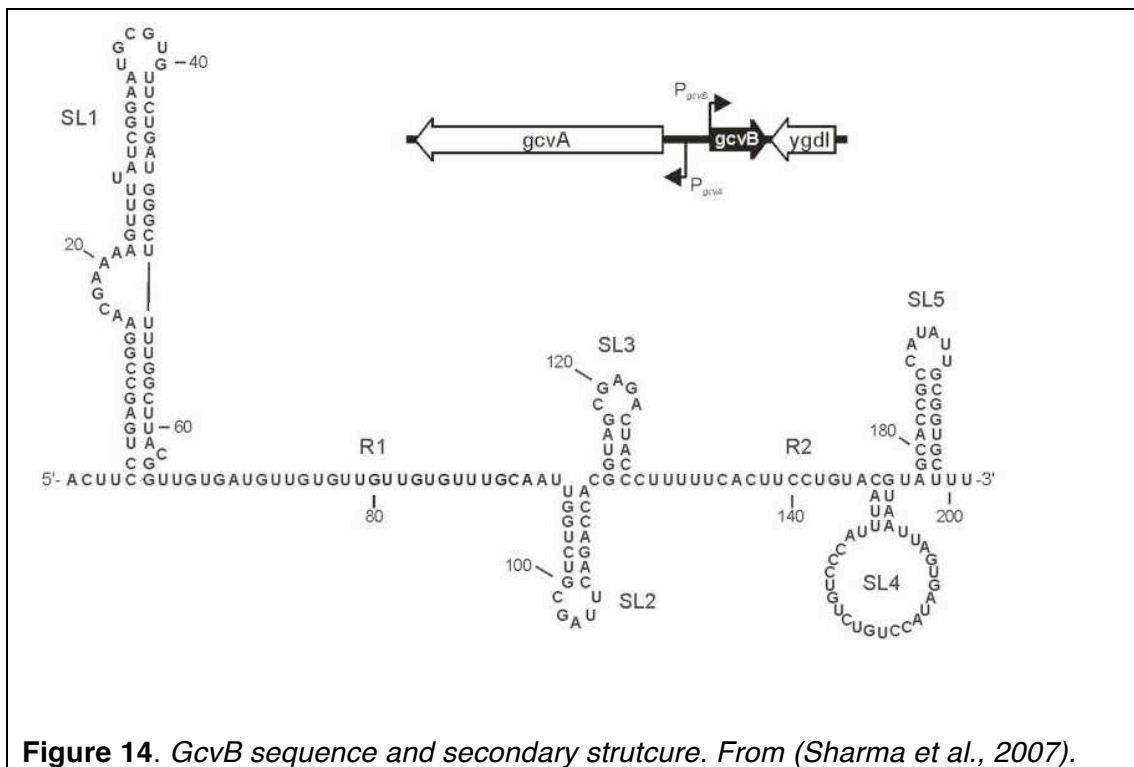
I.5.3.4b. GcvB and amino acid management

GcvB is a 200 nucleotide sRNA that downregulates expression of a large number of genes encoding the transporters for amino acids, dipeptides and oligopeptides (Urbanowski *et al.*, 2000, Sharma *et al.*, 2007, Pulvermacher *et al.*, 2009a, Pulvermacher *et al.*, 2008, Sharma *et al.*, 2011). In addition, GcvB downregulates some amino acid biosynthetic genes (Sharma *et al.*, 2011, Sharma *et al.*, 2007). Although the physiological role of this regulation is not fully understood, it appears to be an important part of amino acid metabolism during growth in nutrient-rich environments (Sharma *et al.*, 2011)

GcvB was discovered fortuitously during the study of the promoter region for the *gcvA* gene, which encodes the main transcriptional regulator of the glycine cleavage system (see section II.3.2.) This work revealed the presence of a small gene divergently transcribed from *gcvA* with its promoter overlapping the *gcvA* promoter (Urbanowski *et al.*, 2000). Further analysis showed that GcvA activates *gcvB* transcription by binding to a region between

positions -29 and -76 relative to the transcription start site (Urbanowski et al., 2000).

The sequence of GcvB is highly conserved in gram-negative bacteria, particularly in the region involved in base-pairing with most target mRNAs. This region, called the R1 region, is characterized by the presence of a long stretch of alternating G and U residues (Fig.14).



As a result, most GcvB regulated mRNAs contain CA-rich sequences at variable distances from the translation initiation sites. In some targets, these elements lie more than 40 nucleotides upstream from the initiator AUG (Sharma et al., 2007). Two exceptions to the R1 pairing rule have been reported. One exception is in the downregulation of the *phoP* mRNA (encoding the response regulator of the PhoP/PhoQ two-components system) by GcvB in *E.coli*. Pairing was shown to involve sequences in stem-loop 4 (see Fig. 14)(Coornaert et al., 2013). The second exception is the *cycA* mRNA (encoding a permease for various amino acids) which pairs with GcvB region R2 (Pulvermacher et al., 2009b). Recently GcvB was also found to negatively control the leucine response regulator Lrp (Sharma et al., 2011). In this case, more than one region of GcvB appears to participate in the interaction with *lrp*

mRNA (Sharma et al., 2011). Overall, the number of GcvB targets is thought to be greater than 40, making the GcvB network the largest of its kind (Fig. 15).

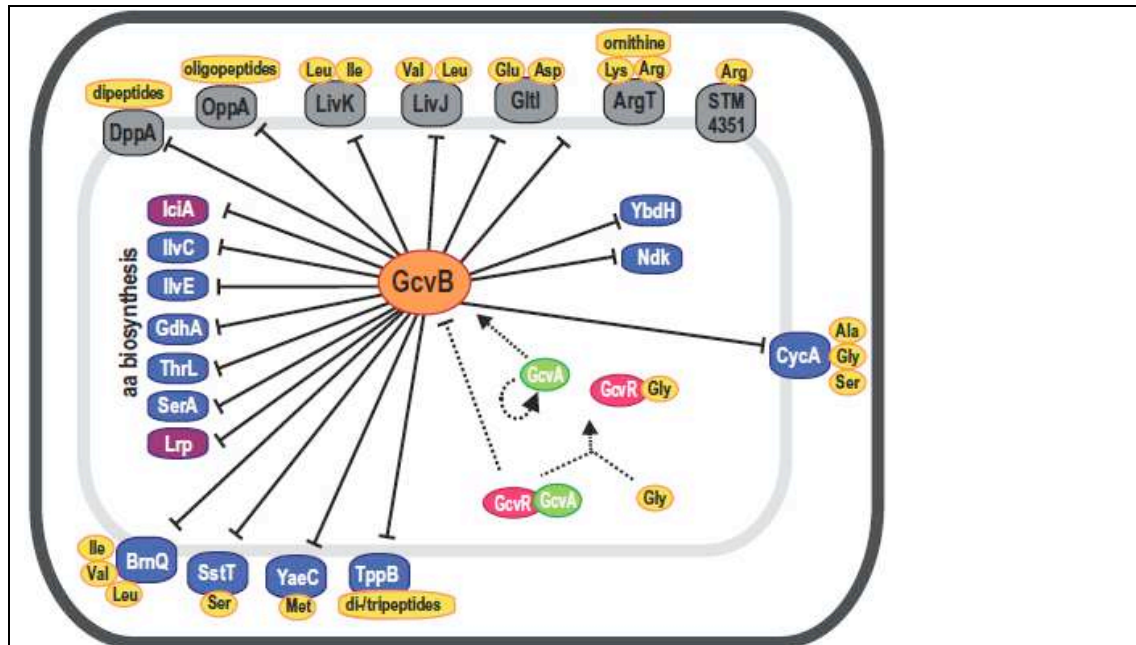


Figure 15. Schematic representation of the GcvB network. GcvB sRNA represses mRNAs of many ABC transporters involved in amino acid uptake and of amino acid biosynthetic enzymes. From : (Sharma et al., 2011).

Several lines of evidence strongly suggest that GcvB is an Hfq-dependent sRNA. GcvB coimmunoprecipitates with Hfq in extracts of *E. coli* (Zhang et al., 2003) and is unstable in Δhfq strains of *E. coli* (Urban & Vogel, 2007) and *Salmonella* (Sharma et al., 2007). In *E. coli*, the GcvB requirement for Hfq has been demonstrated in the regulation of *dppA*, *oppA* (Pulvermacher et al., 2009a, Urban & Vogel, 2007), *sstT* (Pulvermacher et al., 2009c) and *cycA* (Pulvermacher et al., 2009b). For *dppA*, this requirement has also been confirmed in *S. Typhimurium* (Sharma et al., 2007). It was demonstrated that Hfq confers increased the stability of the sRNA and facilitated pairing between GcvB and target mRNA (Pulvermacher et al., 2009a).

II. RESULTS

II.1. Characterization of *yifK*, a Hfq-regulated locus in *Salmonella*

Some years ago, our laboratory developed a genetic screen for identifying genes regulated by sRNA chaperon protein Hfq in *Salmonella* (Figueroa-Bossi et al., 2006a). This approach was aimed at identifying novel genes regulated by Hfq-dependent small RNAs. The strategy employed a transposable element derived of phage Mu, MudK, which carries a *lac* operon with a promoterless *lacZ* gene which also lacks the initiating AUG codon. MudK can give rise to translational fusion upon inserting randomly in the bacterial chromosome (Hughes & Roth, 1988). Presence of a kanamycin-resistance (Kan^R) marker within the transposon allows to positively select for the transposition events. The transposition experiment was performed in a strain with the normal copy of the *hfq* gene deleted and a second copy of *hfq* fused to the promoter of the arabinose operon. Therefore this strain is a conditional *hfq* mutant: Hfq⁻ in the absence of arabinose and Hfq⁺ in the presence of arabinose. Colonies selected on LB plates supplemented with kanamycin were replica-plated on lactose indicator plates (either MacConkey-lactose or LB-Xgal) that contained or lacked arabinose. Colonies whose color differed in the presence or absence of arabinose were characterized further. This study led to the identification of 19 *lac* fusions showing increased expression in the absence of arabinose and 4 fusions expressed less efficiently under these conditions (Figueroa-Bossi et al., 2006a). The members of the first group were candidates for loci downregulated by Hfq-dependent sRNAs, whereas the second group might contain loci requiring sRNAs for activation. In fact, three of the four members of the second group are genes transcribed by alternative sigma factor σ^S , previously known to be activated by sRNAs (Repoila et al., 2003).

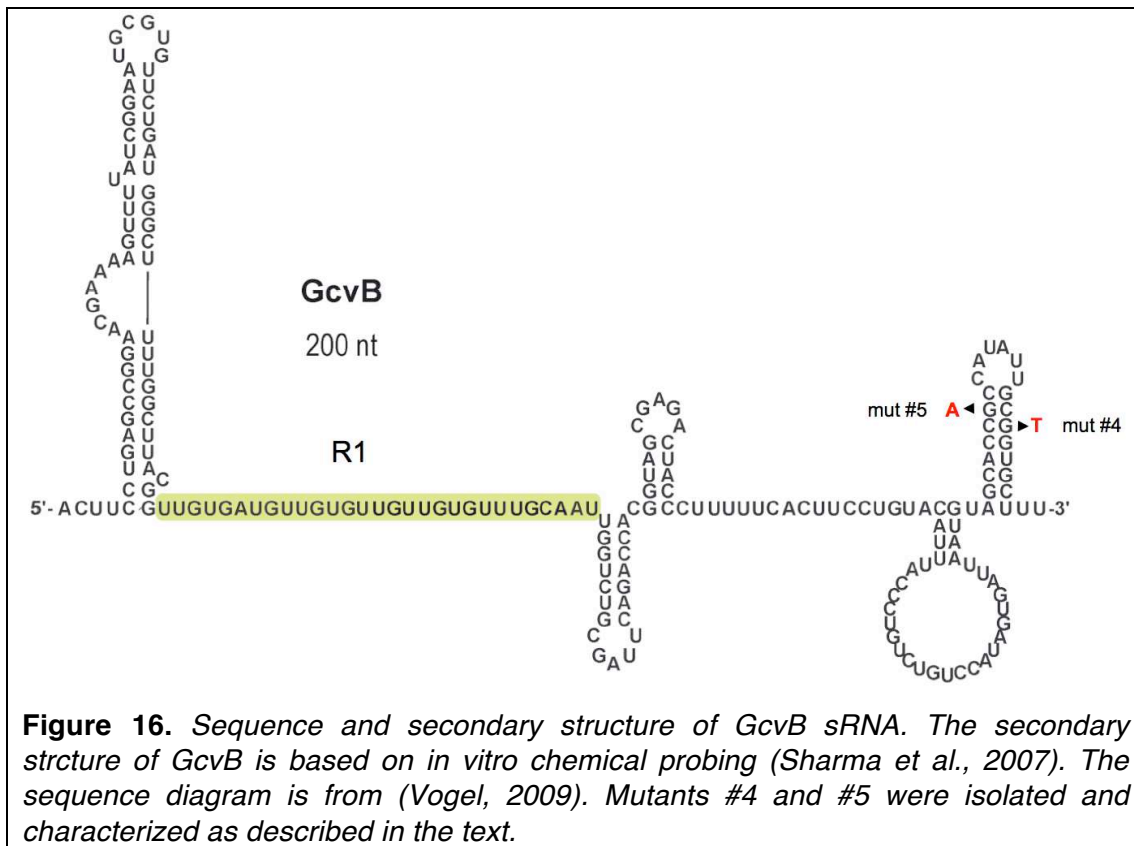
When I joined the laboratory to prepare my thesis, the involvement of small RNAs in the regulation of most of the loci identified above was not yet known. I was assigned the task of determining the sRNA involvement in two such loci. The first locus was *yifK*, which codes for a highly conserved protein annotated as “putative amino acid transporter” in DNA sequence databases.

The second locus was *aphA*, coding for a periplasmic acid phosphatase (Grose *et al.*, 2005). Encouraging results originating from the *yifK* project during the initial phase of my thesis pushed me to concentrate more of this part, which became my official thesis project. The preliminary data obtained in the *aphA* project are described in Appendix #1.

Two MudK insertions generating in-frame *lacZ* fusions to the *yifK* gene had been isolated in the laboratory, one at position 143, the other at position 284 of the coding sequence. Both fusions are activated 5 to 7-fold in a strain deleted for *hfq*. The fusion at position 143 expresses higher β -galactosidase activity and was chosen in most of the experiments described in this thesis. For simplicity, I will often call this fusion *yifK-lacZY*, although one should remember that there is also Kan^R cassette downstream from it. It should be noticed that the presence of the entire *lac* operon in the MudK element allows one to follow the entire lactose utilization pathway – not just the *lacZ* gene product. This will be important in many experiments described here.

II.1.1. *yifK* is repressed by GcvB

In *Salmonella* and *E. coli*, the small RNA GcvB represses translation of several mRNAs encoding the transporters for amino acids and peptides (see section I.5.3.4b). Since *yifK* belongs to this class (see above), GcvB appeared a likely candidate for a *yifK* regulator. To test this hypothesis, I constructed a deletion of the *gcvB* gene by the technique of “Lambda Red” recombination (Datsenko & Wanner, 2000).



I found the *gcvB* deletion to cause a 3-to 5-fold increase in the expression of the *yifK-lacZY* fusion (see section II.2 below). Although this increase was less than that observed with the *hfq* deletion (see above), these findings confirmed that GcvB downregulates *yifK*. The greater effect of the *hfq* deletion might indicate that additional sRNAs participate in the regulation or the Hfq acts directly in the repression mechanism (section II.2). Presence of a sequence segment complementary to GcvB in the region immediately upstream of *yifK* Shine-Dalgarno motif suggested that GcvB represses *yifK* translation by base-pairing. GcvB uses a GU-rich portion of its sequence – called the R1 region (Fig. 16) – for pairing with most of its targets. This also appears to be the case with *yifK*, even though the portion of the R1 region involved in *yifK* pairing differs from that used with other targets (see below) .

In the initial part of my project, we sought to obtain evidence for the GcvB:*yifK* mRNA interaction through mutational analysis. In a first approach, I randomly mutagenized the *gcvB* gene and searched for mutants expressing *yifK-lacZY* at higher levels than the wild-type strain. I followed a procedure

that combines mutagenic PCR and Lambda red recombination. Performing the mutant selection on MacConkey-lactose indicator plates allowed the identification of isolates with higher β -galactosidase activity (red colonies in Fig. 17).

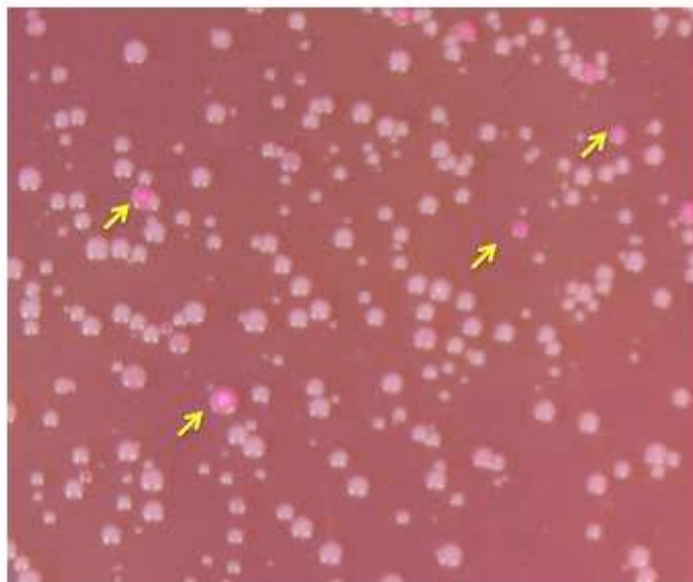


Figure 17. PCR mutagenesis of the *gcvB* gene. A DNA fragment spanning the *gcvB* gene and an adjacent *cat* cassette was amplified by PCR under error-prone conditions (see Materials and Methods) and introduced into a strain carrying a *yifK-lacZY* fusion and lambda red plasmid *pKD46* (Datsenko & Wanner, 2000). Recombinants were selected on MacConkey-lactose indicator plates supplemented with chloramphenicol. B. Red colonies (indicated by arrows) express *yifK-lacZY* at higher level.

Analysis of a number of these mutants showed that most of them carried mutations in the promoter of *gcvB*, or in the promoter of *gcvA* (the gene adjacent to *gcvB* on the 5' side), or within the portion of the *gcvA* gene included in the mutagenized fragment. The *gcvA* gene encodes the activator of *gcvB* transcription (Sharma et al., 2007, Urbanowski et al., 2000). Thus, these results suggested that derepression of *yifK-lacZY* in these mutants was the consequence of a decrease in GcvB concentration due to reduced transcription of the *gcvB* gene. The position and nature of some of these mutations are shown in Fig. S1 of the manuscript (section II.2.1). In addition, two mutations were found to affect the CG-rich stem of *gcvB* Rho-independent transcription terminator (see Fig. 16, above). Similar mutations were previously found in other regulatory sRNAs and were shown to cause a

decrease in sRNA levels, presumably by destabilizing the stem and making the sRNA more susceptible increased 3' end degradation (Balbontin et al., 2010, Bossi & Figueroa-Bossi, 2007, Figueroa-Bossi et al., 2009). This possibility was not tested here.

In parallel to the above analysis, a procedure identical to that depicted in Fig. 17A was carried on a DNA segment spanning the 5' end of the *yifK* gene. In this case, most of the red colonies analyzed were found to contain two or more mutations at different positions. This prevented the interpretation of the sequence data and the mutants were discarded. Only three mutants containing a single base-changes were identified; one carried a mutation in the -35 box of the *yifK* promoter increasing its strength, and two carried mutations inside the *yifK* 5' UTR that appear to increase the stability of the mRNA. The characterization of these mutants can be found in section II.2 below.

II.1.2. Genetic analysis of GcvB:*yifK* interaction unveils a translational enhancer in the target sequence

No mutations affecting the presumptive pairing sequences of GcvB or *yifK* mRNA were found by the above analysis. We interpreted this finding as an indication that perhaps single base changes did not relieve repression enough to cause a detectable color change on MacConkey plates. We thus proceeded to introduce multiple mutations by site-directed mutagenesis. As an initial test, we changed a four-base segment in the pairing interval: the sequence UGUG in GcvB was changed to ACAC; the reciprocal change was introduced at the corresponding position of *yifK* (ACAC to UGUG). The effects of these modifications on the expression of the *yifK-lacZY* fusion were assessed measuring β -galactosidase activity (Fig. 18).

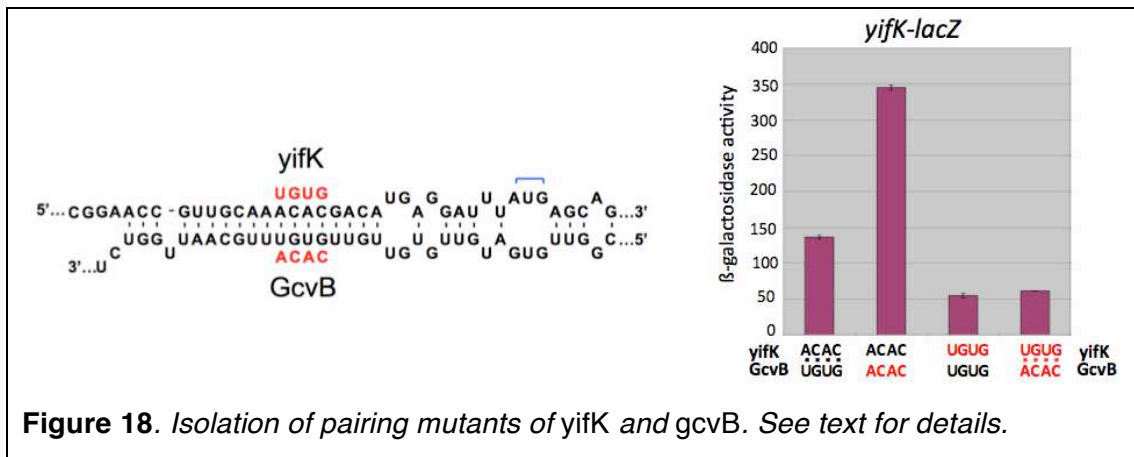
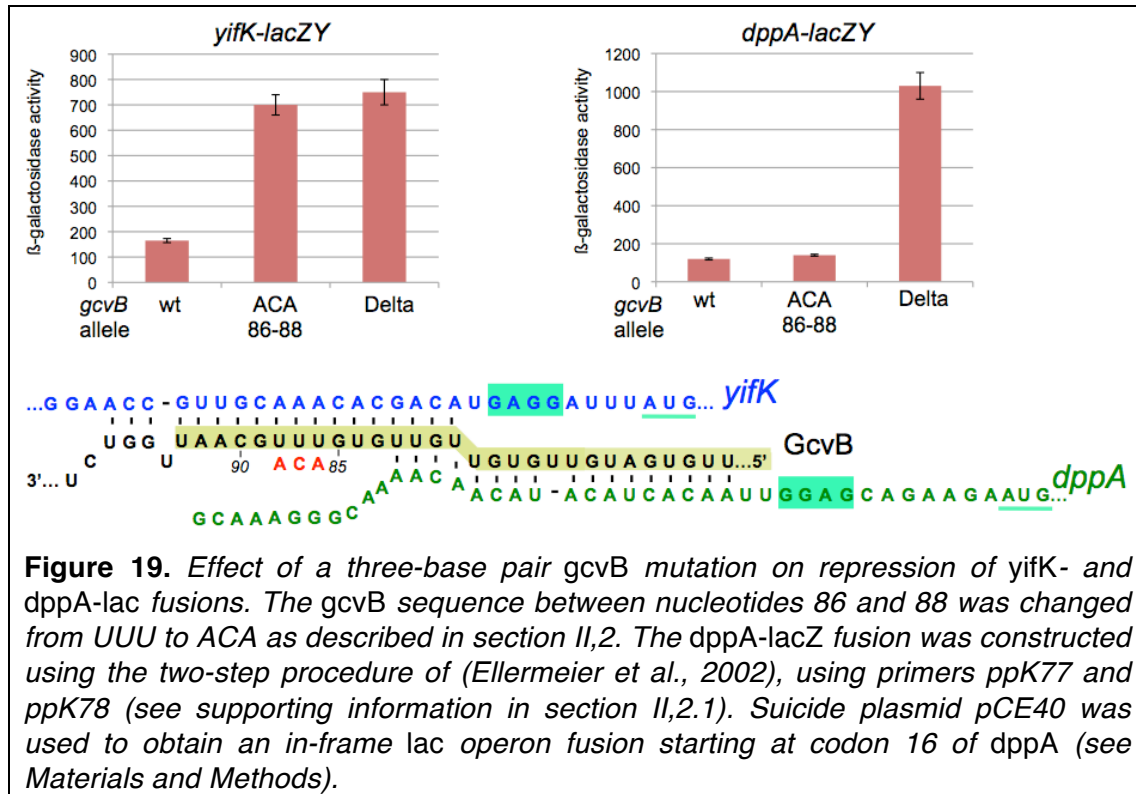


Fig. 18 shows that changing the UGUG segment in *GcvB* to ACAC causes *yifK-lacZY* expression to increase a little more than two-fold. This is less than the increase observed upon deleting the *gcvB* gene (presumably because the four-base change still does not completely abolish the interaction) but it is nonetheless indicative of the involvement of the sequence in *yifK* repression. In contrast, analysis of the reciprocal change in *yifK* gave an unexpected result: when the ACAC sequence was changed to UGUG, *yifK* expression did not increase (as one expected due to reduced pairing with *GcvB*) but actually decreased. In addition, this decrease was not aggravated when we combined the ACAC \rightarrow UGUG *yifK* mutant with the *GcvB* mutant carrying the complementary change, UGUG \rightarrow ACAC (fourth bar in Fig. 18). These findings suggested that the ACAC sequence in *yifK* stimulated *yifK* expression and this effect superseded *GcvB* regulation. To analyze this phenomenon in more detail, the core of the presumptive pairing interval was arbitrarily divided in two parts that were mutagenized separately. The results of this analysis, described in detail in the manuscript in section II.2 provided convincing evidence that *GcvB* represses *yifK* through base-pairing and corroborated the existence of an enhancer element embedded in the *yifK* sequence recognized by *GcvB*.

II.1.3. Separate subdomains of *GcvB* R1 region pair with *yifK* and *dppA*

One of the *gcvB* mutants constructed in the course of this study (a three-base change at positions 86-88) causes a nearly complete relief of *yifK*

repression (comparable to a *gcvB* deletion; Fig. 19). Interestingly, this triple mutant has absolutely no effect on repression of a *dppA-lacZY* (Fig. 19) This confirms that separate R1 domains are involved in the recognition of *yifK* and *dppA* targets.



II.1.4. Role of ACA motifs in translation enhancement

A major part of my thesis project was devoted the characterization of the enhancer element discovered in *yifK* 5' UTR. Experiments described in section II.2 indicated that the enhancer acted at the level of translation initiation. These experiments correlated the enhancer activity with two ACA triplets in the *yifK* target sequence. Initial evidence for of the importance of these motifs in *yifK* expression was obtained in an experiment in which either ACA triplet was replaced by random triplets (in a strain carrying *yifK-lacZY* and deleted for *gcvB*) and mutants were screened on MacConkey-lactose indicator plates. As one can see in Fig. 20 the mutant colonies exhibited a whole range of red color intensities, from light pink to deep red, indicative of a wide range of expression levels.

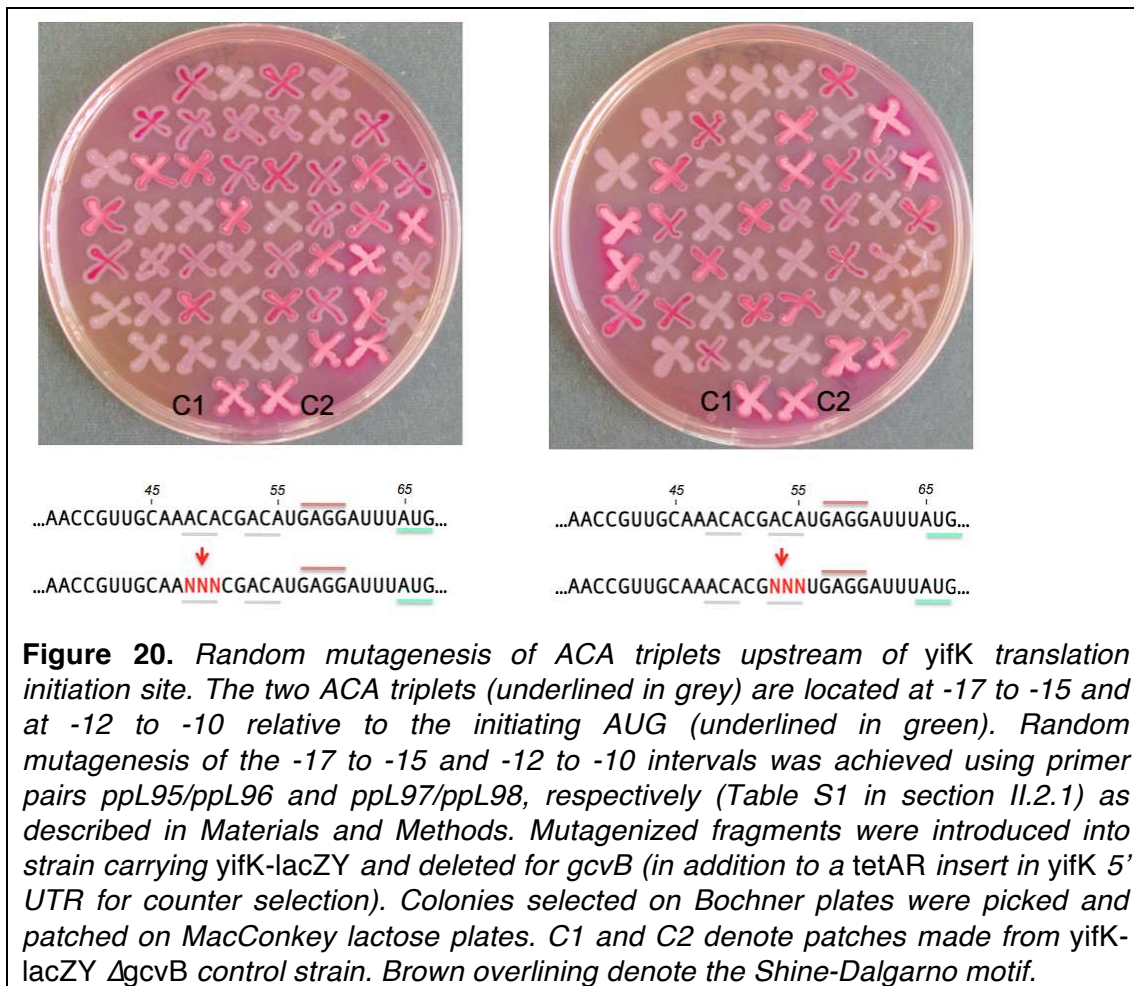


Figure 20. Random mutagenesis of ACA triplets upstream of *yifK* translation initiation site. The two ACA triplets (underlined in grey) are located at -17 to -15 and at -12 to -10 relative to the initiating AUG (underlined in green). Random mutagenesis of the -17 to -15 and -12 to -10 intervals was achieved using primer pairs ppL95/ppL96 and ppL97/ppL98, respectively (Table S1 in section II.2.1) as described in Materials and Methods. Mutagenized fragments were introduced into strain carrying *yifK-lacZY* and deleted for *gcvB* (in addition to a *tetAR* insert in *yifK* 5' UTR for counter selection). Colonies selected on Bochner plates were picked and patched on MacConkey lactose plates. C1 and C2 denote patches made from *yifK-lacZY ΔgcvB* control strain. Brown overlining denote the Shine-Dalgarno motif.

DNA sequence analysis identified triplets which have a particular detrimental effect. For example, replacing the upstream ACA by GGG causes a 92% reduction in *yifK* translation efficiency (see section II.2 below).

II.2. Manuscript submitted to PloS Genetics (under revision; August 10)

**The Translation Enhancer Activity of an ACA Motif Dictates
the Response of a *Salmonella* Gene to Multi-Target Small
RNA Regulator GcvB**

Qi Yang, Nara Figueroa-Bossi and Lionello Bossi*

Centre de Génétique Moléculaire du CNRS, Associé à l'Université Paris-Sud,
91198 Gif-sur-Yvette, France.

Running head: Translation enhancer dictates sRNA response

* E-mail : bossi@cgm.cnrs-gif.fr

Tel (+33) 1 69 82 38 11

Fax (+33) 1 69 82 38 77

Abstract

GcvB is an archetypal multi-target small RNA regulator of genes involved in amino acid uptake or metabolism in enteric bacteria. Included in the GcvB regulon is the *yifK* locus, encoding a conserved putative amino acid transporter. Here we show that GcvB inhibits *yifK* mRNA translation by pairing with a sequence immediately upstream from the Shine-Dalgarno motif. Surprisingly, we found that some target sequence mutations that disrupt pairing, and thus are expected to relieve repression, actually lower *yifK* expression and cause it not to respond to GcvB variants carrying the corresponding compensatory changes. Work prompted by these observations revealed that the GcvB target sequence in *yifK* mRNA includes elements that stimulate translation initiation. Replacing each base of an ACA triplet near the center of the target sequence, by any other base, leads to a decrease in *yifK* expression. Effects are additive, with some triple replacements causing up to a 90% reduction. These findings identify the ACA triplet as a component of a translational enhancer and reveal that GcvB acts by countering the enhancer function. We show that the enhancer activity does not require the ACA motif to be located at a fixed distance from the Shine-Dalgarno sequence nor does it depend on the spacing between the latter and the initiating AUG. Finally, additional data show that besides the GcvB control, *yifK* is regulated at the transcriptional level by the leucine responsive regulator Lrp, which completely silences *yifK* expression in leucine-deprived minimal medium.

Author summary

The majority of small RNA (sRNA) regulators in bacteria act by inhibiting translation initiation in target messenger RNAs. The study of this regulatory mechanism not only allows a better understanding of sRNA function but it can also provide new insight into aspects of the translation initiation process that remain incompletely characterized. This was the case in the work described here. Analyzing the mechanism by which GcvB, a multi-target sRNA, downregulates a putative amino acid transporter in *Salmonella*, we discovered that the sequence base-pairing with GcvB in the target mRNA functions as a translational enhancer. Replacing an ACA triplet near the center of the sequence with unrelated triplets leads to a decrease in translational initiation efficiency that can be as severe as more than 90%. Interestingly, some of these replacements concomitantly render the mRNA insensitive to GcvB variants carrying the appropriate compensatory changes, suggesting that targeting the enhancer element is paramount for GcvB regulatory effectiveness. Overall the data presented in the paper unveil the role of the ACA motif in the translation initiation process and lay the grounds for further analysis of the mechanism involved.

Introduction

A relevant chapter in the expanding field of RNA-mediated gene regulation is devoted to the activities of multi-target trans-encoded small RNAs in bacteria. Acting in concert with chaperon protein Hfq, these RNA regulators function by base-pairing with short, often imperfectly complementary sequences in the 5' untranslated regions (UTR) of target messenger RNAs. They can affect translation and turnover of several mRNAs simultaneously thus reprogramming gene expression of whole gene families in a coordinate manner in response to environmental cues (reviewed in (Gottesman & Storz, 2011b, Storz *et al.*, 2011, Vogel & Luisi, 2011). Archetypal examples of this class of regulators are the RybB small RNA (sRNA) which represses expression of mRNA for dispensable iron-sequestering proteins when iron is limiting (Jacques *et al.*, 2006, Masse *et al.*, 2003, Masse & Gottesman, 2002, Masse *et al.*, 2005, Prevost *et al.*, 2011); RybB, which downregulates several outer membrane protein mRNAs under envelope stress conditions (Balbontin *et al.*, 2010, Gogol *et al.*, 2011, Johansen *et al.*, 2006, Papenfort *et al.*, 2006, Thompson *et al.*, 2007), Spot 42, which amplifies the regulatory range of catabolite repression by targeting several mRNAs involved in sugar uptake and consumption (Beisel & Storz, 2011) and GcvB, which downregulates dozens of different mRNAs involved in amino acid uptake or metabolism in *E. coli* and *Salmonella* (Pulvermacher *et al.*, 2009b, Sharma *et al.*, 2007, Sharma *et al.*, 2011, Urbanowski *et al.*, 2000). GcvB, a 200 nucleotide-long sRNA, was identified serendipitously during a study of *gcvA*, the

gene for the main transcriptional regulator of the glycine cleavage operon *gcvTHP* (Urbanowski *et al.*, 2000). The latter encodes the enzymes of the glycine cleavage system, the pathway generating one-carbon units from the oxidative cleavage of glycine (Stauffer & Stauffer, 2005). The *gcvB* gene is located immediately adjacent to *gcvA* in the opposite orientation with its promoter partially overlapping the *gcvA* promoter. In the presence of excess glycine, the GcvA protein activates transcription of the *gcvTHP* operon as well as of *gcvB* (Urbanowski *et al.*, 2000). Initial characterization of GcvB showed this sRNA to downregulate the synthesis of DppA and OppA proteins, main components of dipeptide- and oligopeptide-transport systems, respectively (Sharma *et al.*, 2007, Urbanowski *et al.*, 2000). Since then, the number of genes found to be regulated by GcvB has increased exponentially. A recent transcriptomic study in *Salmonella enterica* set this number to more than 40, making the GcvB regulon the largest of its kind (Sharma *et al.*, 2011). The vast majority of these loci are linked directly or indirectly to amino acid metabolism and are negatively controlled by GcvB. Typically, regulation is exerted during exponential growth in nutrient rich environments and possibly aimed at coordinating the expression of interconnected metabolic pathways (Sharma *et al.*, 2007, Sharma *et al.*, 2011); however, its precise role remains poorly defined.

GcvB uses a specific sequence region to pair with most, although not all (Coornaert *et al.*, 2013) of its mRNA targets. This pairing domain – named the R1 region (Sharma *et al.*, 2007) – is characterized by its high degree of sequence conservation, the lack of secondary structure and a typical GU-rich sequence bias.

Hence, most sequences targeted by GcvB include CA-rich repeats. They are typically found in 5' UTRs of target mRNAs where GcvB pairing inhibits translation initiation causing the destabilization of the mRNA. In one of these targets - the *gltI* mRNA for a glutamate-aspartate transport protein – the CA-rich element is located 45 nucleotides (nt) upstream from the translation initiation codon. Removal of this sequence (as part of a 27 nt deletion), besides causing the loss GcvB regulation, affected *gltI* translation, suggesting that the CA-rich element acts as a translational enhancer. Consistent with this interpretation, crafting the 27 nt segment at the corresponding position of an unrelated mRNA conferred simultaneously GcvB control and increased translational efficiency (Sharma et al., 2007).

Some years ago, our laboratory performed a *lac*-based genetic screen aimed at identifying genes controlled by trans-encoded small RNAs in *Salmonella*. A random library of *lacZ* fusions to chromosomal genes was generated using a phage Mu-derived transposon (MudK) and screened for isolates whose LacZ levels changed (either increased or decreased) upon inactivating Hfq (Figuroa-Bossi *et al.*, 2006b). Among the candidates that were found, two independent isolates upregulated in the *hfq* mutant background carried the *lacZ* insert translationally fused to the *yifK* gene (Figuroa-Bossi *et al.*, 2006b). Presumptive identification of this gene as an amino acid transporter suggested that *yifK* might be a GcvB target. We thus proceeded to test this hypothesis and characterize *yifK* regulation. While this work was underway, Sharma and coworkers identified *yifK* mRNA as a member of the *gcvB* regulon by microarray analysis; however, these authors

could not confirm direct regulation by GcvB due to low reporter fluorescence of the *yifK-gfp* fusion used in the study (Sharma *et al.*, 2011).

Here we present *in vivo* and *in vitro* evidence that GcvB downregulates *yifK* directly by pairing with a sequence immediately preceding the Shine-Dalgarno (SD) motif in *yifK* mRNA. A surprising observation in the course of this study was that some target sequence mutations that disrupted pairing did not cause *yifK* expression to increase –as expected for the relief of GcvB repression – but had the opposite effect. The drop in expression was not suppressed by deleting *gcvB* nor was it accentuated in a GcvB mutant carrying the appropriate compensatory changes. Closer analysis revealed that GcvB target sequence includes elements that stimulate *yifK* mRNA translation. In the absence of such elements, the role of GcvB pairing in regulation becomes marginal.

Results

Genetic identification of a GcvB-regulated locus

Our original screen for Hfq-regulated genes yielded two isolates carrying the MudK (*lac*) transposon in the *yifK* gene; one predicted to produce a LacZ protein fusion to the 48th amino acid (aa) of the 461 aa YifK (*yifK87::MudK*); the other with LacZ inserted after the 95th aa of YifK (*yifK88::MudK*) (Figuroa-Bossi *et al.*, 2006b). Preliminary tests showed both fusions to be regulated in a closely similar manner; however, *yifK87::MudK* produced significantly higher β -galactosidase activity and was chosen for the present study. A survey of protein sequence

databases showed YifK to be a highly conserved protein with the characteristic signature of amino acid transporters. The known role of GcvB in the regulation of some members of this family made this small RNA the likeliest candidate to control *yifK* expression. This was confirmed by deleting the *gcvB* gene and testing the effects of the deletion on the expression of the *yifK87::MudK* fusion (hereafter referred to as *yifK-lacZY*). As shown in Figure 1, the *gcvB* deletion causes a nearly 5-fold increase of β -galactosidase activity in exponentially growing cells, while effects decline in stationary phase. Somewhat surprisingly, LacZ levels in the *gcvB*-deleted strain are not as high as the levels measured in a strain deleted for *hfq* (Figure 1). This might reflect the existence of one or more additional sRNA(s) participating in *yifK* repression. Alternatively, Hfq could repress *yifK* directly (Desnoyers & Masse, 2012). The data in Figure 1 show that loss of Hfq is epistatic to the *gcvB* deletion.

Mutations affecting of *yifK* expression

Primer extension experiments mapped the 5' end of *yifK* mRNA to 64 nucleotides (nt) upstream from the initiating AUG (Figure 2). This 5' untranslated region (UTR) includes a 14-nucleotide sequence stretch complementary to the 3' half of GcvB's R1 region. As an initial step to characterize GcvB involvement in *yifK* regulation, we tested whether point mutations in the *gcvB* gene or in the promoter-proximal portion of *yifK* relieved GcvB-mediated repression. For this, DNA fragments spanning either of these two regions were randomly mutagenized by the polymerase chain reaction (PCR) under error-prone conditions and introduced into the

chromosome of a strain harboring the *yifK-lacZY* reporter fusion via lambda *red* recombination. Selection of recombinants on MacConkey lactose indicator plates allowed distinguishing derepressed mutant candidates (red colonies) from the background of non mutagenized clones (pink colonies). Most of the isolates originating from the *gcvB* mutagenesis carried changes in the *gcvB* promoter or in the promoter of the adjacent *gcvA* gene (Figure S1). Thus, these mutations appeared to lower the levels rather than the activity of GcvB and were not further considered. No mutations likely to affect GcvB's ability to pair with *yifK* mRNA were found.

Mutagenesis of *yifK* promoter-proximal segment yielded three mutants with elevated *yifK-lacZY* expression. One isolate carried a C:G to A:T change 33 base-pairs upstream from the 5' end of *yifK* mRNA. The position and the nature of the change (producing a -35 promoter consensus match, TTGACA, Figure 2A), strongly suggest that the mutation increases the strength of the *yifK* promoter. The mutation leads to a sharp rise in the intensity of the primer extension product (Figure 2B) and a more than 10-fold increase in β -galactosidase activity (data not shown). These findings confirmed that the 5' end identified by primer extension corresponds to *yifK* transcription initiation site. The remaining two mutations affect residues within the 5' UTR (Figure 2A). One allele, resulting in a U to C change at position +21, falls within a AU-rich segment (AUAACAAUAA) that might constitute a site for Hfq binding (Brennan & Link, 2007, Vogel & Luisi, 2011). Consistent with this interpretation, the mutation has no effect in Δ *hfq* background (Figure 2C). Finally

the third allele (G to A at +27) affects the CG-rich stem of a presumptive secondary structure immediately adjacent to the AU-rich segment. The change causes *yifK-lacZY* expression to increase even in the *hfq* mutant (Figure 2C) suggesting that it affects an Hfq-independent step. To date, however, this step remains unidentified.

GcvB inhibits *yifK* translation by pairing with an enhancer element

The above approach yielded no mutations affecting the presumptive pairing sequences of GcvB or *yifK*. Reasoning that single base changes might not disrupt regulation enough to be revealed by the MacConkey plate screen, we resorted to introducing multiple changes by site-directed mutagenesis. An initial test involved changing a UGUG quadruplet in the GcvB segment thought to pair with *yifK* mRNA. The alteration caused expression of the *yifK-lacZY* fusion to increase approximately threefold, thus corroborating the postulated role of this sequence in *yifK* repression. Unexpectedly, however, when the ACAC sequence at the corresponding position in *yifK* mRNA was changed, *yifK-lacZ* expression did not increase but actually declined (data not shown). Trying to clarify this observation, portions of the region of interest were mutagenized separately. As shown in Figure 3A, converting an AAA triplet in the middle of the *yifK* target sequence to UGU, or making the opposite change (UGU to AAA) in GcvB, similarly relieves *yifK-lacZY* repression. Repression is restored upon combining the compensatory alleles. Thus, this portion of the target sequence behaves as expected, and the compensatory mutant's phenotype strongly suggests that GcvB

represses *yifK* through a base-pair interaction. In vitro toeprint experiments, showing that GcvB inhibits the binding of ribosomal 30S subunit to *yifK* translation initiation site, specifically and in a dose-dependent manner (Figure 4), provided independent support to this conclusion.

Again, however, changing the CA doublet on the 3' side of the AAA triplet to UC produced an unusual pattern: like in the quadruplet mutant above, *yifK-lacZY* expression decreased rather than increase, becoming insensitive to a GcvB variant carrying the compensatory change (Figure 3B) as well as to *gcvB* or *hfq* deletions (Figure 3C). We interpreted these findings to suggest that the CA to UC conversion affects translation and in that under such conditions, GcvB action is no longer rate-limiting for *yifK* expression. The effects of the mutation on translational efficiency were examined *in vitro* using a reconstituted system. Results in Figure 5 showed an epitope-tagged Cat protein to accumulate at significantly greater levels when made from a gene fusion to the wt *yifK* 5' UTR than from an equivalent construct carrying the CA to UC change. These data confirm that the CA doublet stimulates translation and suggest that GcvB effectiveness in regulation reflects the targeting of an activating element.

Anatomy of a translational enhancer

To characterize the enhancer element, selected portions of the ribosome binding sequence of *yifK* were modified by systematically changing individual residues to each of the three alternative bases. In particular, we concentrated on two ACA triplets, one corresponding to the site described above (positions -17 to -15 relative to the

initiating AUG) the other nearly adjacent to the Shine-Dalgarno sequence (-12 to -10). This analysis was performed entirely in a strain deleted for the *gcvB* gene. As shown in Figure 6A, any nucleotide change in the upstream ACA lowers *yifK-lacZY* expression. Variations range between 23% and 62%, with G residues exerting the most adverse effects at any position. The effects appear to be additive since a separate experiment in which all three bases in the ACA triplet were randomized yielded alleles undergoing as much as 92% reduction in *yifK-lacZY* expression (Figure 6B).

Alteration of the downstream ACA produced a somewhat different pattern. Changes in the central C are either neutral or stimulatory; in contrast, having a C at the first triplet position is highly deleterious resulting in nearly 95% reduction of β -galactosidase activity (Figure 6A). Although the physical basis for these effects remains elusive, the data in Figure 6 clearly point to a role of ACA motifs in *yifK* translation.

A peculiarity of the *yifK* translation signal is the unusually short distance (four nucleotides) between the most conserved base of the SD motif (Shultzaberger *et al.*, 2001) and the initiating AUG. We thus envisaged that the role of the enhancer could be to somehow compensate for such suboptimal arrangement. To test this possibility, we generated a 7 nt tandem direct duplication of the SD region and then inactivated either copy of the SD by changing the GAGG motif to GACG (Figure 7). Thus, the resulting constructs have the functional SD sequence positioned either 4 or 11 nt from the AUG. As shown in Fig 7, these two variants (n. 4 and n. 5) express β -galactosidase levels that are

similar to each other and to the strain in which both SD are functional (n. 1). However, when the upstream ACA motif is replaced by the GGG triplet, *yifK-lacZ* expression drops sharply in both constructs (Figure 5, compare n. 4 and n. 6, and n. 5 and n. 7). Thus, these results indicate that the enhancer activity does not require the ACA motif to be positioned at a fixed distance from the initiation site and it does not depend on the spacing between the SD and the initiation codon. In an experiment carried out in parallel to that in Figure 7, the upstream ACA triplet was replaced by random sequences and the mutants obtained screened directly on MacConkey-lactose indicator plates. The color of colonies spanned the whole range of shades between white and red (Figure S2), confirming that the data in Figure 7 do not reflect a specific inhibitory effect of the triple G, and providing a snapshot of the nuances in the enhancer-dependent effects on *yifK-lacZ* expression.

GcvB activity stimulates RNase E-dependent *yifK* mRNA decay

Northern blot analysis was used to assess the effects of GcvB regulation on *yifK* mRNA levels. This study critically benefited from the availability of the -33 promoter mutant (see above), *yifK* mRNA being otherwise undetectable when expressed from the wild-type promoter (compare Figs 7A and 7B). The analysis identified two *yifK* mRNA species, a 1.4 kilobase (Kb) mRNA covering just the *yifK* coding portion and a longer, 2.0 Kb RNA extending into the adjacent *argX-hisR-leuT-proM* tRNA operon. As shown in Figure 7, both RNAs accumulate upon RNase E inactivation, whereas only the shorter species accumulates in cells lacking GcvB or Hfq. We

interpret these results to indicate that derepression of *yifK* translation in Δ *gcvB* or Δ *hfq* cells protects the 1.4 Kb RNA against RNase E cleavage, and that this protection does not extend to the untranslated, tRNA coding portion of the 2.0 Kb RNA. Absence of any obvious transcription termination signals in the intercistronic region between *yifK* and the tRNA operon suggests that the 1.4 Kb RNA originates from processing of the longer form. Likely, under normal conditions (*i.e.*, wt *yifK* promoter) *yifK* transcription contributes to only a small fraction of the four tRNAs, as most the tRNA operon transcription results from a strong promoter located immediately upstream from the *argX* gene (Bossi & Smith, 1984). The approximately 500 nt RNA accumulates in the RNase E mutant (Figure 7C). Previous work in *E.coli*, showed that this tRNA precursor is processed by the concerted actions RNase E and RNase P in a pathway that, intriguingly, also sees the participation of Hfq (Zhang et al., 2003).

yifK is repressed by Lrp

Early on in this study, it became apparent that *yifK* expression was exquisitely sensitive to the growth medium and virtually silenced in minimal medium. As a result, a strain with the *yifK-lacZY* fusion is phenotypically Lac⁻ when plated in minimal medium. We exploited this phenotype to positively select for spontaneous Lac⁺ mutants. The selection yielded two classes of mutations, one genetically linked to the *yifK-lacZY* locus, the second mapping elsewhere. All of the linked mutants that were analyzed were found to harbor the -33 C:G to A:T promoter change obtained previously (see above). The unlinked mutations mapped in a chromosomal interval encompassing the gene

for leucine response regulator, Lrp. Prompted by this observation, we introduced an *lrp* insertion mutation into the *yifK-lacZY*-containing strain. The resulting strain acquired a Lac⁺ phenotype (Figure 8A), indicating that *yifK* silencing in minimal medium results from Lrp repression. Addition of leucine efficiently relieves repression (Figure 8). The data in Figure 8 also show that GcvB does not contribute to *yifK* repression to any significant extent in minimal medium. This is not surprising as GcvB is transcribed at very low level under these conditions (Sharma et al., 2007) and inactivating Lrp does not reverse this pattern (Figure S3). The data in Figure S3 differ from those of Modi et al (Modi *et al.*) who reported an approximate 30-fold increase in GcvB levels in an *lrp* deletion mutant of in *E.coli*. This discrepancy might reflect differences in the organisms used or in media composition.

Discussion

In the present work, we have characterized the regulation of *Salmonella*'s *yifK* locus encoding a putative amino acid transporter highly conserved in *Enterobacteriaceae*. Our analysis showed *yifK* to be negatively controlled at the transcriptional level by the leucine response regulator Lrp, and at the post-transcriptional level by GcvB sRNA. These findings place *yifK* at the intersection of two global regulatory networks devoted to amino acid management (Calvo & Matthews, 1994, Sharma et al., 2011). The relative impacts of two systems on *yifK* expression vary as a function of growth conditions, with the Lrp control predominating in leucine-deprived

poor media and the GcvB control operating when amino acids are plentiful, possibly in excess. The sole condition where *yifK* appears to escape negative control is leucine-supplemented minimal medium, where Lrp repression is relieved. This response closely parallels that of the oligopeptide permease operon, *oppABCDF* (Calvo & Matthews, 1994) whose transcript is also a target of GcvB repression (Sharma et al., 2007, Urbanowski et al., 2000). Likely, the overlap of Lrp and GcvB networks reflects the link between amino acid metabolism and one-carbon units production; however, the precise physiological role and the implications of the above responses remain incompletely understood.

Genetic analysis of GcvB:*yifK* mRNA interactions revealed that the GcvB target sequence in *yifK* mRNA contains an enhancer element. Intriguingly, mutations that disrupt the enhancer - and lower *yifK* expression as a result - render *yifK* expression totally insensitive to GcvB repression. This suggests that the effectiveness of GcvB regulation is dependent on the enhancer function and that when this component is removed, GcvB-mediated repression no longer constitutes a rate-limiting step in *yifK* expression. Sharma and coworkers (2007) previously showed that GcvB's target sequence in the *gltI* gene of *Salmonella* acts as transferable translation enhancer (see Introduction). Unlike in our study, the effects of GcvB as a translational repressor were much greater than the effects of removing the enhancer, leading the authors to conclude that GcvB did not simply block the enhancer effect (Sharma et al., 2007). It seems possible that the plasmid-borne nature of the *gcvB* gene in the study by Sharma and coworkers made the GcvB repression tighter

than when the sRNA is expressed from the chromosome. Alternatively, the enhancer activity might be less conspicuous in *gltI* than in *yifK*. The *gltI* enhancer, located 45 nt upstream from the initiation codon, was characterized as part of a 27 nt segment and not analyzed in any further detail (Sharma et al., 2007). Here we found that nucleotide replacement in either of two ACA triplets within GcvB target site in *yifK* can result in more than 90% reduction in *yifK* expression. Although our data do not allow defining the contours of the enhancer element, they unequivocally identify the ACA motif as a determinant of its activity. We also found that the enhancer activity is maintained following a 7 nt shift in the position of the initiation site, suggesting the absence of strict spatial requirements for the functioning of the element. This is consistent with the data from the *gltI* system and with a report showing CA repeats to stimulate translation even when placed downstream from the start codon (Martin-Farmer & Janssen, 1999).

Translation initiation efficiencies have been known to vary greatly as a function of the sequence context of the initiation region (Dreyfus, 1988, Yarchuk *et al.*, 1992). Computational analysis of sequences surrounding translation initiation sites of *E.coli* genes showed that the spacing between the SD and the initiation codon affects SD sequence conservation and its pattern. This study did not reveal significant biases outside these main elements (Shultzaberger et al., 2001); however, conserved patterns occurring at variable positions might have been difficult to identify by the statistical analysis. Indeed, separate lines of evidence point to the role of the ACA motif in translation initiation. The motif is

found in other translation enhancer sequences (Komarova et al., 2002, McCarthy *et al.*, 1985) and, as an ACAA repeat, was shown to promote translation initiation in the absence of a SD sequence (Tzareva et al., 1994). ACA is also found in the loops of pseudoknots formed by RNA ligands to ribosomal protein S1, obtained through Systematic Evolution of Ligands by Exponential Enrichment (SELEX) (Ringquist *et al.*, 1995) and is part of the SELEX-determined consensus sequence for binding of protein CsrA, a translational regulator (Dubey *et al.*, 2005). Finally, ACA is the recognition sequence of the MazF endonuclease that inactivates *E.coli* mRNAs by preferentially cleaving near the translation initiation codon (Vesper *et al.*, 2011).

The lack of position requirements for the functioning of the enhancer suggests that its role is to provide an anchor point for the 30 S ribosomal subunit so as to facilitate subsequent recognition of the SD sequence. Some of the evidence reviewed above tentatively identifies protein S1 as the possible candidate for the interaction. In vitro S1-binding studies with some of the mutants constructed in the course of this work should allow testing of this idea. Combined with the mutational analysis of other GcvB-regulated mRNAs, this approach might provide further insight into how the ACA motif participates in the translation initiation step.

Materials and Methods

Bacterial strains and culture conditions

Strains used in this study were derivatives of *Salmonella enterica* serovar Typhimurium strain LT2 (Lilleengen, 1948). Strain SV4280 was a gift of J. Casadesús. Except for the latter strain and for strain MA7224, all other strains were derived from MA3409, an LT2 derivative cured for the Gifsy-1 prophage (Figueroa-Bossi *et al.*, 1997). The genotypes of the relevant strains used are listed in Table 1. Bacteria were cultured at 37°C in liquid media or in media solidified by the addition of 1.5% Difco agar. LB broth (Bertani, 2004) was used as complex medium. Carbon-free medium (NCE) (Maloy & Roth, 1983), supplemented with 0.2% glycerol or 0.2% lactose was used as minimal medium. Antibiotics (Sigma-Aldrich) were included at the following final concentrations: chloramphenicol, 10 µg ml⁻¹; kanamycin monosulphate, 50 µg ml⁻¹; sodium ampicillin 100 µg ml⁻¹; spectinomycin dihydrochloride, 80 µg ml⁻¹; tetracycline hydrochloride, 25 µg ml⁻¹. MacConkey agar plates containing 1% lactose (Macconkey, 1905) were used to monitor *lacZ* expression in bacterial colonies. Liquid cultures were grown in New Brunswick gyratory shakers and growth was monitored by measuring the optical density at 600 nm with a Shimadzu UV-mini 1240 spectrophotometer.

Relevant enzymes and chemicals

T4 polynucleotide kinase and Taq DNA polymerase were from New England Biolabs, Pfu-Turbo DNA polymerase was from Stratagene, T4 DNA ligase was from New England Biolabs. DNA oligonucleotides were

custom synthesized by Sigma Aldrich or Eurofins MWG/Operon. The complete list of DNA oligonucleotides used in this study is shown in Table S1. DNA sequencing was performed by GATC biotech. Acrylamide-bisacrylamide and other electrophoresis reagents were from BioRad. Agarose was from Invitrogen. Hybond-N⁺ membranes and hybridization buffer used for Northern blot analysis were from GE Healthcare and from Applied Biosystems-Ambion, respectively. The rNTPs were from Promega and the ³²P-NTPs were from PerkinElmer or Hartmann Analytic. ³²P-labeled nucleic acids were detected by phosphorimaging using ImageQuant software.

Genetic techniques

Generalized transduction was performed using the high-frequency transducing mutant of phage P22, HT 105/1 *int-201* (Schmieger, 1972) as described (Lemire *et al.*, 2011). Chromosomal engineering (recombineering) was carried out by the λ *red* recombination method (Datsenko & Wanner, 2000, Murphy *et al.*, 2000, Yu *et al.*, 2000) implemented as in (Datsenko & Wanner, 2000). Donor DNA fragments were generated by PCR using plasmid DNA or chromosomal DNA or DNA oligonucleotides as templates. Amplified fragments were electroporated into appropriate strains harboring the conditionally replicating plasmid pKD46, which carries the λ *red* operon under the control of the P^{BAD} promoter (Datsenko & Wanner, 2000). Bacteria carrying pKD46 were grown at 30°C in the presence of ampicillin and exposed to arabinose (10 mM) for 3 hours prior to preparation of electrocompetent cells. Electroporation was carried out using a Bio-Rad MicroPulser under

the conditions specified by the manufacturer. Recombinant colonies were selected on LB plates containing the appropriate antibiotic. Constructs were verified by PCR and DNA sequence analysis (performed by GATC company).

Random PCR mutagenesis

PCR amplification of DNA fragments under error-prone conditions was carried out as previously described (Bossi & Figueroa-Bossi, 2007).

“Scarless” DNA recombineering

Scarless modification of chromosomal DNA sequences at the single base-pair level was achieved with a two-step recombineering procedure as previously described (Bossi *et al.*, 2012). Briefly, this involved: 1) inserting a *tetAR* module (produced by PCR) at the chromosomal site to be modified and: 2) replacing the *tetAR* module by a DNA fragment carrying the desired change through positive selection tetracycline-sensitive recombinants (Bochner *et al.*, 1980). Typically, the DNA fragment in the second step was also obtained by PCR using oligonucleotides with complementary sequences at their 3' ends priming DNA synthesis on each other (“reciprocal priming”). In site-directed mutagenesis experiments, one of the two primers contained the desired nucleotide changes or randomized sequence stretches. All constructs were verified by DNA sequencing. Table S2 shows the list of alleles made by standard or scarless recombineering.

RNA extraction and analysis by primer extension and Northern blotting

RNA was prepared by the acid-hot-phenol method from exponentially growing cells (OD₆₀₀ of 0.35) as previously described (Bossi & Figueroa-Bossi, 2007). Reverse transcriptase reactions (enzyme Superscript II from Invitrogen) were carried out using 5 µg of bulk RNA and ³²P-labeled primer ppF49. The same DNA primer was used for the sequencing reactions. Reactions were performed with the *fmol* DNA Cycle Sequencing System from Promega, according to the manufacturer's protocol. Reaction products were fractionated on a 10% polyacrylamide-8 M urea gel. For Northern blot analysis, RNA was fractionated on a 1% agarose-formaldehyde gel, blotted onto a nylon membrane, and hybridized to the appropriate ³²P-labeled DNA oligonucleotide probes.

In vitro translation

In vitro coupled transcription/translation was performed using New England Biolabs' PURExpress *In vitro* Protein Synthesis kit (NEB #E6800) according to the manufacturer instructions. Genes to be analyzed were cloned under T7 promoter control in the DFRH plasmid provided with the kit. The hybrid genes carried *yifK* wt or mutant 5' UTR sequences fused to the *cat*-3xFLAG coding sequence (chloramphenicol acetyl transferase in-frame fusion to the 3xFLAG epitope). Final volume of the transcription/translation reaction was 25 µl in all cases. In addition to kit solutions A and B, reaction mix contained, 10 U of RNase inhibitor SUPERase (Ambion) and template plasmid DNA added to either 0.5 or 5 pM final concentration. Incubation times at 37 °C varied from 15 to 90 min. Reactions were stopped by addition of equal volume of 2X Laemmli buffer and immediate freezing.

Aliquots were loaded on 12.5% Acrylamide gels and Western analysis performed as previously described (Uzzau *et al.*, 2001).

Toeprinting assay

Toeprinting reactions were carried out as described by Darfeuille *et al* (Darfeuille *et al.*, 2007) with minor modifications. RNA fragments spanning positions +1 to +135 of *yjfK* mRNA were synthesized *in vitro* from T7 DNA templates generated by PCR amplification of chromosomal DNA (from strains MA8020 or MA11793) with primers ppI22 and ppI23. 2 pmol of RNA were annealed with 5'-end-labeled primer ppI23 (1 pmol) in 10 mM Tris-acetate [pH7.6], 0.1 M potassium acetate, and 1 mM DTT for 1 min at 90°C and chilled in ice for 5 min. Then, all dNTPs (final concentration 1 mM), Mg Acetate (10mM final) were added; this was followed by preincubation with 2 pmol of 30S ribosomal subunit (a gift of Dominique Fourmy and Satoko Yoshizawa) at 37°C for 5 min. In experiments involving GcvB, 5, 1 or 0.5 pmol of sRNA were added prior to both, addition of the 30S ribosomal subunit and the preincubation step. After the 5-min period, 2 pmol of tRNA^{fMet} were added and preincubation at 37°C continued for 15 additional min. Finally, Reverse Transcriptase (Superscript II, Invitrogen, 200U) was added and samples incubated for 15 min at 37°C. Following phenol chloroform extraction and ethanol precipitation, resuspended samples were loaded onto a 10% polyacrylamide-8 M urea gel along with the sequencing reaction samples generated with the same primer.

We are grateful to Dominique Fourmy and Satoko Yoshizawa for the generous gift of purified 30 S ribosomal subunits and to Josep Casadesús for strain SV4280. We thank Fabien Darfeuille for advice on the toeprinting protocol. Q. Yang was the recipient of a China Scholarship Council fellowship. This work was supported by grant BLAN07-1_187785 from the French National Research Agency (ANR) to LB.

Acknowledgements

References

- Gottesman S, Storz G (2011) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a003798.
- Storz G, Vogel J, Wassarman KM (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 43: 880-891.
- Vogel J, Luisi BF (2011) Hfq and its constellation of RNA. *Nat Rev Microbiol* 9: 578-589.
- Jacques JF, Jang S, Prevost K, Desnoyers G, Desmarais M, et al. (2006) RyhB small RNA modulates the free intracellular iron pool and is essential for normal growth during iron limitation in *Escherichia coli*. *Mol Microbiol* 62: 1181-1190.
- Massé E, Escorcía FE, Gottesman S (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* 17: 2374-2383.
- Massé E, Gottesman S (2002) A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. *Proc Natl Acad Sci U S A* 99: 4620-4625.
- Massé E, Vanderpool CK, Gottesman S (2005) Effect of RyhB small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187: 6962-6971.
- Prevost K, Desnoyers G, Jacques JF, Lavoie F, Massé E (2011) Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev* 25: 385-396.
- Balbontín R, Fiorini F, Figueroa-Bossi N, Casades J, Bossi L (2010) Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in *Salmonella enterica*. *Mol Microbiol* 78: 380-394.
- Gogol EB, Rhodius VA, Papenfort K, Vogel J, Gross CA (2011) Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proc Natl Acad Sci U S A* 108: 12875-12880.
- Johansen J, Rasmussen AA, Overgaard M, Valentin-Hansen P (2006) Conserved small non-coding RNAs that belong to the σ^E regulon: role in down-regulation of outer membrane proteins. *J Mol Biol* 364: 1-8.
- Papenfort K, Pfeiffer V, Mika F, Lucchini S, Hinton JC, et al. (2006) σ^E -dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay. *Mol Microbiol* 62: 1674-1688.
- Thompson KM, Rhodius VA, Gottesman S (2007) σ^E regulates and is regulated by a small RNA in *Escherichia coli*. *J Bacteriol* 189: 4243-4256.
- Beisel CL, Storz G (2011) The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in *Escherichia coli*. *Mol Cell* 41: 286-297.
- Pulvermacher SC, Stauffer LT, Stauffer GV (2009) Role of the sRNA GcvB in regulation of *cycA* in *Escherichia coli*. *Microbiology* 155: 106-114.
- Sharma CM, Darfeuille F, Plantinga TH, Vogel J (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev* 21: 2804-2817.
- Sharma CM, Papenfort K, Pernitzsch SR, Mollenkopf HJ, Hinton JC, et al. (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* 81: 1144-1165.
- Urbanowski ML, Stauffer LT, Stauffer GV (2000) The *gcvB* gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*. *Mol Microbiol* 37: 856-868.
- Stauffer GV (1987) Biosynthesis of Serine and Glycine. In: Neidhart FC, Ingraham JL, Low KB, Magasanik B, Schaechter M et al., editors. *Escherichia coli* and *Salmonella: Cellular and Molecular Biology*. Washington, DC: American Society for Microbiology Press. pp. 412-418.
- Coornaert A, Chiaruttini C, Springer M, Guillier M (2013) Post-Transcriptional Control of the *Escherichia coli* PhoQ-PhoP Two-Component System by Multiple sRNAs Involves a Novel Pairing

- Region of GcvB. PLoS Genet 9: e1003156.
21. Figueroa-Bossi N, Lemire S, Maloriol D, Balbontín R, Casadesús J, et al. (2006) Loss of Hfq activates the σ^E -dependent envelope stress response in *Salmonella enterica*. Mol Microbiol 62: 838-852.
 22. Desnoyers G, Masse E (2012) Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. Genes Dev 26: 726-739.
 23. Brennan RG, Link TM (2007) Hfq structure, function and ligand binding. Curr Opin Microbiol 10: 125-133.
 24. Shultzaberger RK, Bucheimer RE, Rudd KE, Schneider TD (2001) Anatomy of *Escherichia coli* ribosome binding sites. J Mol Biol 313: 215-228.
 25. Bossi L, Smith DM (1984) Conformational change in the DNA associated with an unusual promoter mutation in a tRNA operon of *Salmonella*. Cell 39: 643-652.
 26. Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, et al. (2003) Global analysis of small RNA and mRNA targets of Hfq. Mol Microbiol 50: 1111-1124.
 27. Modi SR, Camacho DM, Kohanski MA, Walker GC, Collins JJ Functional characterization of bacterial sRNAs using a network biology approach. Proc Natl Acad Sci U S A 108: 15522-15527.
 28. Calvo JM, Matthews RG (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. Microbiol Rev 58: 466-490.
 29. Martin-Farmer J, Janssen GR (1999) A downstream CA repeat sequence increases translation from leadered and unleadered mRNA in *Escherichia coli*. Mol Microbiol 31: 1025-1038.
 30. Dreyfus M (1988) What constitutes the signal for the initiation of protein synthesis on *Escherichia coli* mRNAs? J Mol Biol 204: 79-94.
 31. Yarchuk O, Jacques N, Guillerez J, Dreyfus M (1992) Interdependence of translation, transcription and mRNA degradation in the *lacZ* gene. J Mol Biol 226: 581-596.
 32. Komarova AV, Tchufistova LS, Supina EV, Boni IV (2002) Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. RNA 8: 1137-1147.
 33. McCarthy JE, Schairer HU, Sebald W (1985) Translational initiation frequency of *atp* genes from *Escherichia coli*: identification of an intercistronic sequence that enhances translation. EMBO J 4: 519-526.
 34. Tzareva NV, Makhno VI, Boni IV (1994) Ribosome-messenger recognition in the absence of the Shine-Dalgarno interactions. FEBS Lett 337: 189-194.
 35. Ringquist S, Jones T, Snyder EE, Gibson T, Boni I, et al. (1995) High-affinity RNA ligands to *Escherichia coli* ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites. Biochemistry 34: 3640-3648.
 36. Dubey AK, Baker CS, Romeo T, Babitzke P (2005) RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. RNA 11: 1579-1587.
 37. Vesper O, Amitai S, Belitsky M, Byrgazov K, Kaberdina AC, et al. (2011) Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. Cell 147: 147-157.
 38. Lilleengen K (1948) Typing of *Salmonella typhimurium* by means of bacteriophage. Acta Pathol Microbiol Scand 77: 2-125.
 39. Figueroa-Bossi N, Coissac E, Netter P, Bossi L (1997) Unsuspected prophage-like elements in *Salmonella typhimurium*. Mol Microbiol 25: 161-173.
 40. Bertani G (2004) Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. J Bacteriol 186: 595-600.
 41. Maloy SR, Roth JR (1983) Regulation of proline utilization in *Salmonella typhimurium*: characterization of *put::Mu d(Ap, lac)* operon fusions. J Bacteriol 154: 561-568.
 42. Macconkey A (1905) Lactose-Fermenting Bacteria in Faeces. J Hyg (Lond) 5: 333-379.
 43. Schmieger H (1972) Phage P22-mutants with increased or decreased transduction abilities. Mol Gen Genet 119: 75-88.
 44. Lemire S, Figueroa-Bossi N, Bossi L (2011) Bacteriophage crosstalk: coordination of prophage induction by trans-acting antirepressors. PLoS Genet 7: e1002149.

45. Datsenko KA, Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* 97: 6640-6645.
46. Murphy KC, Campellone KG, Poteete AR (2000) PCR-mediated gene replacement in *Escherichia coli*. *Gene* 246: 321-330.
47. Yu D, Ellis HM, Lee EC, Jenkins NA, Copeland NG, et al. (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A* 97: 5978-5983.
48. Bossi L, Figueroa-Bossi N (2007) A small RNA downregulates LamB maltoporin in *Salmonella*. *Mol Microbiol* 65: 799-810.
49. Bossi L, Schwartz A, Guillemardet B, Boudvillain M, Figueroa-Bossi N (2012) A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes Dev* 26: 1864-1873.
50. Bochner BR, Huang HC, Schieven GL, Ames BN (1980) Positive selection for loss of tetracycline resistance. *J Bacteriol* 143: 926-933.
51. Uzzau S, Figueroa-Bossi N, Rubino S, Bossi L (2001) Epitope tagging of chromosomal genes in *Salmonella*. *Proc Natl Acad Sci U S A* 98: 15264-15269.
52. Darfeuille F, Unoson C, Vogel J, Wagner EG (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell* 26: 381-392.
53. Figueroa-Bossi N, Valentini M, Malleret L, Fiorini F, Bossi L (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev* 23: 2004-2015.
54. Camacho EM, Casadesus J (2002) Conjugal transfer of the virulence plasmid of *Salmonella enterica* is regulated by the leucine-responsive regulatory protein and DNA adenine methylation. *Mol Microbiol* 44: 1589-1598.

Table 1. Relevant *Salmonella enterica* strains used in this work

Strain ^a	Genotype	Source or reference
MA3398	<i>zfh-8157::Tn10dTc</i>	(Figueroa-Bossi et al., 1997)
MA3409	wild-type	(Figueroa-Bossi et al., 1997)
MA7224	<i>sodCI::3xFLAG ilvI3305::Tn10dTac-cat-3xFLAG</i>	(Uzzau et al., 2001)
MA7455	wild-type / pKD46	(Figueroa-Bossi et al., 2006b)
MA8020	<i>yifK87::MudK</i>	(Figueroa-Bossi et al., 2006b)
MA8021	<i>yifK87::MudK Δhfq67::cat</i>	(Figueroa-Bossi et al., 2006b)
MA8201	<i>eptB115::MudK hfq13::Tn5-T-POP (TcR)</i>	(Figueroa-Bossi et al., 2006b)
MA9791	<i>[rluC-rne]IG::cat rne-3071 (ts)</i>	(Figueroa-Bossi et al., 2009)
SV4280	<i>lrp-42::Tn10dCm</i>	(Camacho & Casadesus, 2002)
MA10241	<i>ΔgcvB141::cat</i>	this work
MA10242	<i>ΔgcvB142::aadA (Spec^R)</i>	this work
MA10280	<i>yifK87::MudK / pKD46</i>	this work
MA10377	<i>yifK87::MudK ΔgcvB141::cat</i>	this work
MA10403	<i>yifK87::MudK ΔgcvB141::cat Δhfq67::cat</i>	this work
MA11526	<i>yifK145::tetAR / pKD46</i>	this work
MA11594	<i>yifK145::tetAR ΔgcvB141::cat / pKD46</i>	this work
MA11779	<i>ygdI143::cat</i>	this work
MA11780	<i>wecG144::cat</i>	this work
MA11781	<i>yifK[-33A] yifK87::MudK wecG144::cat</i>	this work
MA11782	<i>yifK[U21C] yifK87::MudK wecG144::cat</i>	this work
MA11783	<i>yifK[G27A] yifK87::MudK wecG144::cat</i>	this work
MA11784	<i>yifK[-33A] yifK87::MudK</i>	this work
MA11785	<i>yifK[-33A]</i>	this work
MA11786	<i>yifK[U21C] yifK87::MudK wecG144::cat ΔgcvB141::scar</i>	this work
MA11787	<i>yifK[G27A] yifK87::MudK wecG144::cat ΔgcvB141::scar</i>	this work
MA11788	<i>yifK[U21C] yifK87::MudK wecG144::cat hfq13::Tn5-T-POP</i>	this work
MA11789	<i>yifK[G27A] yifK87::MudK wecG144::cat hfq13::Tn5-T-POP</i>	this work
MA11790	<i>yifK[A46U,A47G,A48U] yifK87::MudK</i>	this work
MA11791	<i>yifK87::MudK gcvB[U86A,U87C,U88A] ygdI143::cat</i>	this work
MA11792	<i>yifK[A46U,A47G,A48U] yifK87::MudK gcvB[U86A,U87C,U88A] ygdI143::cat</i>	this work
MA11793	<i>yifK[C49U,A50C] yifK87::MudK</i>	this work
MA11794	<i>yifK87::MudK gcvB[U84G,G85A] ygdI143::cat</i>	this work
MA11795	<i>yifK[C49U,A50C] yifK87::MudK gcvB[U84G,G85A] ygdI143::cat</i>	this work
MA11796	<i>yifK[C49U,A50C] yifK87::MudK ΔgcvB141::cat</i>	this work
MA11797	<i>yifK[C49U,A50C] yifK87::MudK Δhfq67::cat</i>	this work
MA11798	<i>yifK[-33A]ΔyifK146::cat-3xFLAG</i>	this work
MA11799	<i>yifK[-33A] yifK[C49U,A50C] ΔyifK146::cat-3xFLAG</i>	this work
MA11800	<i>ΔgcvB141::scar [rluC-rne]IG::cat rne-3071 (ts)</i>	this work
MA11801	<i>yifK[-33A] ΔgcvB141::cat</i>	this work
MA11802	<i>yifK[-33A] [rluC-rne]IG::cat rne-3071 (ts)</i>	this work

MA11792	<i>yifK</i> [-33A] <i>hfq13</i> ::Tn5-T-POP	this work
MA11802	<i>yifK</i> [-33A] Δ <i>gcvB142</i> :: <i>aadA</i> [<i>rluC-rne</i>]IG:: <i>cat rne-3071</i> (ts)	this work
MA11803	<i>yifK</i> [-33A] Δ <i>gcvB141</i> :: <i>cat hfq13</i> ::Tn5-T-POP	this work
MA11804	<i>yifK87</i> ::MudK <i>lrp-42</i> ::Tn10dCm	this work
MA11805	<i>yifK87</i> ::MudK Δ <i>gcvB142</i> :: <i>aadA</i>	this work
MA11806	<i>yifK87</i> ::MudK Δ <i>gcvB142</i> :: <i>aadA lrp-42</i> ::Tn10dCm	this work

^a Strains are derived from *Salmonella enterica* serovar Typhimurium strain MA3409 (Figuroa-Bossi et al., 1997) except for strains SV4280 (Camacho & Casadesus, 2002), MA3398 (Figuroa-Bossi et al., 1997) and MA7224 (Uzzau et al., 2001). The term “scar” denotes the DNA sequence left following FLP-mediated excision of antibiotic-resistance cassettes introduced by the procedure of Datsenko and Wanner (Datsenko & Wanner, 2000). Additional details on relevant alleles in this Table can be found in Table S2.

Figures

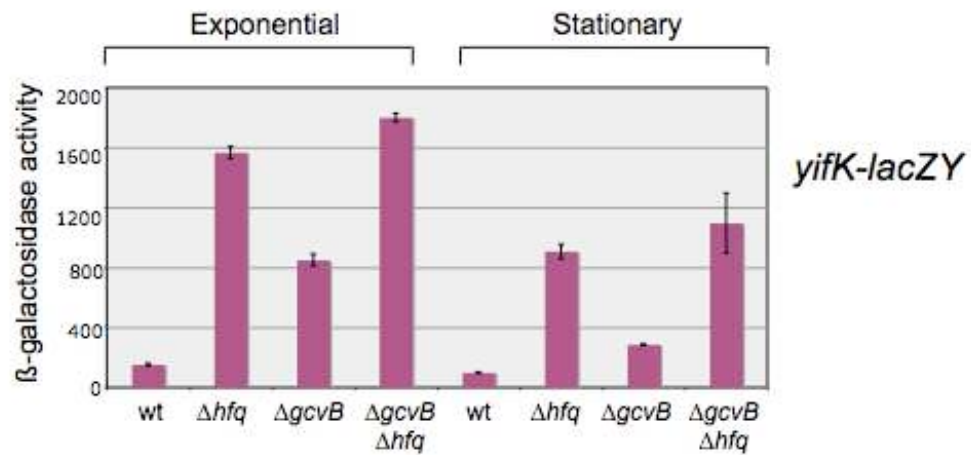


Figure 1. Effect of *hfq* and *gcvB* deletions on the expression of a *yifK-lacZ* fusion. Deleting *gcvB* causes *yifK* expression to increase approximately 5-fold in exponential cultures (OD600 \approx 0.4) and less than three-fold in stationary overnight cultures (OD600 \approx 2.0). A greater increase is observed in the *hfq* deletion mutant, suggesting the involvement of a separate Hfq-dependent step in *yifK* regulation. Strains used were MA8020 (wt), MA8021 (Δhfq), MA10377 ($\Delta gcvB$) and MA10403 ($\Delta hfq \Delta gcvB$). All strains carry the *yifK*::MudK lac fusion. Their full genotypes are listed in Table 1.

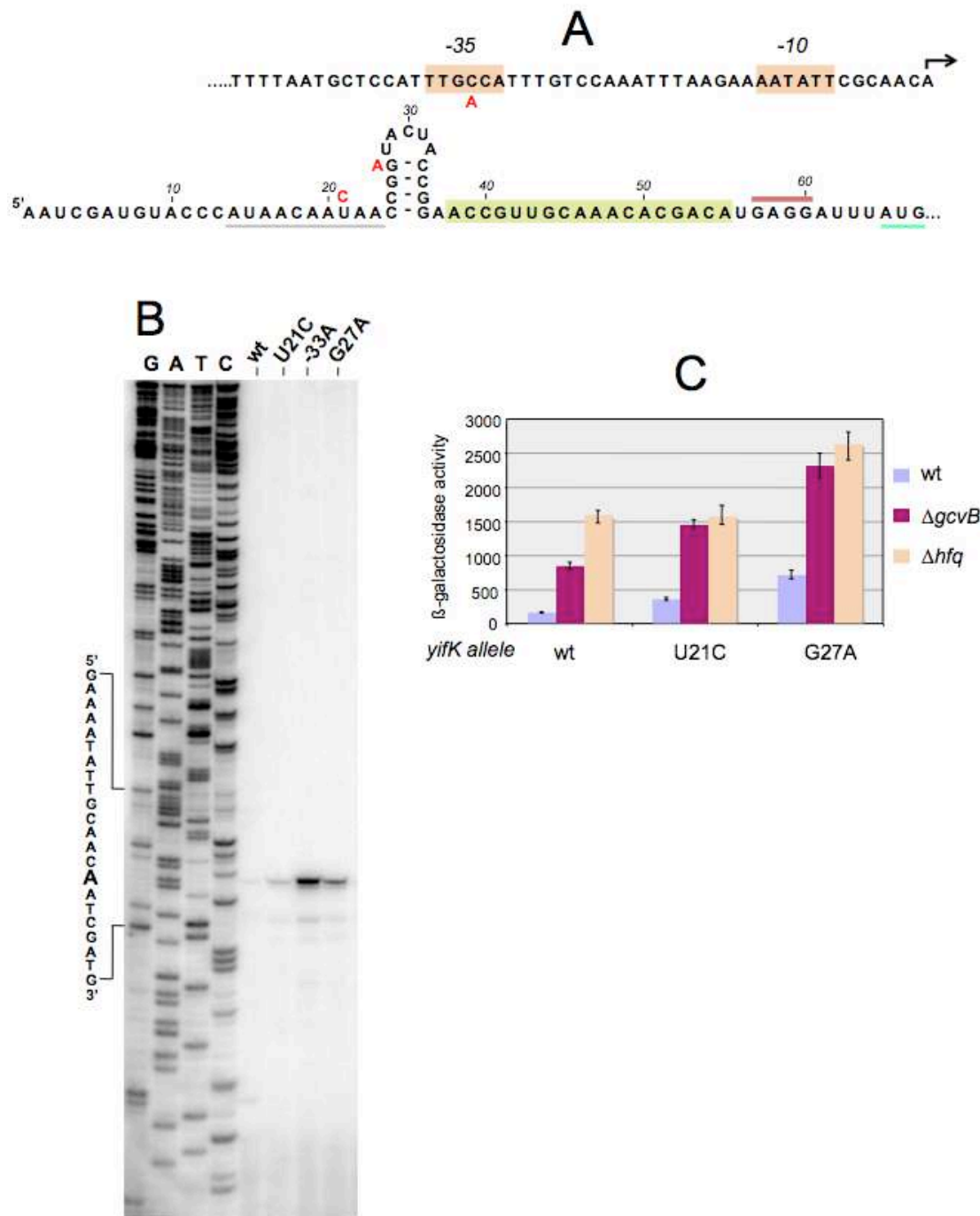


Figure 2. Characterization and mutational analysis of the leader region of the *yifK* gene. The promoter-proximal region of the *yifK* gene was randomly mutagenized by error-prone PCR using primers ppF45 and ppF47 and chromosomal DNA from a strain carrying a *cat* cassette 78 bp upstream from the *yifK* promoter (in opposite orientation; strain MA11780) as template. The mutagenized fragment was introduced into strain MA10280 (*yifK-lacZY* / pKD46) and recombinants were selected as described in the text. Three mutants expressing higher β-galactosidase activity were identified: one carrying a *yifK* promoter change that

causes the -35 box of *yifK* to match the consensus sequence (TTGACA) (A, top); the other two isolates carrying mutations in the 5'UTR (U21C and G27A) (A, bottom). Grey and green underlining denote a putative Hfq binding site and the initiating AUG, respectively. Brown overlining denotes the Shine-Dalgarno sequence. A sequence stretch complementary to GcvB is boxed in light green. B. Primer extension of *yifK* mRNA. Total RNA extracted from wild-type and mutant strains was used to map the 5' end of the *yifK* mRNA by reverse transcription (primer ppF49)(B). This analysis identifies the 5' end of *yifK* mRNA and shows that all three mutations lead to higher mRNA levels. Measurements of β -galactosidase activity (C) show that U21C, but not G27A, does not cause any further increase in *yifK-lacZ* expression in the Δhfq background, suggesting that U21C affects Hfq binding to *yifK* mRNA. β -galactosidase activity was measured in exponentially growing cultures ($OD_{600} \approx 0.4$).

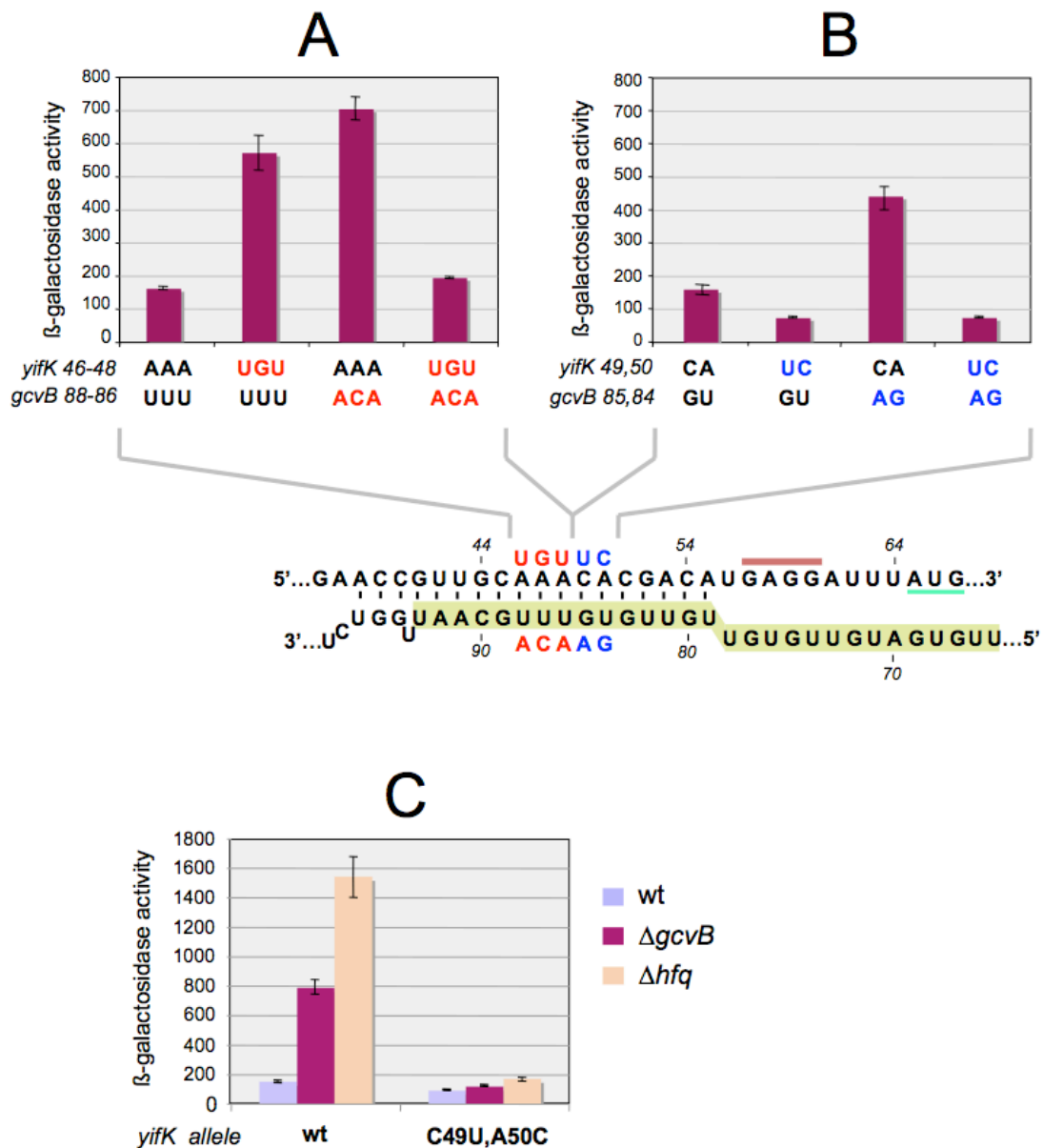


Figure 3. Differential effects of *yifK* target sequence mutations on *yifK-lacZ* expression and on its response to compensatory changes in GcvB. Adjacent portions of the sequence presumed to base-pair with GcvB in *yifK* mRNA were randomly mutagenized by a “scarless” λ *red* recombineering procedure (see Materials and Methods). Briefly, DNA fragments amplified by “reciprocal priming” with oligonucleotide pairs ppG48/ppG49 (mutagenesis of the 46-48 interval) and ppH12/ppH13 (mutagenesis of the 49-50 interval) were introduced into strain MA11526 and tetracycline-sensitive recombinants selected one plates supplemented with fusaric acid (12 μ g/ml). Two of the mutants obtained were chosen for further study. Compensatory changes in GcvB were obtained by standard recombineering using fragments amplified from the chromosome of strain MA11779 with primer pairs

ppG63/ppF18 (mutagenesis of the 86-88 interval) and ppH61/ppF18 (mutagenesis of the 84-85 interval). Results above show that changing *yifK* mRNA sequence from positions +46 to +48, or making the opposite change in GcvB, both relieve repression (A). Repression is restored in a strain carrying the two sets of changes combined, showing that base-pairing is required for repression (A). In contrast, replacing the CA doublet at +49, +50 by UC causes a reduction, rather than an increase, of *yifK-lacZ* expression; introduction of the compensatory changes in GcvB does not accentuate this trend (B). Thus, the CA to UC conversion impairs *yifK* expression and renders it insensitive to GcvB repression. Furthermore, the CA to UC change causes *yifK* to no longer respond to *gcvB* or *hfq* deletions (C).

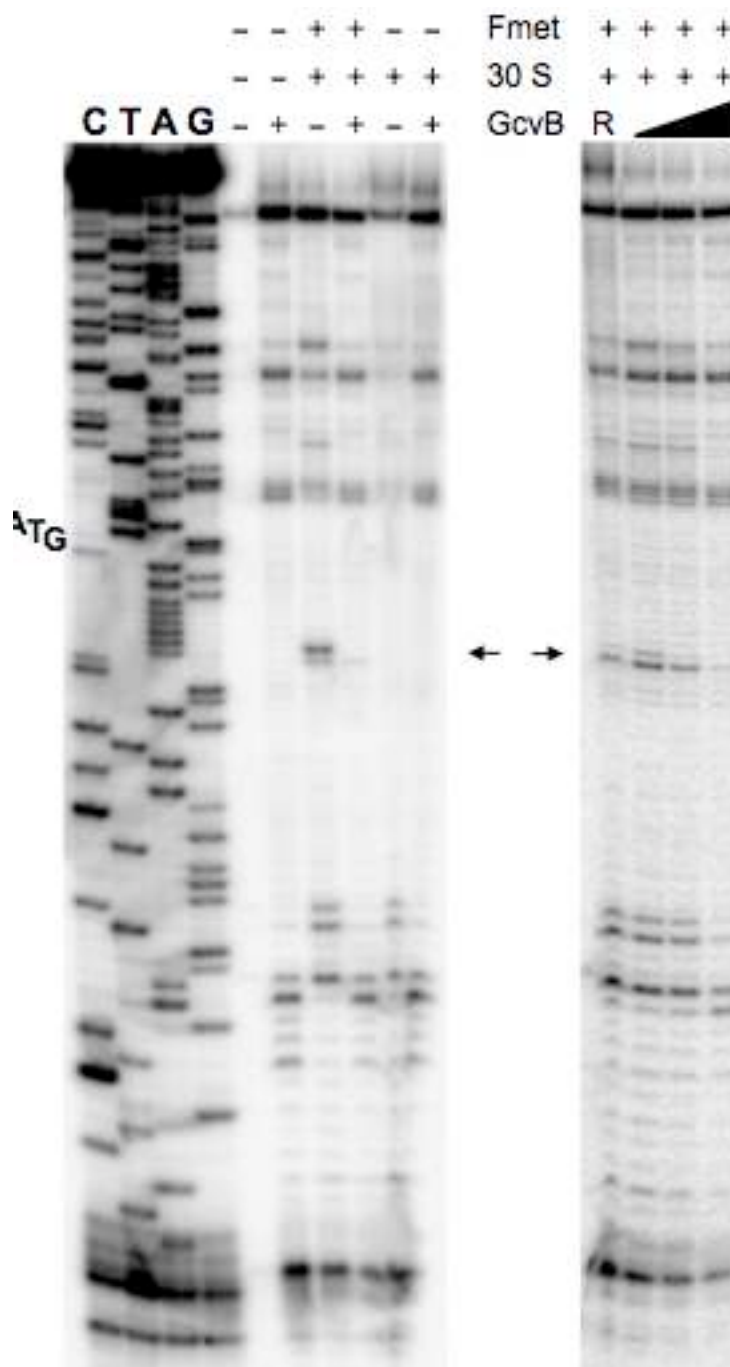


Figure 4. Toeprinting analysis of *yifK* mRNA. 30S ribosomal toeprinting was carried out in the absence or in the presence of GcvB RNA or of RyhB RNA as described in the Materials and Methods. “+” and “-” signs denote the presence or absence of indicated components. The decline and disappearance of the toeprint (shown by arrows) at increasing GcvB concentrations (50 nM, 100 nM and 500 nM), is indicative of interference with the 30S subunit binding to *yifK* mRNA. Failure of RyhB to do so at the concentration of 6.0 μ M (lane “R”) confirms the specificity of the effect.

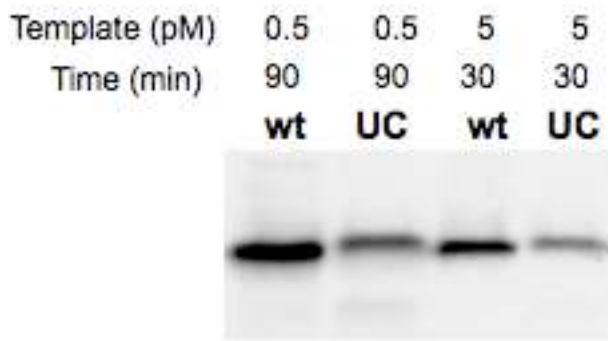


Figure 5. Effect of *yifK* 5' UTR's changes on mRNA translation. *In vitro* translation was performed using the coupled transcription/translation PURExpress kit (see Materials and Methods). DNA templates were from plasmids carrying the entire *yifK* 5' UTR from wild-type, or from the C49U,A50C mutant, fused to the coding sequence of a 3xFLAG epitope-tagged version of the *cat* gene. Fusions were initially obtained as chromosomal constructs using DNA fragments amplified from strain MA7224 with primer pairs ppL50/ppL52 (wt) and ppL51/ppL52 (C49U,A50C). Subsequently, the fusions were cloned into plasmid DHFR following amplification (ppM29 / ppM30) and double Xba I / Pst I digestion. Transcription / translation reactions were carried out at a template DNA concentration of 0.5 pM for 90 min or 5 pM for 30 min and products analyzed by Western blotting using anti-FLAG monoclonal antibodies (Uzzau et al., 2001). Under both conditions, higher amounts of *cat*-3xFLAG protein were synthesized from the construct with the wild-type *yifK* sequence than from the construct harboring the double CA to UC change.

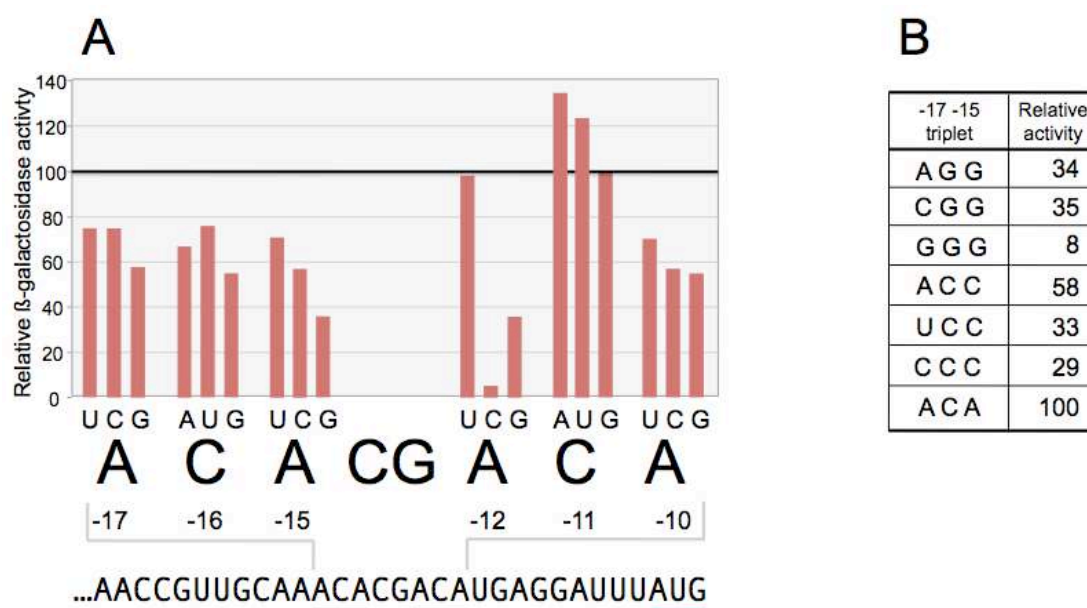


Figure 6. Mutational analysis of ACA motifs. Each base position in two ACA triplets upstream of the *yifK*'s SD sequence was randomized and changes were introduced in the chromosome of strain MA11594 (*yifK-lacZY* Δ *gcvB*) as described in the legend to Fig. 3 (primers used to generate the set of mutagenized fragments are listed in Table S2). Mutants obtained were screened by PCR (primers ppF45/ppF62) and DNA sequencing. All possible replacements (9 variants for each triplet) were identified. These strains were assayed for β -galactosidase activity (A). Typically, duplicate or triplicate β -galactosidase measurements were carried out in parallel for all variants of any given position and the wild-type strain, whose value was set to 100. Standard deviations were less than 5% of the mean in all cases. Data in (A) show that any change in the upstream ACA causes a decrease in *yifK* expression, with G residues exerting the most negative effects. Replacements at the downstream ACA have contrasting effects: replacing the central C with A or U increases expression while having a C at the first position is highly deleterious. Representative examples of triple substitutions in the upstream ACA are also shown (B). β -galactosidase measurements were as in (A).



Figure 7. Increasing the distance between the enhancer and the translation initiation region. Constructs were made as described in the legends to Figs. 3 and 6 using strain MA11594 (*yifK-lacZY* Δ *gcvB*) as recipient and fragments amplified by reciprocal priming of oligonucleotides described in Table S2. A 7-nucleotide segment (boxed in light green), duplicating the SD sequence (overlined in brown), was inserted between the ACA triplets (underlined in grey) and the SD sequence (construct n. 3). A G to C change was then introduced in either copy of the SD (constructs n. 4 and 5; construct n. 2 shows the effect of this change in a strain with a single SD). The upstream ACA triplet was converted to GGG in the constructs carrying either SD sequence mutated (constructs n. 6 and 7). β -galactosidase activity was measured as described in the legend to Fig. 6. The activity of the wild-type strain (construct n. 1) was set to 100. Standard deviations were less than 5% of the mean in all cases. The data (see also strains' phenotypes on MacConkey-lactose plates) show that the upstream ACA maintains its enhancer effect when placed further upstream from the initiation region, independent of the spacing between the SD sequence and the starting AUG.

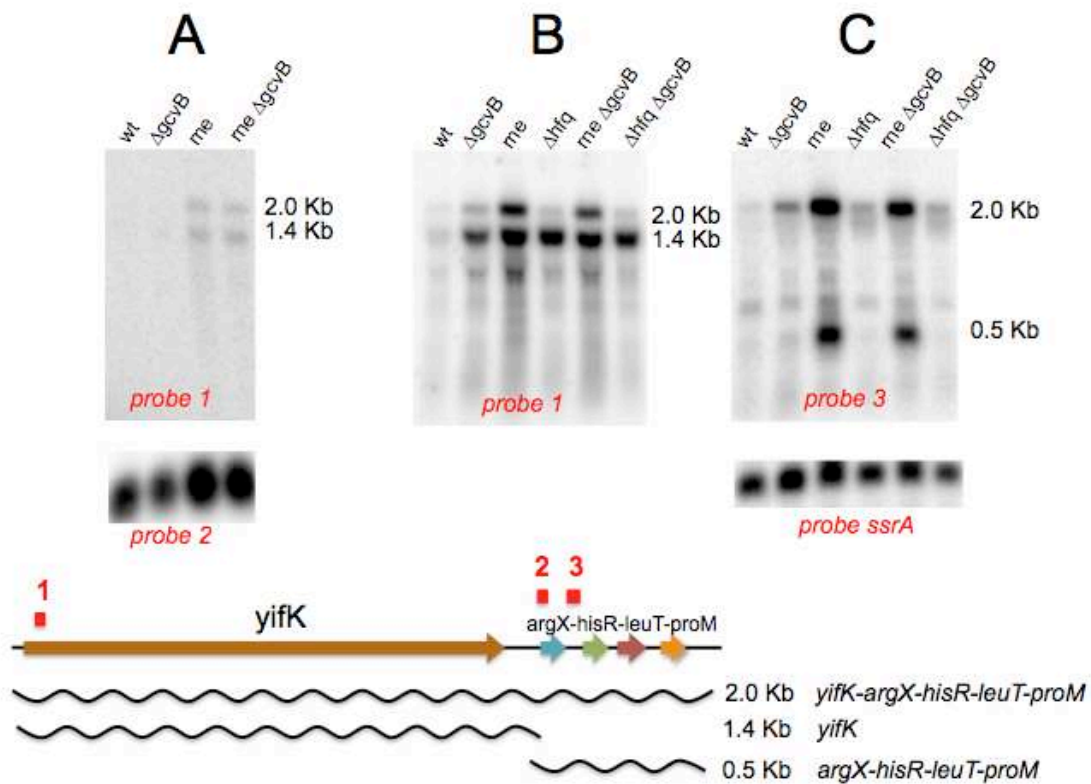


Figure 8. Northern blot analysis of *yifK* transcription. RNA was extracted from strains carrying the wild-type *yifK* promoter (A) or the -33 promoter “up” mutation (B and C). RNA was separated on a 1% agarose-formaldehyde gel and probed with ^{32}P -labeled DNA oligonucleotides complementary to a sequence near the 5' end of *yifK* mRNA (ppF16; probe 1 above) or to an internal portion of the *argX* tRNA (ppG42, probe 2), or to a sequence in the *argX-hisR* intercistronic region (ppH27, probe 3). Probing for the SsrA RNA (pp813) served as loading control. The strains used as source of RNA are listed in Table 1. DNA oligonucleotides used as hybridization probes are listed in Table S1.

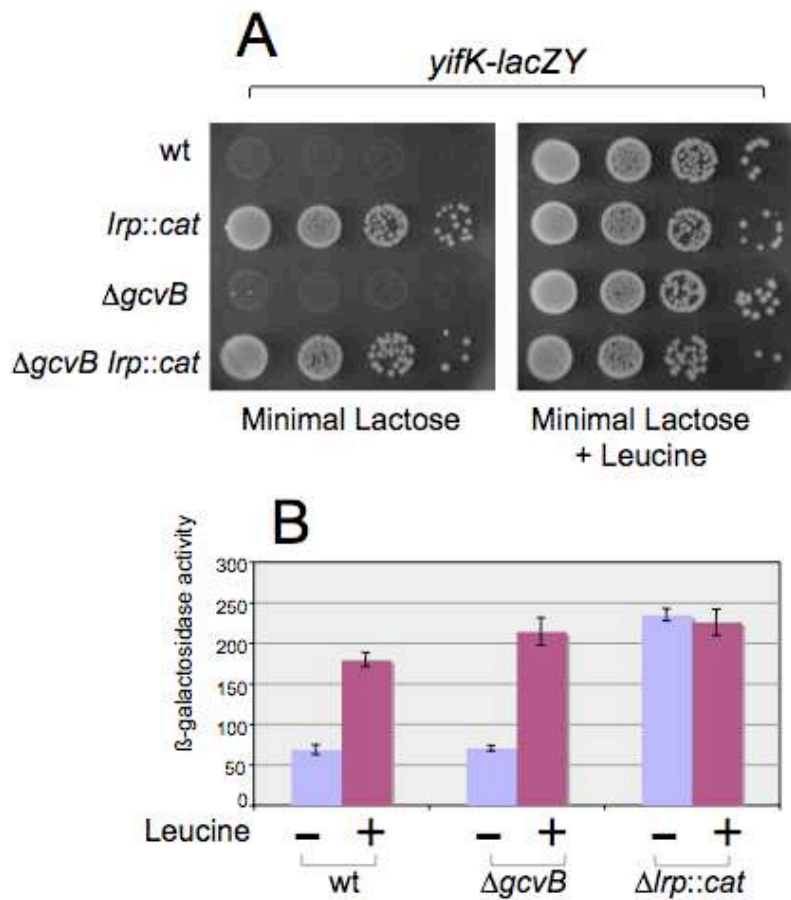


Figure 9. Lrp control of *yifK-lacZY* expression. The *yifK* gene is expressed at very low levels in minimal medium. As a result, strains carrying the *yifK-lacZY* fusion are phenotypically Lac⁻ in this medium, regardless of the *gcvB* allele (A, left panel). Inactivation of Lrp confers a Lac⁺ phenotype (A, left panel). This suggests that low *yifK* expression results from Lrp repression. The Lac⁺ phenotype is also restored upon addition of leucine (0.3 mM) (A, right panel), indicating that leucine relieves Lrp repression. β -galactosidase measurements (B) confirm that GcvB plays no significant role in *yifK* regulation in minimal medium and provide a quantitative estimate of the Lrp effects. NCE medium (Maloy & Roth, 1983) supplemented with 0.2% lactose (A) or 0.2% glycerol (B) was used as minimal medium.

II.2.1. Supporting information of manuscript

The translation enhancer activity of an ACA motif dictates the response of a *Salmonella* gene to multi-target small RNA regulator GcvB

Qi Yang, Nara Figueroa-Bossi and Lionello Bossi*

Centre de Génétique Moléculaire, CNRS, Gif-sur-Yvette, France

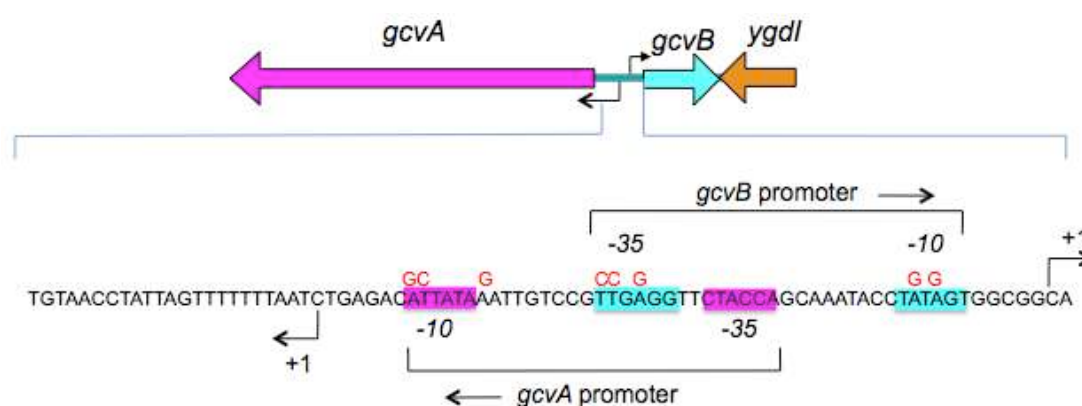


Figure S1. *gcvB*-linked mutations relieving *yifK* repression. A DNA fragment spanning the *gcvB* gene and a linked *cat* marker (*ygdI::cat*; placed 71 bp downstream in a parallel orientation) was amplified by PCR under error-prone conditions using oligonucleotides ppF17 and ppF18 as primers (Table S1) and chromosomal DNA from strain MA1179 (Table 1) as template. The amplified fragment was introduced into strain MA10280 (*yifK-lacZY* / pKD46) and recombinants were selected on MacConkey lactose plates supplemented with chloramphenicol as described in the text. Red colored colonies were picked and the region of the *gcvB* locus analyzed by DNA sequencing. Most of the isolates were found to harbor DNA sequence changes in the *gcvA-gcvB* intergenic region, which affected either the -35 or -10 box of the *gcvB* promoter, or the -10 box of the *gcvA* promoter [1]. Mutations within the initial portion of the *gcvA* coding sequence and a mutation affecting the CG-rich stem of *gcvB*'s Rho-dependent terminator were also identified (data not shown).

1. Urbanowski ML, Stauffer LT, Stauffer GV (2000) Mol Microbiol 37: 856-868.

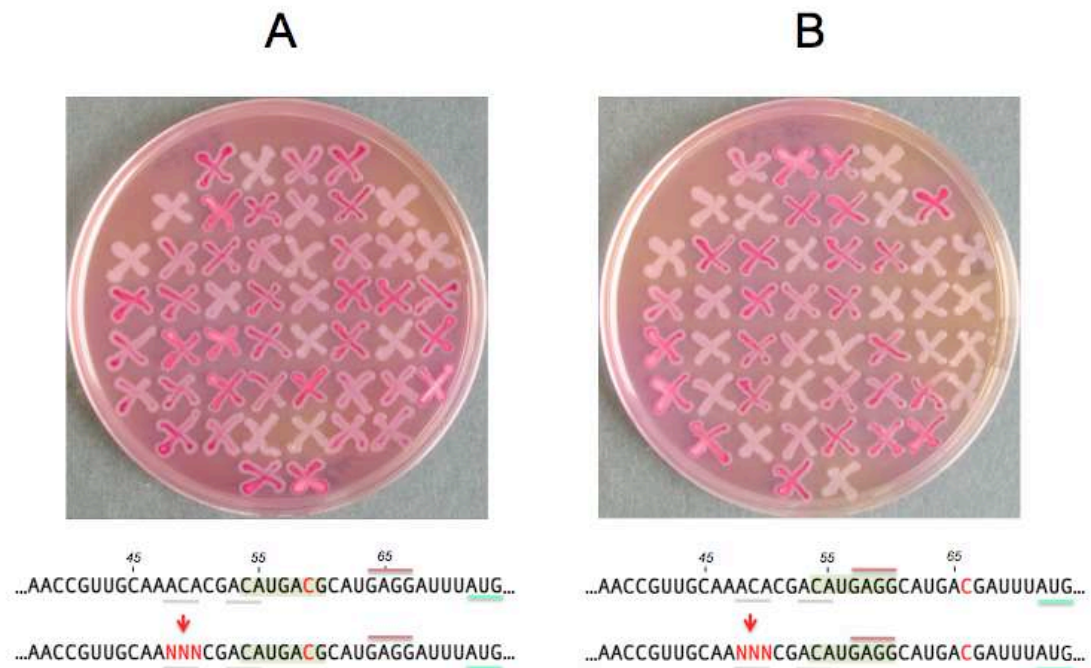


Figure S2. Effects of replacing the ACA motif at positions 48-50 by random triplets. The sequence of the 48-50 interval was randomized in strains carrying a 7 nt tandem direct duplication of the SD region with either the AUG-distal SD sequence inactivated (A) or the AUG-proximal SD sequence inactivated (B). Mutagenesis involved introducing DNA fragments amplified by reciprocal priming of DNA oligonucleotides ppN76 and ppN77 (A) or ppN78 and ppN79 (B) into strain MA11594 (*yifK::tetAR yifK-lacZY ΔgcvB*) selecting fusaric acid resistance (see text) and patching recombinants onto MacConkey lactose plates. A sample of 50 recombinants from each cross is shown. For more details on alleles and primer used, see Tables S1 and S2.

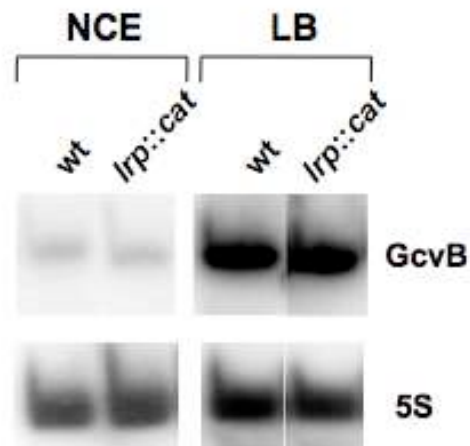


Figure S3. Comparing GcvB sRNA levels in wild-type and in an *Irp* insertion mutant as a function of the growth medium. Bacteria were grown in minimal medium (NCE [1]) supplemented with 0.2 % glycerol or in LB to an $OD_{600} \approx 0.4$. RNA was extracted, fractionated on an 8% polyacrylamide-8 M urea gel and subjected to Northern blot hybridization. Blot was hybridized to DNA oligonucleotides complementary to GcvB and to 5S RNA (for loading control). The probes used were ppl67 (GcvB) and ppB10 (5S)(Table S1).

1. Maloy SR, Roth JR (1983) J Bacteriol 154: 561-568

Table S1. DNA oligonucleotides used in this work

name	Sequence (5'-3') ^a
pp813	GCGGAGGCTAGGGAGAGAGG
ppB10	ACACTACCATCGGCGCTACG
ppF08	ATTGTCCGTTGAGGTTCTACCAGCAAATACCTATAGTGGCCCATGGTCCATATGAATATC
ppF09	TACTGACGTGAAAGAGATGGTGGCTCTGGAAAATAACCGTGTGTAGGCTGGAGCTGCTT
ppF10	ATTGTCCGTTGAGGTTCTACCAGCAAATACCTATAGTGGCACCTTGCCGTAGAAGAACAG
ppF11	TACTGACGTGAAAGAGATGGTGGCTCTGGAAAATAACCGTTTGGCTGTGAGCAATTATG
ppF16	GGCAATCAATTCAATATGACGAGCTTCCAGCCCACGC
ppF17	CATCTCTTTCACGTCAGTACGATTGATCTGCTGTTTGTGCCATGGTCCATATGAATATC
ppF18	GTCCTAATTACGTTATGCACACCAATGATGGACGTAGCATTGTGTAGGCTGGAGCTGCTT
ppF19	TGCTCGCGTAAAGCTCAAAT
ppF20	CGAAACTGGCTACGATGAA
ppF43	TATCTTCGCTGGCACTATACTGGCGATCTCTAATTCTCCCTGTGTAGGCTGGAGCTGCTT
ppF44	CATCGATTATGTTGTAATGAGCAGCACCATAAGCACAAATCCATGGTCCATATGAATATC
ppF45	CGTGGTGC GTTCAATACGTA
ppF46	CGCCATCCACATAAACCAGT
ppF49	TCTCTGCCATAAATCCTCAT
ppF62	CGCCATTTTCAGCGTACT
ppG42	AGGGCAGCGCTCTATCCAGCTGAGCTACGGGCGCTTAG
ppG44	AATCGATGTACCCATAACAATAACCGGTA CTACCGGAACCTTAAGACCCACTTTCACATT
ppG45	ACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATCTAAGCACTTGTCTCCTG
ppG48	TACCCATAACAATAACCGGTA CTACCGGAACCGTTGCNNNCACGACATGAGGATTTATGGCAG
ppG49	CGAGCTTCCAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATGTGCTG
ppG51	TATGACGAGCTTCCAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATG
ppG63	TTCTGATGGGCTTTTGGCTTACGGTTGTGATGTTGTGTTGTTGTGACAGCAATTGGTCTGCGATT
ppH12	TCGATGTACCCATAACAATAACCGGTA CTACCGGAACCGTTGCAANNCGACATGAGGATTTATGGCAG
ppH13	TATGACGAGCTTCCAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATGTGCTG
ppH27	TGAAACTTTTTACGCGGTATTAACCACCGCAGCTCAAGCACCTAAATAAA
ppH61	CTGATGGGCTTTTGGCTTACGGTTGTGATGTTGTGTTGTTGGAATTTGCAATTGGTCTGCGATT
ppH81	CGATGATTTCTTACACAATAAGTGCATTTTTTTAATGCTCCATTTGACATTTGTCCAAATTTAAG
ppI22	GGTAATACGACTCACTATAGGGCAATCGATGTACCCATAACAATAACCG
ppI23	CCCGAGGGCAATCAATTCAA
ppI67	AGACCAATTGCAAACACAACAACAACATCACAACCGTAAGCCA
ppL50	TACTACCGGAACCGTTGCAAACACGACATGAGGATTTATGGAGAAAAAATCACTGGATATACCA
ppL51	TACTACCGGAACCGTTGCAAAATCCGACATGAGGATTTATGGAGAAAAAATCACTGGATATACC
ppL52	AAATGCGGTTTTGTTATGCTTTGCGCATTTTGGCGCTTACTATTTATCGTCGCATCTTTGTAG
ppL95	AATCGATGTACCCATAACAATAACCGGTA CTACCGGAACCGTTGCAA
ppL96	ACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATGTGCGNNNTTGCAACGGTCCGGTAGTAC

Table S1, continued

name	Sequence (5'-3') ^a
ppL97	AATCGATGTACCCATAACAATAACCGGTACTACCGGAACCGTTGCAAACACG
ppL98	CACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTC ANN CGTGTGGCAACGGTCCGGT
ppM29	AATTTATATTTAACTATTCTAGACAATCGATGTACCCATAACAATAACC
ppM30	ATTTATATTTTCGGCCGCTCGAGCTGCAGTGGCGCTTACTATTTATCGTCG
ppM80	ACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATGTCGT V TTGCAACGGTCCGGTAGTAC
ppM81	ACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATGTCGT H TTTGAACGGTCCGGTAGTAC
ppM82	ACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATGTCG V TTTGAACGGTCCGGTAGTAC
ppM83	CACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATG V CGTGTGGCAACGGTCCGGT
ppM84	CACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCAT H TCGTGTGGCAACGGTCCGGT
ppM85	CACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTC V GTCGTGTGGCAACGGTCCGGT
ppN58	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATG <u>CCTCATG</u> TCGTGTGGCAACGGTCCG
ppN59	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATC G TCATG <u>CCTCATG</u> TCGTGTGGCAACGGTCCG
ppN60	AATCGATGTACCCATAACAATAACCGGTACTACCGGAACCGTTGCAAACACGACA
ppN63	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATC G TCATGTCGTGTGGCAACGGTCCG
ppN64	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATG <u>CGTCATG</u> TCGTGTGGCAACGGTCCG
ppN73	AATCGATGTACCCATAACAATAACCGGTACTACCGGAACCGTTGCA GGG CGACATG
ppN74	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATG <u>CGTCATG</u> TCG CCC TTGCAACGGTTC
ppN75	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATC G TCATG <u>CCTCATG</u> TCG CCC TTGCAACGGTTC
ppN76	TGTACCCATAACAATAACCGGTACTACCGGAACCGTTGCA ANN CGA <u>CATGACG</u> CATGAGGATTTATGG
ppN77	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATG G TCATGTCG
ppN78	TGTACCCATAACAATAACCGGTACTACCGGAACCGTTGCA ANN CGA <u>CATGAGG</u> CATG C GATTTATGG
ppN79	CAGCCCACGCTGTAGCTCCGGTTTTTCTCTGCCATAAATC G TCATGCCTCATGTCG

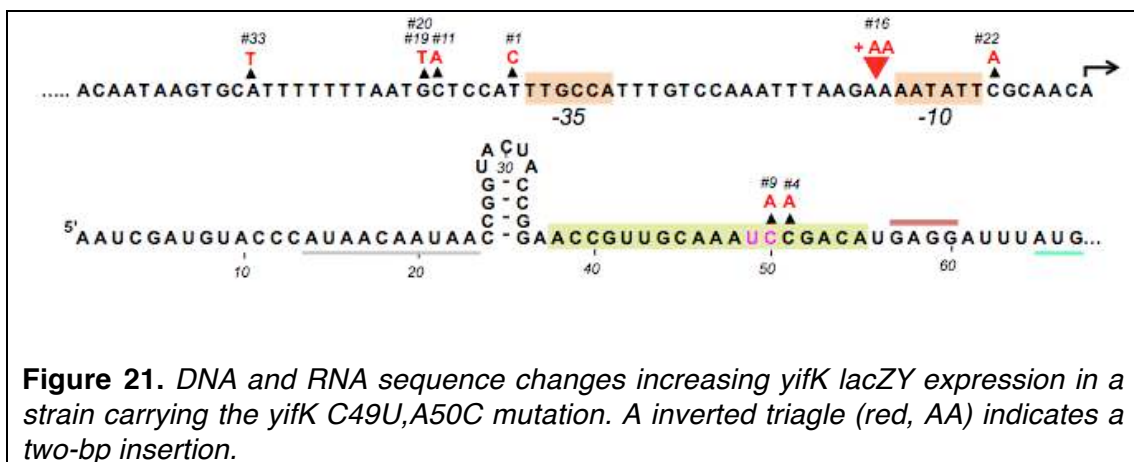
^a Changes from the wild-type sequence are in red. Nucleotide insertions are in bold and underlined. “N” denotes an equimolar mixture of all four nucleotides, “V” denotes an equimolar mixture of A, C and G; “H” denotes an equimolar mixture of A, T and C.

Table S2. Relevant alleles constructed in this work

Primer pair	Template ^a	Allele	Description
ppF08 / ppF09	pKD3	$\Delta gcvB141::cat$	<i>cat</i> gene in place of <i>gcvB</i>
ppF10 / ppF12	pSEB5	$\Delta gcvB142::aadA$	Spec ^R gene in place of <i>gcvB</i>
ppF17 / ppF18	pKD3	<i>ygdI143::cat</i>	<i>cat</i> on 3' side of <i>gcvB</i> (parallel orientation)
ppF43 / ppF44	pKD3	<i>wecG144::cat</i>	<i>cat</i> on 5' side of <i>yifK</i> (opposite orientation)
ppG44 / ppG45	MA3397	<i>yifK145::tetAR</i>	<i>tetAR</i> insertion 10 bp upstream of <i>yifK</i> AUG
ppG48 / ppG49	self	<i>yifK</i> [46-48 UGU]	Replacement of <i>yifK</i> mRNA segment between +46 and +48 (AAA to UGU)
ppG63 / ppF18	MA11779	<i>gcvB</i> [86-88 ACA]	Replacement of GcvB segment between +86 and +88 (UUU to ACA)
ppH12/ ppH13	self	<i>yifK</i> [49-50 UC]	Replacement of <i>yifK</i> mRNA segment between +49 and +50 (CA to UC)
ppH61 / ppF18	MA11779	<i>gcvB</i> [84-85 GA]	Replacement of GcvB segment between +84 and +85 (UG to GA)
ppH81 / ppG51	MA3409	<i>yifK</i> [-33A]	<i>yifK</i> promoter "up" mutation (G to A at -33) in wild-type background
ppL50 / ppL52	MA7224	<i>yifK::cat-3xFLAG</i>	Replacement of <i>yifK</i> orf by <i>cat-3xFLAG</i> orf
ppL51 / ppL52	MA7224	<i>yifK</i> [49-50 UC]:: <i>cat-3xFLAG</i>	Replacement of <i>yifK</i> orf by <i>cat-3xFLAG</i> orf in the C49U,A50C mutant
ppL95 / ppL96	self	<i>yifK</i> [48-50 NNN]	Randomized sequence in 48-50 segment of <i>yifK</i> mRNA (AUG-distal ACA)
ppL97 / ppL98	self	<i>yifK</i> [53-55 NNN]	Randomized sequence in 53-55 segment of <i>yifK</i> mRNA (AUG-proximal ACA)
ppL95 / ppM80	self	<i>yifK</i> [48 V]	C, U, or G at position 48 of <i>yifK</i> mRNA
ppL95 / ppM81	self	<i>yifK</i> [49 H]	A, U, or G at position 49 of <i>yifK</i> mRNA
ppL95 / ppM82	self	<i>yifK</i> [50 V]	C, U, or G at position 50 of <i>yifK</i> mRNA

II.3. An intriguing regulator of *yifK* expression

Some of *yifK*'s translation enhancer mutations lower the expression of a *yifK-lacZY* fusion to the extent that the strain becomes Lac⁻ even if deleted for the *lrp* gene (for Lrp effects on *yifK* see section II.2). We considered that possible factors mediating the enhancer activity could be identified through the analysis of suppressor mutations restoring a Lac⁺ phenotype. Strain YQ114 (*yifK* C49U,A50C::MudK *lrp*::*cat*; was used to select spontaneous mutants capable of using lactose as sole carbon source (note: the MudK element in YQ114 includes the *lacZY* operon plus a kanamycin-resistance gene; see II.1). A total of 36 Lac⁺ colonies were picked and genetically characterized. Anticipating that a fraction of these mutants might contain from cis-acting alleles, a preliminary test was performed to identify mutations genetically linked to the *yifK-lacZY* fusion. For this, P22 transducing lysates were made on all mutants and used to transduce strain 8920 selecting kanamycin resistance (Kan^R). Transductant colonies were replica-plated onto NCE-lactose plates. With 9 of the 36 lysates more than 95% of Kan^R transductants were Lac⁺ indicating that the mutation conferring the Lac⁺ phenotype was genetically linked to *yifK*. These mutants were characterized further by amplifying and sequencing the region at the 5' end of the *yifK-lacZY* fusion. As expected all of the mutants carried sequence changes in this region (Fig. 21).



In isolates #1, #11, #16, #19, #20, #22 and #33 the change is found upstream of the transcription start site, strongly suggesting that the mutations do not actually correct the translation initiation defect; rather they compensate for it by increasing transcription levels. Intriguingly, five of these mutants have a sequence change in the region upstream from the -35 box of the *yifK* promoter. These mutants will be further discussed below. Only two mutations appear to act by relieving the translation defect. The change in isolate #9 is a partial reversion of the starting C49U,A50C allele (A50 restored) whereas the C51A change in isolate #4 (C51A) must somehow relieve the defect associated with the C49U,A50C allele.

With the lysates from the remaining 27 mutants, the vast majority of the Kan^R transductants remained Lac⁻, indicating that the mutation causing the Lac⁺ phenotype in the donor strain was not near to the *yifK* locus. Still, a small fraction of transductants (between 2 and 4 % of the total) were Lac⁺ suggesting that the Lac⁺ mutation was in the same transduction interval as the *yifK*::MudK insertion, albeit distant from this site. We therefore performed additional mapping experiments tests using other genetic markers in the region that were available in our laboratory. This work revealed the mutation to be approximately 50% linked to a Tn10 insertion in the *ilvA* gene. To map the mutation more precisely we sought additional markers from other laboratories. One such marker, obtained from the laboratory of Prof. Casadesús at the University of Seville, Spain, was a *cat* insertion in the *yifA* gene (also named *hdfR*) about 10 Kb counterclockwise from *ilvA*. Much to our surprise, when we used the *hdfR*::*cat* insertion in transduction experiments, we discovered that this insertion by itself conferred a Lac⁺ phenotype to a strain with *yifK* C49U,A50C::MudK and Δlrp . Postulating that our Lac⁺ mutations could be in this gene, we sequenced the entire *hdfR* gene from two of the mutants isolates (#2 and #14). This analyses showed no *hdfR* sequence changes in either strain. However, sequencing runs that extended to the region upstream from the *hdfR* promoter revealed the presence of mutations in the adjacent gene, *yifE*, oriented in the opposite direction from *hdfR* (Fig. 22).

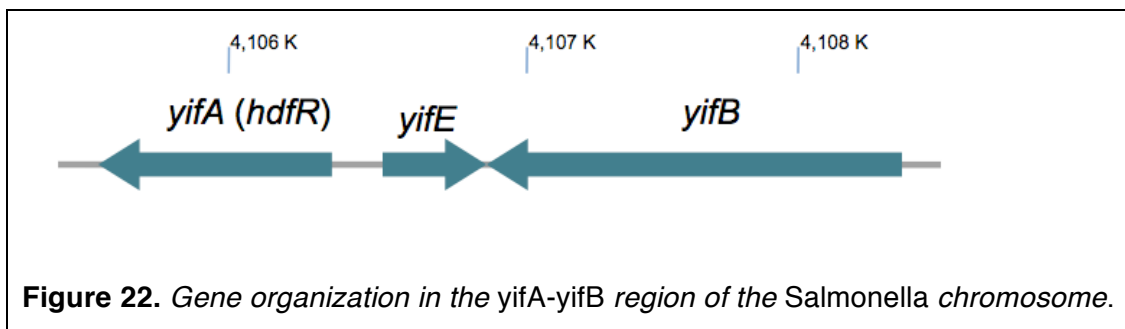


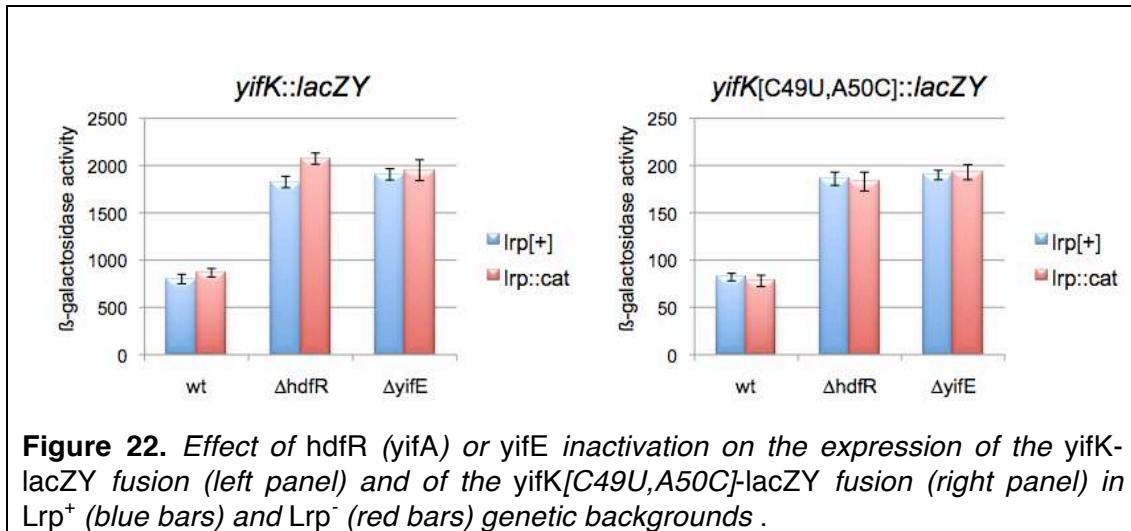
Figure 22. Gene organization in the *yifA*-*yifB* region of the *Salmonella* chromosome.

Isolate #2 carried a 10 bp deletion approximately in the middle of the 336 nucleotide *yifE* orf, while isolate #14 had a nonsense mutation in the 15th codon (AAA to UAA). To confirm these findings, new *hdfR* (*yifA*) and *yifE* mutants were constructed by the lambda Red method. Introduction of either of these constructs ($\Delta hdfR::tetAR$ or $\Delta yifE::tetAR$) into strain YQ114 (*yifK* C49U,A50C::MudK *Irp*::*cat*) caused the strain to become Lac⁺.

The *hdfR* gene encodes a poorly characterized LysR-type transcriptional regulator originally identified in *E.coli* as a repressor of the flagellar master operon *flhDC* (Ko & Park, 2000). Purified HdfR protein was shown to bind to the *flhDC* promoter region and the *hdfR* gene to be itself negatively regulated by H-NS (Ko & Park, 2000). In *Salmonella*, *hdfR* was identified a locus responsible for the high-level expression of the *std* fimbrial operon in *dam* methylase mutants (Jakomin *et al.*, 2008). In this case, HdfR functioned as an activator of transcripton, although it appeared to do so only when the *std* promoter DNA was unmethylated (Jakomin *et al.*, 2008). To date the environmental signals HdfR modulating expression and/or activity of HdfR remain unknown.

The *yifE* gene encodes a small protein (112 aa) highly conserved in the bacterial kingdom, but whose function is currently unknown.

The roles of HdfR and YifE in *yifK* expression were further characterized measuring the effects of the *hdfR* or *yifE* insertions on *yifK* expression in various genetic backgrounds in cells growing exponentially in LB medium (Fig. 23).



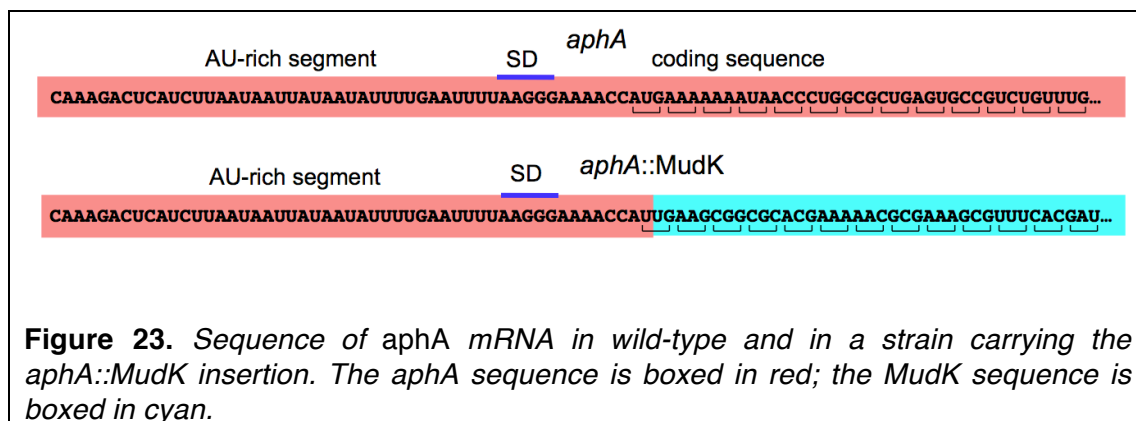
The results in Fig. 22 allow some important points to be made: *i*) inactivation of either gene causes *yifK-lacZY* expression to increase regardless of whether *yifK* 5'UTR is normal or contains the C49U,A50C mutation. In other words, the effects of the *hdfR* or *yifE* are not specific to the C49U,A50C allele. Although constructs with the C49U,A50C mutation have an about 10-fold lower overall activity, the derepression ratios in the presence or absence of HdfR or YifE are the same as in *yifK* wild-type; *ii*) HdfR/YifE-mediated regulation is active in cells growing in rich medium (LB). This is unlike what observed with Lrp regulation which only operates in minimal medium (note the absence of Lrp effects above). In fact, HdfR/YifE-mediated regulation appears to be completely independent of Lrp regulation.

In conclusion, these data suggest that *hdfR* and *yifE* gene products act in concert to repress *yifK* expression. The most likely scenario is that repression takes place at the transcriptional level and that the active repressor is a heterodimer containing HdfR and YifE polypeptides. This scenario can incorporate the finding that the majority of cis-acting mutations from the Lac⁺ selection (above) map on the 5' side of *yifK* promoter -35 box (Fig. 21, above). This could be the site where the repressor binds and the mutations in Fig. 21 could affect the binding. However, alternative scenarios are also possible. For example, either HdfR or YifE could act alone in repressing *yifK* and the other protein be required for expression of the repressor. Since YifE shows no similarity to known transcriptional regulators,

this includes the possibility that YifE acts post-transcriptionally, and HfdR is required for *yifE* transcription. Work aimed at characterizing this system is currently under way in the laboratory.

II.4. A specialized mutant ribosome improves initiation from a UUG start codon

One of the MudK inserts identified in the original screen for Hfq-regulated *lac* fusions was in the *aphA* gene (Figueroa-Bossi et al., 2006a). This gene encodes a periplasmic acid phosphatase involved in the uptake of nicotinamide mononucleotide (Grose et al., 2005). The MudK insert position is unusual as it is inside the AUG initiation codon of *aphA*, between the second and third nucleotide position. The insert generates a UUG codon in-frame with the *lacZ* coding sequence, which appears to function as initiation codon (Fig. 23).



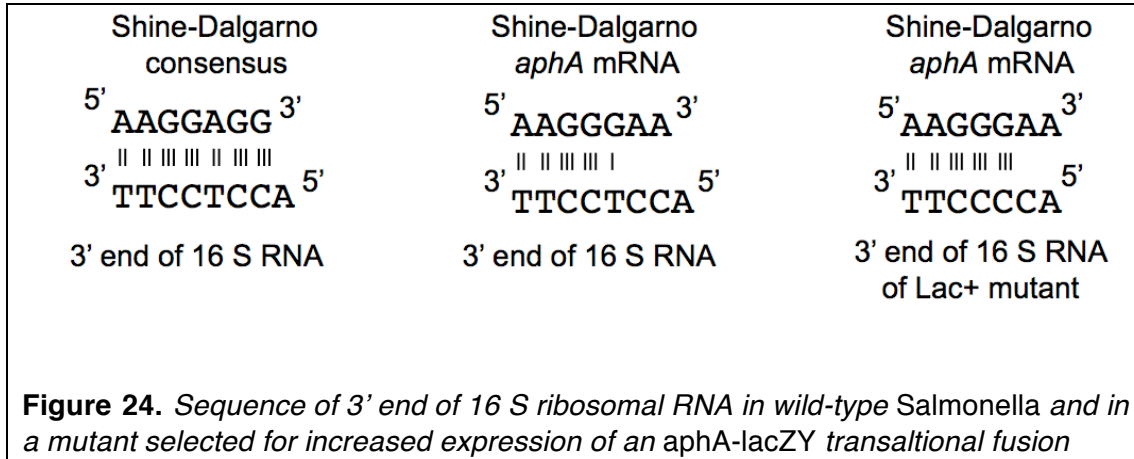
The *aphA-lacZ* fusion is poorly expressed - about 20 Miller units of β -galactosidase in exponentially growing cells - and the expression levels increase 3.5-fold upon inactivating *hfq* (Figueroa-Bossi et al., 2006a). These findings, together with the presence of a long AU-rich sequence that could constitute a Hfq binding site upstream from the Shine-Dalgarno motif (Fig. 23), suggested that a Hfq-dependent regulatory sRNA might contribute to the poor expression of the *aphA-lacZY* fusion.

The low β -galactosidase activity does not allow growth on lactose as sole carbon source (strain is phenotypically Lac⁻). However, the 3.5-fold

increase that results from the inactivation of Hfq is sufficient to restore growth on lactose (Lac⁺). This suggested that a *lac*-based selection could be used to analyze the possible participation of a small RNA in *aphA* repression. We postulated that mutations inactivating the hypothetical sRNA regulator would have effects similar to that of Hfq inactivation and confer a Lac⁺ phenotype. I therefore used the strain carrying *aphA-lacZY* to select spontaneous mutants forming colonies on NCE-lactose minimal plates. A total of 36 mutants were picked and genetically characterized. In all of them, except one, the mutation conferring Lac⁺ mapped very close to *aphA*. The unique unlinked mutation appeared as the best candidate for an sRNA mutant. I therefore undertook its characterization. Using a P22 transducing lysate made on a pool of random T-POP (Tet^R) transposon insertions in the *Salmonella* chromosome (Rappleye & Roth, 1997), I was able to isolate an insertion closely linked to the site of the Lac⁺ mutation. I then used the technique of inverse PCR (Hartl & Ochman, 1996) to amplify and sequence the region adjacent to the insertion site. This analysis showed the T-POP element to be in close proximity of the 3' end of a 16 S ribosomal RNA gene. The sequencing revealed that the Lac⁺ mutation caused the deletion of the U residue in the middle of anti-Shine-Dalgarno sequence of 16S RNA, changing the sequence from CCUCCU to CCCC. To confirm that this mutation was responsible for the Lac⁺ phenotype, I prepared a P22 lysate on the Lac⁺ mutant with the T-POP insertion and transduced the *aphA-lacZY* strain (Lac⁻) selecting Tet^R. More than 95% of the Tet^R transductants became Lac⁺, conclusively demonstrating the role of the 16 S RNA change in this phenotype.

In conclusion, this analysis did not allowed me to confirm the existence of *aphA*-regulating sRNA(s). Nonetheless, the study provided an interesting example of ribosomal "specialisation" (Hui & de Boer, 1987). We interpret our findings postulating that the poor expression of the *aphA-lacZY* fusion is due to an imperfect interaction between the 3' end of the 16 RNA and the Shine-Dalgarno sequence (see Fig. 24), combined to the presence of a weak initiation codon. Deleting a U residue in the anti-Shine Dalgarno sequence results in the replacement of a G-U base pair by a G:C base pair (Fig. 24).

This could strengthen the interaction enough to positively affect translation initiation rates and increase β -galactosidase synthesis.



Like *Escherichia coli*, *Salmonella enterica* contains 7 copies of the 16 S RNA gene. From examination of the sequence adjacent to the mutant site in our Lac⁺ strain, we can conclude that the mutation is located in either *rrsA* or *rrsB* loci.

II.5. Article II

Lack of interchangeability of Hfq-like proteins

article by Tatiana Rochat¹, Philippe Bouloc¹, Qi Yang², Lionello Bossi² and Nara Figueroa-Bossi^{2*}

¹ Institut de Génétique et Microbiologie, CNRS, UMR 8621, IFR115, Université Paris-Sud, Orsay, France

² Centre de Génétique Moléculaire, CNRS, UPR3404, IFR115, Université Paris-Sud, Gif-sur-Yvette, France

Published in *Biochimie* 94: 1554-1559, 2012

Brief presentation of the article:

As discussed in the Introduction section of this manuscript, the Sm-like Hfq protein is a key player in the regulatory mechanism involving small non-coding RNAs and target mRNAs. This is particularly the case for the most common bacterial models, namely, Gram-negative enteric bacteria such as *Salmonella* and *E. coli*. Indeed, the stability of the sRNAs and their successful interaction with the target sequences in mRNAs are largely dependent on the availability of Hfq see (De Lay *et al.*, 2013). The current view is that the different RNAs could compete with each other for Hfq binding. Thus, the cellular concentration of Hfq and the differential affinity of the sRNA for the chaperon protein could play a role in the regulation itself (Wagner, 2013). In Gram-positive bacteria, the precise role of Hfq remains ill defined in many instances (Jousselin *et al.*, 2009). Paradoxically however, the first high-resolution crystallographic data were obtained with *Staphylococcus aureus* Hfq (Schumacher *et al.*, 2002). and subsequently, *S. aureus* Hfq structure together with *E. coli*'s have played a determinant role in helping understand the mechanistic aspects of Hfq:RNA interactions (Link *et al.*, 2009, Sauter *et al.*, 2003, Schumacher *et al.*, 2002). In spite of being so the well

characterized structurally, the levels of expression of the *hfq* gene and the role of *S aureus* Hfq in bacterial physiology have remained poorly defined .

The Bossi laboratory has been interested in Hfq function as part of the sRNA regulatory function from the very beginning of their studies on sRNA regulation in *Salmonella enterica*. Indeed, the first experiments designed in the laboratory were aimed at identifying sRNA target genes using the dependence of target gene expression on Hfq activity as an indicator of sRNA-mediated regulation. This genetic approach yielded dozens of reporter gene fusions whose level of expression depend on Hfq function (Figueroa-Bossi et al., 2006a). Using these tools, the functional properties of *Staphylococcus aureus* Hfq protein were studied, when the protein is expressed from the *Salmonella* chromosome, in replacement of the endogenous protein. To further extend the work, the *B. burgdorferi* Hfq protein was also included in the project. By the same experimental approach we could examine this Hfq protein, a recent report having indicated the functionality of *Borrelia's* Hfq and its ability to complement an *E. coli* null mutant (Lybecker et al., 2010).

The experimental design also took advantage of the “no-mark” recombineering approach based on Lambda red mediated gene replacement. The coding sequence of *S. Typhimurium* Hfq (Hfq_{STM}) was replaced by the respective coding sequences of *S. aureus* (Hfq_{SA}) or *B burgdorferi* (Hfq_{BB}). The work in which I participated examined three main aspects related to Hfq in these strains: first, expression of the heterologous proteins under *S. Typhimurium* promoter control was confirmed by Western analysis of epitope tagged proteins; second, reporter fusions to sRNA controlled genes known to be dependent on Hfq function served to evaluate the relative activity of the proteins; and third, Northern analysis of selected sRNAs, whose stability is known to depend on the presence of Hfq, were performed in the different strains. The ensemble of the results indicates that, even though the heterologous proteins were expressed to detectable levels, comparable to those of the endogenous protein, no functional complementation could be detected. Finally, results from Northern blot analysis showed the sRNA

patterns in the heterologous *hfq* background to be close to those of the S. Typhimurium *hfq* deletion strain.

Lack of interchangeability of Hfq-like proteins

Tatiana Rochat¹, Philippe Bouloc¹, Qi Yang²,

Lionello Bossi² and Nara Figueroa-Bossi^{2*}

¹ Institut de Génétique et Microbiologie, CNRS, UMR 8621, IFR115, Université Paris-Sud, Orsay, France

² Centre de Génétique Moléculaire, CNRS, UPR3404, IFR115, Université Paris-Sud, Gif-sur-Yvette, France

* Corresponding author

Centre de Génétique Moléculaire du CNRS, 91198 Gif-sur-Yvette, France

Email: figueroa@cgm.cnrs-gif.fr

Phone: + 33 169 823 811 Fax: +33 169 823 877

Abstract

Hfq is an RNA-binding protein that participates in the regulatory activity of small non-coding RNAs (sRNAs) in many species of bacteria. Hfq protein was first crystallized from *Staphylococcus aureus* and this crystal structure constitutes a hallmark for bacterial Sm-like proteins. Paradoxically, however, the functional relevance/role of *S. aureus* Hfq (Hfq_{SA}) remains uncertain, as growing evidence suggests that the *hfq*_{SA} gene is expressed at very low levels or unexpressed in many *S. aureus* strains. To gather further insight, in the present work we exchanged the structural portion of the *hfq* gene of *Salmonella enterica* serovar Typhimurium (*hfq*_{STM}) with *hfq*_{SA} and analyzed the effects of the replacement on various Hfq-related phenotypes. Our results show that the replacement strain – in spite of expressing Hfq_{SA} at levels comparable to Hfq_{STM} in wild-type *Salmonella* – behaves as an *hfq* null mutant in three discrete small RNA-mediated regulatory responses. These defects correlate with an abrupt reduction in the intracellular concentration of sRNAs, as observed in an *hfq* null mutant. Failure of Hfq_{SA} to protect *Salmonella* sRNAs from degradation suggests that Hfq_{SA} does not bind to these sRNAs. A parallel study with the *Borrelia burgdorferi* *hfq* gene (*hfq*_{BB}) gave essentially identical results: when made from a single copy chromosomal gene, Hfq_{BB} fails to substitute for Hfq_{STM} in sRNA-mediated regulation.

Keywords: Hfq, *Salmonella*, small RNA, *Staphylococcus aureus*, *Borrelia burgdorferi*

1. Introduction

Base pairing between regulatory RNAs and complementary sequences in messenger RNAs is a highly conserved mechanism controlling gene expression at the post-transcriptional level in all forms of life. In bacteria, best-studied regulatory small RNAs (sRNAs) are encoded at separate locations from their target genes and interact with target mRNAs through short and imperfect stretches of complementarity (reviewed in [1]). In some bacterial species, like *Escherichia coli*, *Salmonella enterica* and *Vibrio* species, these *trans*-encoded sRNAs require chaperon protein Hfq for activity. Discovered in the late nineteen sixties as a host factor needed for in vitro replication of RNA phage Q β [2], Hfq was later found to be a key player in a number of RNA transactions (reviewed in [3-5]). In particular, Hfq participates in sRNA-mediated regulation by binding both sRNAs and cognate mRNAs and stimulating their association. For many sRNAs, Hfq binding is also essential to confer protection against degradation by ribonucleases [6]. Hfq and Hfq-like proteins belong to the Sm-like (Lsm) family of RNA-binding proteins characterized by a ring-shaped multimeric architecture. X-ray crystal structure analysis of the Hfq-like protein from *Staphylococcus aureus* (Hfq_{SA}) and of *E. coli* Hfq (Hfq_{EC}) showed that both proteins assemble in homohexameric rings approximately 70 Å in diameter [7-9]. A short U-rich synthetic RNA co-crystallized with Hfq_{SA} was found to circle around the positively charged central pore of the torus on the so-called proximal face [9]. In contrast, A-containing RNA oligomers bind to the distal face of Hfq_{EC} [7]. The notion of opposite Hfq surfaces having different ligand specificities is independently supported by mutational studies [10]. It should be noticed, however, that while similar on the proximal face, Hfq from *S. aureus* and *E. coli* differ sharply in their charge distribution on the distal face and in the trough that connects proximal and distal faces. This latter region has a positively charged surface in Hfq_{EC} and a negatively charged surface in Hfq_{SA} [3]. Finally, while Hfq-like proteins have an evolutionarily conserved core of 65 amino acids, the C-terminus is variable in length, leading to a controversy about its function [11, 12]. The Hfq extended C-terminus is found in γ - and β -proteobacteria whereas in the case of Gram-positive bacteria such as in Hfq_{SA}, extensions are short.

In most bacteria, loss of Hfq function, albeit not a lethal event, causes a variety of pleiotropic defects and renders strains particularly susceptible to environmental stress

[13]. Many of these phenotypes result from the loss of sRNA activities. For example, the σ^S -dependent stress response is poorly induced in *hfq* mutants due to the lack of sRNA-mediated activation of *rpoS* mRNA translation (reviewed in [14]). At the same time, the σ^E -dependent envelope stress response is chronically induced due to over-accumulation of outer membrane proteins that are normally downregulated by sRNAs [15-20].

Genes encoding Hfq-like proteins are found in about half of the sequenced genomes [21]. In many pathogens, they are required for virulence and were shown to participate in sRNA-mediated regulatory processes (reviewed in [22]). The functional proficiency of Hfq homologues was inferred from their ability to complement the loss of Hfq in *E. coli*. For example, *Pseudomonas aeruginosa* Hfq, which shares 92% identity with Hfq_{EC} in the initial 68 amino acids, fully replaced Hfq_{EC} in terms of its requirement from Q β replication and *rpoS* expression [23]. Likewise, the *Moraxella catarrhali* *hfq*-like gene, in spite of being twice the size of the *hfq*_{EC} (but highly similar in the N-terminal encoding domain) partially complemented the growth defect and the stress sensitivity of an *E. coli* *hfq* mutant [24]. Intriguingly, even a protein with very limited sequence relatedness to Hfq (only 12% identity), encoded by *Borrelia burgdorferi* (Hfq_{BB}), was recently reported to complement an *E. coli* *hfq* mutant [25]. On the other hand, Hfq-like proteins from *Neisseria meningitidis* and *Aquifex aeolicus*, and from archaeon *Methanocaldococcus jannaschii* did not reverse the chronic σ^E induction of a *Salmonella* strain lacking Hfq [26]. However, these proteins were capable of binding some *Salmonella* sRNAs and also caused specific RNA processing defects [26].

The function of Hfq-like proteins remains unclear for some bacterial species. For example, deletion of the *Bacillus subtilis* *hfq*-like gene (*ymaH*) does not affect growth rate, stress adaptation, or activities of all sRNAs tested [27-29]. Similar results were reported for *S. aureus*. Deletion of the *hfq*_{SA} gene in several pathogenic isolates did not impair or in any way impact their physiology [30] possibly because *hfq*_{SA} is poorly expressed or not expressed in the strains used for these studies [30, 31](see also [32]). On the other hand, in *S. aureus* strains where Hfq is detected, deletion of its coding gene reportedly resulted in decreased toxicity and virulence, suggesting that Hfq is a global regulator [33].

To gather insight into the functional status and regulatory properties of *S. aureus* Hfq, in the present study, we introduced the sequence encoding this protein in place

of the endogenous *hfq* gene in the *Salmonella* chromosome. (Hfq_{STM} is 100% identical to Hfq_{EC} in the initial 78 amino acids). In parallel with the above work, a similar exchange was performed using a DNA fragment spanning the *Borrelia hfq*-like gene. Both constructs were made with surgical precision replacing only protein coding portions – *i.e.*, the segment between initiation and termination codons – to allow the heterologous sequences to fall under the control of signals normally devoted to *hfq*_{STM} expression. We show below that Hfq_{SA} and Hfq_{BB} in spite of being made at levels comparable to Hfq_{STM}, fail to replace the latter in sRNA-mediated regulation as well as in protecting representative sRNAs from degradation.

2. Materials and methods

Bacterial strains and growth conditions

Strains used in this study, listed in Table 1, were all derived from *Salmonella enterica* serovar Typhimurium strain MA3409, a derivative of strain LT2 cured for the Gifsy-1 prophage [34]. Bacteria were cultured at 37°C in liquid media or in media solidified by the addition of 1.5% (w/v) Difco agar. LB broth [1% bacto tryptone (w/v), 0.5% Difco yeast extract (w/v), 0.5% NaCl (w/v)] was used as complex medium. When needed, LB medium was supplemented with 0.2% (w/v) arabinose. Antibiotics (Sigma-Aldrich) were included at the following final concentrations: chloramphenicol, 10 µg/ml; kanamycin monosulphate, 50 µg/ml; sodium ampicillin 100 µg/ml; tetracycline hydrochloride, 25 µg/ml. Liquid cultures were grown in New Brunswick gyratory shakers and growth was monitored by measuring the optical density at 600 nm with a Shimazu UV-mini 1240 spectrophotometer.

Enzymes and chemicals

Restriction enzymes, T4 polynucleotide kinase and Taq DNA polymerase were from New England Biolabs, Pfu-Turbo DNA polymerase was from Stratagene. DNA oligonucleotides were obtained from Sigma Aldrich. Acrylamide-bisacrylamide (30%, 29:1) and other electrophoresis reagents were from BioRad. Hybond-N⁺

membranes and hybridization buffer used for Northern blot analysis were from GE Healthcare and from Applied Biosystems-Ambion, respectively.

Genetic techniques

Generalized transduction was carried out using the high frequency transducing mutant of phage P22, HT 105/1 *int-201* [35]. “ λ Red”-mediated chromosomal recombineering was carried out by the method of Datsenko and Wanner [36] implemented as in [37]. Donor DNA fragments were generated by the polymerase chain reaction (PCR) using plasmid or chromosomal DNA templates. DNA oligonucleotides used as primers for PCR amplification are listed in Table 2. Amplified fragments were electroporated into the appropriate strains using a Bio-Rad MicroPulser under the conditions specified by the manufacturer. Constructs were verified by PCR and DNA sequence analysis (performed by GATC company).

Construction of relevant strains

Salmonella strains carrying the structural portions of the *hfq*-like genes from *Staphylococcus aureus* or *Borrelia burgdorferi* were constructed with a two-step recombineering procedure as described [38]. Firstly, a *tetAR* module (amplified with primers ppH71 and ppH72) was inserted in the *hfq* gene in the *Salmonella* chromosome. Subsequently, the entire *hfq::tetAR* was crossed out selecting for the loss of tetracycline resistance [39] using DNA fragments amplified from chromosomal DNA of *S. aureus* RN4220 and *B. burgdorferi* clinical isolate 28354 with primer pairs ppI06/ppI07 and ppJ44/ppJ45, respectively (Table 2). Introduction of the 3xFLAG epitope at the 3' ends of the coding sequence of *hfq*_{STM}, *hfq*_{SA} and *hfq*_{BB} was carried out using DNA fragments amplified from plasmid pSUB11 [37] with primer pairs pp913/pp914, ppJ41/pp914 and ppJ40/pp914, respectively, as previously described [37].

RNA extraction and Northern blot analysis

RNA was prepared by the acid-hot-phenol method from exponentially growing cells (OD₆₀₀ of 0.35) as previously described [40]. RNA was separated on an 8% polyacrylamide-8 M urea gel and electro-blotted onto a nylon membrane. Blots were

hybridized to 5' end-labelled DNA oligonucleotide probes specific for the sRNAs under study (listed in Table 3). Hybridization signals were analyzed by Phosphorimaging using ImageQuant software.

Western blot analysis

Bacterial cells harvested by centrifugation were resuspended in 250 μ l of Laemmli protein gel loading buffer. Whole-cell extracts were fractioned in a 15% polyacrylamide-SDS gel. Proteins were transferred to poly-vinylidene difluoride (PVDF) membranes and probed with anti-FLAG M2 monoclonal antibodies from Sigma. Horse Radish Peroxidase (HRP) conjugated goat anti-mouse IgG (KPL) was used as a secondary antibody. Detection was performed by Enhanced Chemiluminescence (ECL, Amersham) with ECL films (Amersham).

β -galactosidase assays

Activity of β -galactosidase was measured in toluene-permeabilized cells as described [41] and is expressed in Miller units. Reported values were the average of at least two independent determinations, each involving duplicate or triplicate samples.

3. Results and discussion

Exchanging hfq coding sequences

Replacement of the structural portion of the *Salmonella hfq* gene with the corresponding regions from the *hfq* genes of *Staphylococcus aureus* and *Borrelia burgdorferi* was achieved by a two-step recombineering procedure (see Materials and Methods). DNA sequence analysis of the constructs (strains MA10741 and MA11042) showed that they carry no extra material or scars and that the *hfq*_{SA} and *hfq*_{BB} coding sequences are precisely positioned in replacement of the endogenous *hfq* coding sequence in the *Salmonella* chromosome (Figs. 1 and S1). We therefore expected that *hfq*_{SA} and *hfq*_{BB} would be expressed from the transcription/translation initiation signals of *hfq*_{STM}. To confirm this, strain derivatives carrying 3xFLAG epitope fusions to the 3' ends of all three *hfq* genes were constructed, and expression of the carboxy-terminally tagged proteins assessed by Western blot analysis. As

shown in Fig. 2, all three proteins accumulate at significant levels. Some differences in the intensity of the hybridization signals are apparent; however, the fact that they correlate with the molecular weights of the proteins suggest that they are intrinsic to the Western protocol (increasingly less quantitative with smaller molecules) and do not reflect a difference in expression rates. Overall, the data in Fig. 2 strongly suggest that *hfq*_{SA} and *hfq*_{BB} are expressed at levels closely comparable to that of *hfq*_{STM}.

*Hfq*_{SA} and *Hfq*_{BB} fail to replace *Hfq*_{STM} in sRNA-mediated regulation

Genes regulated directly or indirectly by small RNA constitute reporter systems for Hfq activity. Three representative loci were chosen for this analysis. The *chiP* gene of *Salmonella* encodes a chitoporin whose synthesis is repressed by a constitutively made sRNA, ChiX, which blocks *chiP* mRNA translation under most laboratory growth conditions. ChiX is the tightest Hfq-binding RNA known to date [42] and strongly dependent on Hfq binding for stability [43]. Thus, a chromosomal translational *chiP-lacZ* fusion is expressed at very low level when Hfq is functional, but becomes derepressed nearly 40-fold in a strain deleted for *hfq* (Fig3A). As shown in Fig3A, strains carrying *hfq*_{SA} or *hfq*_{BB} in place of *hfq*_{STM} have β -galactosidase activities similar that of the *hfq* deleted strain. Consistent with these findings, Northern blot analysis reveals that ChiX sRNA is nearly undetectable in the *hfq*_{SA} or *hfq*_{BB} expressing strains, like what observed in the absence of Hfq (Fig. 4). Failure of Hfq_{SA} or Hfq_{BB} to protect ChiX from degradation suggests that neither of the two proteins binds to this sRNA *in vivo*.

The second reporter system used in this study is the *eptB* gene, which encodes phosphoethanolamine transferase, an enzyme involved in lipopolysaccharide modification. A previous study identified *eptB* as one of the genes most dramatically upregulated in a Δhfq mutant [18]. This results from the combined effect of two apparently independent mechanisms: activation of the σ^E regulon [18] and loss of repression by MgrR sRNA [44]. To examine whether Hfq_{SA} or Hfq_{BB} corrected these regulatory defects in any way, a translational *eptB-lacZ* fusion was introduced into in the *hfq*_{SA}- or *hfq*_{BB}-expressing strains and β -galactosidase activity measured. Again, the results in Fig. 3B show that the two strains are indistinguishable from the *hfq* null mutant in this test.

Both the above systems represent examples of negative regulation. We thus included a third reporter system in which Hfq activity is needed for gene activation. Due to the requirement of Hfq-dependent sRNAs for optimal translation of σ^S mRNA, genes controlled by σ^S are typically less expressed in the absence of Hfq [18]. One such gene encodes the KatE catalase. A further set of strains was therefore constructed carrying a *katE-lacZ* fusion in the *hfq_{SA}* or *hfq_{BB}* genetic background. As shown in Fig. 3C neither of the heterologous Hfq proteins can compensate for the loss of the endogenous Hfq as far as *katE-lacZ* expression is concerned.

Hfq_{SA} and Hfq_{BB} fail to protect sRNAs from degradation

As described above, ChiX sRNA is almost undetectable in the strains carrying *hfq_{SA}* or *hfq_{BB}* (Fig. 4). This analysis included additional representative sRNAs: GcvB, MicA, MicC and MicF. As shown in Fig. 4, the Northern hybridization patterns from the *hfq_{SA}*- and *hfq_{BB}*-carrying strains in all cases are indistinguishable from that of *hfq* deleted strain: MicC and MicF are both undetectable, while GcvB decreases significantly. In the case of MicA, the slight increase in sRNA levels is ascribable to σ^E activation since the *micA* gene is a member of the σ^E regulon [18, 20]. The broadening of the MicA band, on the other hand, reflects the increased decay resulting from the Hfq defect (see [18]). These findings are reminiscent of those by Sittka and co-workers who found *Methanococcus* Hfq incapable of reversing the chronic σ^E activation of a *Salmonella* strain lacking Hfq [26]. Altogether, the above data suggest that neither Hfq_{SA} nor Hfq_{BB} are capable of binding *Salmonella* sRNAs *in vivo*.

Staphylococcus aureus Hfq is representative of bacterial Sm-like proteins and an inspirational reference in the small RNA field. Somewhat ironically, however, the function of this protein remains enigmatic. Although the protein was shown to bind RNAIII, a major RNA regulator in *S. aureus* [45], it does not appear to facilitate the RNAIII's interaction with any of the mRNA targets that have been looked at (reviewed in [46]). Data on the effects of *hfq* deletions on various aspects of *S. aureus* physiology are contradictory, possibly due to strain variability in the expression levels of *hfq* [30, 33]. Considering that a possible redundancy in Hfq-like functions might underlie such variability in *S. aureus*, we chose to examine the functional status of Hfq_{SA} in *Salmonella enterica*, where the role of the unique Hfq protein is well

established. Our results showed that Hfq_{SA} does not participate in any of the sRNA-mediated regulatory mechanisms that were tested. Furthermore, we found no evidence of Hfq_{SA} binding to sRNAs *in vivo*. Altogether, these data indicate that Hfq_{SA} is functionally very different from its homologues in enteric bacteria. Such functional divergence might be in relation with the structural differences noted earlier on the distal face of the protein.

A second set data from this study shows that the Hfq-like protein from *Borrelia burgdorferi* is also unable to replace Hfq_{STM} in sRNA-mediated regulation and sRNA stabilization. These findings are surprising as a recent publication reported Hfq_{BB} to complement the defect in *rpoS* mRNA regulation of an *E. coli hfq* mutant [25]. The discrepancy is difficult to reconcile. The authors of the above report performed their analysis using a plasmid-born *hfq*_{BB} gene; thus, in principle, the discrepancy could arise from gene dosage differences. Somehow, however, this explanation seems unlikely. An alternative explanation might lie in the nature of the *E. coli hfq* allele - *hfq-1* - used in the above study. The *hfq-1* allele was originally constructed by inserting an Ω (KnR) cassette into the BclI site spanning positions 117-122 of the *hfq*_{EC} coding sequence [13]. Therefore, the mutant is expected to express a truncated Hfq fragment of about 40 amino acids, which would include the N-terminal α 1 helix and two of the β strands of the Sm1 motif. Perhaps this fragment could somehow oligomerize with the Hfq_{BB} monomers to reconstitute a functional Hfq protein. Testing complementation in an *E. coli* strain in which the *hfq* gene is completely deleted would help remove this doubt.

Concluding remarks

The data presented here convincingly show that the Hfq proteins from *S. aureus* and *B. burgdorferi* are not interchangeable with their *Salmonella* counterpart as far as sRNA-mediated regulation is concerned. Clearly, more work is needed to elucidate the basis and biological rationale for the differences. Some of the strains constructed in the course of this work might prove useful to this end. The 3xFLAG-tagged derivative of Hfq_{SA} and Hfq_{BB} could be used to test whether there exist RNAs that bind either of these proteins in *Salmonella* and possibly identify them [26]. As a complementary strategy, one could take advantage of the wide array of *Salmonella* genetic tools to search for genes whose expression might change in the presence or

absence of Hfq_{SA} or Hfq_{BB}. These studies should help improve our understanding of the function, mode of action and ligand specificity of Sm-like proteins in bacteria.

Acknowledgements

This work was supported by a grant from the IFR 115, Genomes, Transcriptomes, Proteomes to NFB and a grant from the Agence Nationale pour la Recherche grant (ANR-2010-BLAN-1602-01 'Duplex-Omics') to PB. We thank Christine Pourcel (Orsay, France) for providing *B. burgdorferi* genomic DNA, and S. Yoshizawa and D. Fourmy (CGM, Gif-sur-Yvette) for helpful discussions.

4. Legends to the Figures

Fig. 1. Schematic representation of recombineering events replacing the coding segment of the *hfq* gene in the *Salmonella* chromosome (309 bp) with the 234 bp DNA fragment and the 480 bp DNA fragment encoding the structural portions of the *hfq*-like gene from *S. aureus* and *B. burgdorferi*, respectively. The construction was carried out as described in Materials and methods.

Fig. 2. Immunodetection of 3xFLAG-tagged Hfq_{STM}, Hfq_{SA} and Hfq_{BB} proteins in *Salmonella* extracts. Bacteria harvested from stationary cultures were lysed [37] and crude extracts were fractionated by electrophoresis in a 15% SDS-polyacrylamide gel. The gel was processed for the immunodetection of epitope-tagged proteins [37]. The strains used were (from left to right): MA9132 (wt, no FLAG), MA11054 (*hfq*_{STM}-3xFLAG), MA11055 ($\Delta hfq::hfq_{SA}$ -3xFLAG) and MA11056 ($\Delta hfq::hfq_{BB}$ -3xFLAG). The identity and origin of the non-specific, low molecular weight band present in all strains analyzed is unknown.

Fig. 3. Expression of translational *lacZ* fusions to chromosomal genes sensitive to Hfq inactivation in *Salmonella enterica*. β -galactosidase activity was measured in exponentially growing LB cultures ($OD_{600} \approx 0.3$) (A and B) or in stationary phase cultures ($OD_{600} \approx 2$) (C). Strains used were: A. MA9132, MA10744, MA10747 and MA11044; B. MA8028, MA8029, MA10746 and MA11043; C. MA8149, MA8679, MA11057 and MA11058 (see Table 1 for full genotypes).

Fig. 4. Northern blot analysis of small RNAs. Total RNA extracted from exponentially growing cells ($OD_{600} = 0.35$) was fractionated on an 8% polyacrylamide-8 M urea gel and electrotransferred onto a Hybond-N⁺ membrane. Blots were hybridized to 5' end-labeled DNA oligonucleotide probes (Table 3). *Salmonella* strains used were: MA8028 (*hfq*_{STM} wt), MA8029 ($\Delta hfq_{STM}::cat$), MA10746 ($\Delta hfq_{STM}::hfq_{SA}$), MA11044 ($\Delta hfq_{STM}::hfq_{BB}$) and MA9132 (*hfq*_{STM} wt).

Table 1. *Salmonella enterica* serovar Typhimurium strains used in this work.

Strain ^a	Genotype	Source or reference
MA3409	wild-type	[34]
MA7455	wild-type / pKD46	[18]
MA8028	<i>eptB115::MudK</i>	[18]
MA8029	<i>eptB115::MudK Δhfq67::cat</i>	[18]
MA8149	<i>katE561::MudK</i>	[18]
MA8679	<i>katE561::MudK Δhfq67::cat</i>	[18]
MA9132	<i>chiP91::pCE40(lac)</i>	[43]
MA10675	<i>Δhfq116::tetAR</i>	this work
MA10740	<i>Δhfq116::tetAR</i> / pKD46	this work
MA10741	<i>Δhfq_{STM}::hfq_{SA}</i>	this work
MA10744	<i>chiP91::pCE40(lac) Δhfq116::tetAR</i>	this work
MA10746	<i>eptB115::MudK Δhfq_{STM}::hfq_{SA}</i>	this work
MA10747	<i>chiP91::pCE40(lac) Δhfq_{STM}::hfq_{SA}</i>	this work
MA11042	<i>Δhfq_{STM}::hfq_{BB}</i>	this work
MA11043	<i>eptB115::MudK Δhfq_{STM}::hfq_{BB}</i>	this work
MA11044	<i>chiP91::pCE40(lac) Δhfq_{STM}::hfq_{BB}</i>	this work
MA11054	<i>hfq-3xFLAG-aph</i> (KnR)	this work
MA11055	<i>Δhfq_{STM}::hfq_{SA}-3xFLAG-aph</i> (KnR)	this work
MA11056	<i>Δhfq_{STM}::hfq_{BB}-3xFLAG-aph</i> (KnR)	this work
MA11057	<i>katE561::MudK Δhfq_{STM}::hfq_{SA}</i>	this work
MA11058	<i>katE561::MudK Δhfq_{STM}::hfq_{BB}</i>	this work

^a All strains are derived from *Salmonella enterica* serovar Typhimurium strain MA3409. The latter is a derivative of strain LT2 cured for the Gifsy-1 prophage [34].

Table 2. DNA oligonucleotides used as PCR primers for λ Red-mediated recombination.

<i>Primer^a</i>	<i>Sequence (5'–3')</i>
ppH71 (fw)	ATTGCGTCGGGAACGTGTTCCAGTTTCTATTTATTTGGTG <i>TTAAGACCCACTTTCACATT</i>
ppH72 (rv)	AACTGATCAAAGGACTCGATTTGACCTTGCAGCTTAATACC <i>CTAAGCACTTGTCTCTG</i>
ppI06 (fw)	AGGTTCAAAGTACAAATAAGCATATAAGGAAAAGAGA <i>ATGATTGCAAACGAAAACATCCA</i>
ppI07 (rv)	ATTATCCGACGCCCCCGACATGGATAAACAGCGCGTGAAC <i>TTATTCCTTCACTTTCAGTAGATGC</i>
ppJ44 (fw)	AAAGTTCAAAGTACAAATAAGCATATAAGGAAAAGAGA <i>ATGTTTATAAGCAGGGAATTGAAG</i>
ppJ45 (rv)	ATTATCCGACGCCCCCGACATGGATAAACAGCGCGTGAAC <i>TTATTCCTTCTTGCTCATTAAAG</i>
pp913 (fw)	GCAGGGGTCTACTGCGCAACAGGACAGCGAAGAGACTGAA <i>GACTACAAAGACCATGACGG</i>
pp914 (rv)	ATTATCCGACGCCCCCGACATGGATAAACAGCGCGTGAAC <i>CATATGAATATCCTCCTTAG</i>
ppJ40 (fw)	AAAGTTGTTGCAGATGCTATTA AAACTTTAATGAGCAAGAAGGAA <i>GACTACAAAGACCATGACGG</i>
pp914 (rv)	above
ppJ41 (fw)	TATACAGTAGAAACTGAAGGTCAAGCATCTACTGAAAGTGAAGAA <i>GACTACAAAGACCATGACGG</i>
pp914 (rv)	above

^a Primers are defined as “forward” (fw) or “reverse” (rv) depending on whether they have same or opposite orientation (5'-3') relative to the gene being modified. The portions of primers annealing to template DNA are in red italics.

Table 3. DNA oligonucleotides used as probes for Northern analysis.

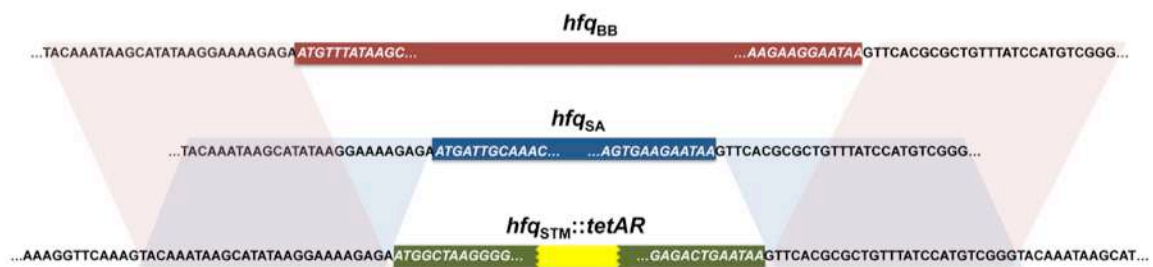
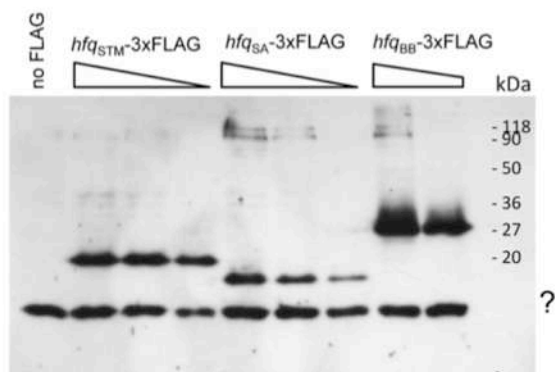
<i>Name</i>	<i>Sequence (5'–3')</i>	<i>Specificity</i>
pp814	ATGATGATAACAAATGCGCG	MicA
pp831	AGGTTAACGCAATGGCCCAG	MicC
pp832	AGGGGTAAACAGACATTCAG	MicF
ppB07	CGTCAAAGAGGAATTTTCATCGTTATTATTATCCCGACGCTTTCGCTTC	ChiX
ppB10	ACACTACCATCGGCGCTACG	5 S
ppI67	AGACCAATTGCAAACACAACAACATCACAACCGTAAGCCA	GcvB

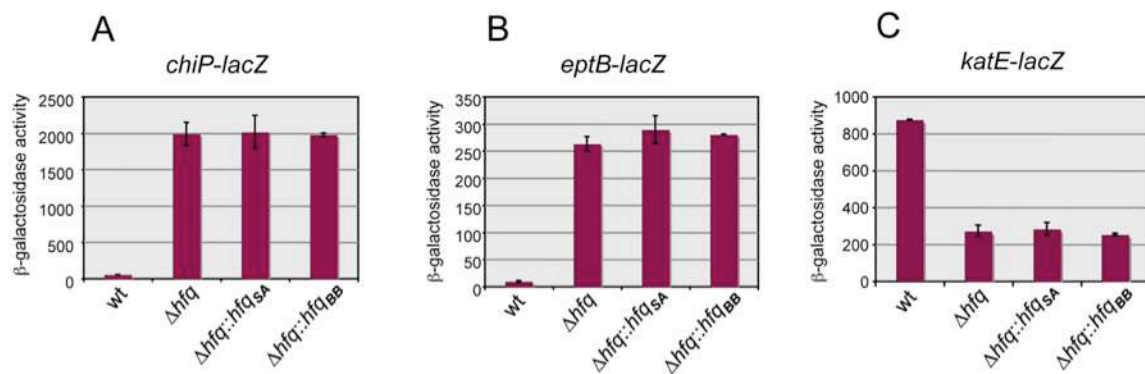
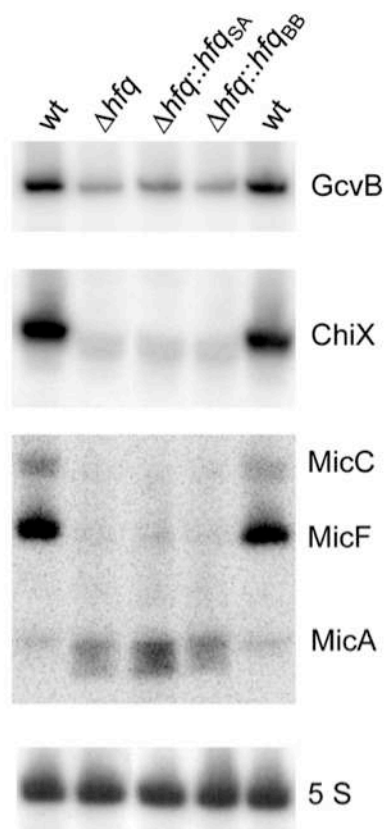
References

- [1] G. Storz, J. Vogel, K.M. Wassarman, Regulation by small RNAs in bacteria: expanding frontiers, *Mol Cell* 43 (2011) 880-891.
- [2] M.T. Franze de Fernandez, L. Eoyang, J.T. August, Factor fraction required for the synthesis of bacteriophage Q β -RNA, *Nature* 219 (1968) 588-590.
- [3] R.G. Brennan, T.M. Link, Hfq structure, function and ligand binding, *Curr Opin Microbiol* 10 (2007) 125-133.
- [4] P. Valentin-Hansen, M. Eriksen, C. Udesen, The bacterial Sm-like protein Hfq: a key player in RNA transactions, *Mol Microbiol* 51 (2004) 1525-1533.
- [5] J. Vogel, B.F. Luisi, Hfq and its constellation of RNA, *Nat Rev Microbiol* 9 (2011) 578-589.
- [6] H. Aiba, Mechanism of RNA silencing by Hfq-binding small RNAs, *Curr Opin Microbiol* 10 (2007) 134-139.
- [7] T.M. Link, P. Valentin-Hansen, R.G. Brennan, Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA, *Proc Natl Acad Sci U S A* 106 (2009) 19292-19297.
- [8] C. Sauter, J. Basquin, D. Suck, Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*, *Nucleic Acids Res* 31 (2003) 4091-4098.
- [9] M.A. Schumacher, R.F. Pearson, T. Moller, P. Valentin-Hansen, R.G. Brennan, Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein, *Embo J* 21 (2002) 3546-3556.
- [10] P.J. Mikulecky, M.K. Kaw, C.C. Brescia, J.C. Takach, D.D. Sledjeski, A.L. Feig, *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, *rpoS* and poly(A) RNAs, *Nat Struct Mol Biol* 11 (2004) 1206-1214.
- [11] A.S. Olsen, J. Moller-Jensen, R.G. Brennan, P. Valentin-Hansen, C-terminally truncated derivatives of *Escherichia coli* Hfq are proficient in riboregulation, *J Mol Biol* 404 (2010) 173-182.
- [12] B. Vecerek, L. Rajkowitsch, E. Sonnleitner, R. Schroeder, U. Blasi, The C-terminal domain of *Escherichia coli* Hfq is required for regulation, *Nucleic Acids Res* 36 (2008) 133-143.
- [13] H.C. Tsui, H.C. Leung, M.E. Winkler, Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12, *Mol Microbiol* 13 (1994) 35-49.
- [14] F. Repoila, N. Majdalani, S. Gottesman, Small non-coding RNAs, coordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm, *Mol Microbiol* 48 (2003) 855-861.
- [15] L. Bossi, D. Maloriol, N. Figueroa-Bossi, Porin biogenesis activates the σ^E response in *Salmonella hfq* mutants, *Biochimie* 90 (2008) 1539-1544.
- [16] Y. Ding, B.M. Davis, M.K. Waldor, Hfq is essential for *Vibrio cholerae* virulence and downregulates sigma expression, *Mol Microbiol* 53 (2004) 345-354.
- [17] V. Douchin, C. Bohn, P. Boulloc, Down-regulation of porins by a small RNA bypasses the essentiality of the regulated intramembrane proteolysis protease RseP in *Escherichia coli*, *J Biol Chem* 281 (2006) 12253-12259.
- [18] N. Figueroa-Bossi, S. Lemire, D. Maloriol, R. Balbontín, J. Casadesús, L. Bossi, Loss of Hfq activates the σ^E -dependent envelope stress response in *Salmonella enterica*, *Mol Microbiol* 62 (2006) 838-852.

- [19] E. Guisbert, V.A. Rhodius, N. Ahuja, E. Witkin, C.A. Gross, Hfq modulates the σ^E -mediated envelope stress response and the σ^{32} -mediated cytoplasmic stress response in *Escherichia coli*, *J Bacteriol* 189 (2007) 1963-1973.
- [20] K. Papenfort, V. Pfeiffer, F. Mika, S. Lucchini, J.C. Hinton, J. Vogel, σ^E -dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global *omp* mRNA decay, *Mol Microbiol* 62 (2006) 1674-1688.
- [21] X. Sun, I. Zhulin, R.M. Wartell, Predicted structure and phyletic distribution of the RNA-binding protein Hfq, *Nucleic Acids Res* 30 (2002) 3662-3671.
- [22] Y. Chao, J. Vogel, The role of Hfq in bacterial pathogens, *Curr Opin Microbiol* 13 (2010) 24-33.
- [23] E. Sonnleitner, I. Moll, U. Blasi, Functional replacement of the *Escherichia coli* *hfq* gene by the homologue of *Pseudomonas aeruginosa*, *Microbiology* 148 (2002) 883-891.
- [24] A.S. Attia, J.L. Sedillo, W. Wang, W. Liu, C.A. Brautigam, W. Winkler, E.J. Hansen, *Moraxella catarrhalis* expresses an unusual Hfq protein, *Infect Immun* 76 (2008) 2520-2530.
- [25] M.C. Lybecker, C.A. Abel, A.L. Feig, D.S. Samuels, Identification and function of the RNA chaperone Hfq in the Lyme disease spirochete *Borrelia burgdorferi*, *Mol Microbiol* 78 (2010) 622-635.
- [26] A. Sittka A., C.M. Sharma, K. Rolle, J. Vogel, Deep sequencing of *Salmonella* RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes, *RNA Biol* 6 (2009) 266-275.
- [27] A. Gaballa, H. Antelmann, C. Aguilar, S.K. Khakh, K.B. Song, G.T. Smaldone, J.D. Helmann, The *Bacillus subtilis* iron-sparing response is mediated by a Fur-regulated small RNA and three small, basic proteins, *Proc Natl Acad Sci U S A* 105 (2008) 11927-11932.
- [28] N. Heidrich, A. Chinali, U. Gerth, S. Brantl, The small untranslated RNA SR1 from the *Bacillus subtilis* genome is involved in the regulation of arginine catabolism, *Mol Microbiol* 62 (2006) 520-536.
- [29] J.M. Silvaggi, J.B. Perkins, R. Losick, Small untranslated RNA antitoxin in *Bacillus subtilis*, *J Bacteriol* 187 (2005) 6641-6650.
- [30] C. Bohn, C. Rigoulay, P. Boulloc, No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*, *BMC Microbiol* 7 (2007) 10.
- [31] E. Geisinger, R.P. Adhikari, R. Jin, H.F. Ross, R.P. Novick, Inhibition of *rot* translation by RNAIII, a key feature of *agr* function, *Mol Microbiol* 61 (2006) 1038-1048.
- [32] B. Felden, F. Vandenesch, P. Boulloc, P. Romby, The *Staphylococcus aureus* RNome and its commitment to virulence, *PLoS Pathog* 7 (2011) e1002006.
- [33] Y. Liu, N. Wu, J. Dong, Y. Gao, X. Zhang, C. Mu, N. Shao, G. Yang, Hfq is a global regulator that controls the pathogenicity of *Staphylococcus aureus*, *PLoS One* 5 (2010).
- [34] N. Figueroa-Bossi, E. Coissac, P. Netter, L. Bossi, Unsuspected prophage-like elements in *Salmonella typhimurium*, *Mol Microbiol* 25 (1997) 161-173.
- [35] H. Schmieger, Phage P22-mutants with increased or decreased transduction abilities, *Mol Gen Genet* 119 (1972) 75-88.
- [36] K.A. Datsenko, B.L. Wanner, One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products, *Proc Natl Acad Sci U S A* 97 (2000) 6640-6645.

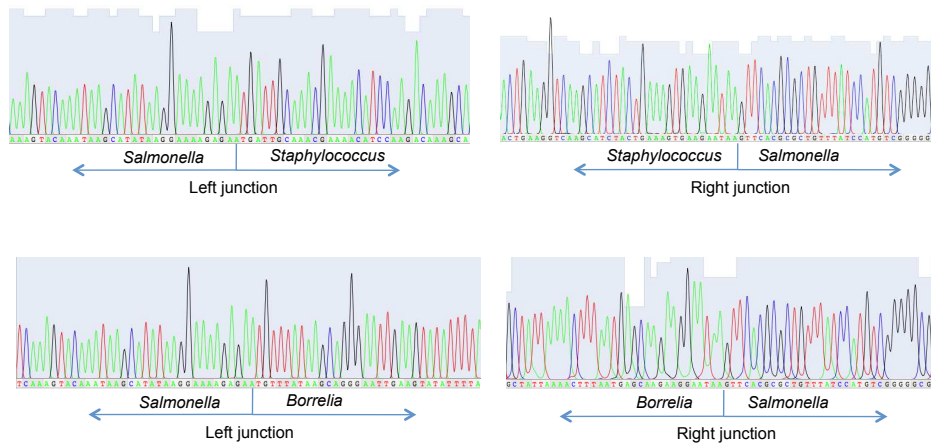
- [37] S. Uzzau, N. Figueroa-Bossi, S. Rubino, L. Bossi, Epitope tagging of chromosomal genes in *Salmonella*, Proc Natl Acad Sci U S A 98 (2001) 15264-15269.
- [38] J.E. Karlinsey, λ -Red genetic engineering in *Salmonella enterica* serovar Typhimurium, Methods Enzymol (2007) 199-209.
- [39] B.R. Bochner, H.C. Huang, G.L. Schieven, B.N. Ames, Positive selection for loss of tetracycline resistance, J Bacteriol 143 (1980) 926-933.
- [40] L. Bossi, N. Figueroa-Bossi, A small RNA downregulates LamB maltoporin in *Salmonella*, Mol Microbiol 65 (2007) 799-810.
- [41] J.H. Miller, A Short Course in Bacterial Genetics. A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1992.
- [42] K. Moon, S. Gottesman, Competition among Hfq-binding small RNAs in *Escherichia coli*, Mol Microbiol (2011).
- [43] N. Figueroa-Bossi, M. Valentini, L. Malleret, F. Fiorini, L. Bossi, Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target, Genes Dev 23 (2009) 2004-2015.
- [44] K. Moon, S. Gottesman, A PhoQ/P-regulated small RNA regulates sensitivity of *Escherichia coli* to antimicrobial peptides, Mol Microbiol 74 (2009) 1314-1330.
- [45] E. Huntzinger, S. Boisset, C. Saveanu, Y. Benito, T. Geissmann, A. Namane, G. Lina, J. Etienne, B. Ehresmann, C. Ehresmann, A. Jacquier, F. Vandenesch, P. Romby, Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression, Embo J 24 (2005) 824-835.
- [46] A. Jousset, L. Metzinger, B. Felden, On the facultative requirement of the bacterial RNA chaperone, Hfq, Trends Microbiol 17 (2009) 399-405.

Rochat *et al.*, Fig. 1Rochat *et al.*, Fig. 2

Rochat *et al.*, Fig. 3Rochat *et al.*, Fig. 4

Rochat et al., Supplementary Figure 1 (COLOR)

Figure S1. Sequence verification of *hfq* replacement strains. The regions encompassing the *hfq* genes in strains MA10741 ($\Delta hfq_{STM}::hfq_{SA}$) and MA11042 ($\Delta hfq_{STM}::hfq_{BB}$) were amplified by PCR with primers pp613 (CCCGAGGAGATTGCATA) and pp616 (GGGCTGGACTCAATGCAT) and the fragments obtained subjected to DNA sequence analysis. Only the chromatogram portions covering the left and right junctions of the heterologous *hfq* inserts are shown. Internal sequences fully matched those of *hfq_{SA}* and *hfq_{BB}* genes (data not shown).



III. DISCUSSION AND PERSPECTIVES

The work described in this thesis can be regarded from different angles. On one hand, the study deals with a basic topic like the initiation of translation and its regulation. On the other hand, the study contributes to a better characterization of processes related to bacterial physiology and metabolic regulation.

III.1. The role of ACA motifs in translation initiation

A major finding in this work was the discovery that GcvB sRNA represses expression of the *yifK* gene by countering the action of a translation enhancer embedded in its the target sequence. Having the recognition sequence of a regulatory sRNA coincide with a translation enhancer seems a particular effective way to achieve regulation and we might predict that this feature will be shared by other targets of GcvB and possibly by targets of other sRNAs.

Two ACA triplets were identified as major contributors to the translational enhancer activity. ACA motifs are found in multiple copies at other GcvB targets. We do not yet know if the enhancing effects are a general property of these motifs or if their effects somehow depend on the local sequence context. In this respect, it is important to notice that the *yifK* translation initiation signal remains relatively weak - in spite of including the enhancer element - as compared to other mRNAs studied in the laboratory. This raises the possibility that the function of the enhancer is to compensate for some weakening elements elsewhere in the initiation region. If that is the case, in other contexts, ACA motifs could play a much less important role. Experiments addressing this point are currently being performed in the laboratory. In one approach, we are testing the effects of progressively removing the four ACA triplets found at the GcvB target site in the *dppA* mRNA (in the region immediately upstream of the Shine-Dalgarno sequence) on the translation efficiency of this mRNA. In a complementary approach, we are testing if introducing ACA triplets near the Shine-Dalgarno sequence of an

mRNA that does not normally contain ACA will increase its translation efficiency. We have chosen for this analysis the mRNA of the *chiP* gene (Figuroa-Bossi et al., 2009). To produce minimal perturbation, the ACA triplets will be derived from preexisting triplets, namely UCA and AUA found within a 10 nt segment immediately upstream of *chiP* Shine-Dalgarno sequence.

The other important question that remains open concerns the mechanism by which the ACA motif improves translation initiation. Results described in this thesis (section II. 2), together with data from the Vogel's laboratory (Sharma et al., 2007) and other data in the literature (Martin-Farmer & Janssen, 1999), indicate that the enhancer effects do not require a precise positioning of ACA relative to the translation start site. This suggests that ACA does not directly participate in the architecture of the initiation complex between the mRNA and the 30 S ribosomal subunit. Rather, the ACA motif could serve as an anchoring point for some component of the 30 S subunit and stimulate the initial recognition of the initiation region. Presence of the ACA sequence in RNA aptamers selected by the SELEX procedure against to ribosomal protein S1 (Ringquist et al., 1995) makes this protein a candidate to be such component. A study comparing the binding affinities of wild-type *yifK* 5' UTR and of the mutants constructed in this study to purified S1 protein *in vitro* should allow it to verify this hypothesis.

III. 2. Role of GcvB in bacterial physiology.

In spite of extensive work, some basic questions regarding the physiological role of GcvB are still unanswered: why does GcvB inhibit amino acid uptake when bacteria grow at fast rates in rich medium and amino acids are maximally required for protein synthesis? What is the relationship between GcvB function and the glycine cleavage pathway?

A tentative answer to the former question is that GcvB activity is required to optimize the intracellular concentrations of amino acids to their ratios of occurrence in bacterial proteins. Alternatively, GcvB could balance amino acid levels with those of other macromolecular precursors such nucleotides or

carbohydrates. Most GcvB studies, including those involving transcriptomic analyses (Sharma et al., 2011, Modi *et al.*, 2011) have been performed using bacteria growing in Luria-Bertani (LB) medium. This medium contains tryptone (a pancreatic digest of casein from cow milk) and yeast extract (autodigest of the yeast *Saccharomyces cerevisiae*). Although permitting fast growth rates, this medium may not be optimized for bacterial metabolism. For example, a study by R. D'Ari group showed that LB is limiting for carbohydrates, which causes a metabolic switch to take place during the exponential phase when sugars are exhausted and bacteria begin using amino acids as carbon sources (Sezonov *et al.*, 2007). It would be interesting to study *gcvB* expression and activity under more defined growth conditions, such as a synthetic medium supplemented with all amino acids, nucleotide precursors and a carbon source. This would allow varying the ratios of some of these supplements and studying the effects on *gcvB* expression.

The fact that both the *gcvB* gene and the *gcvTHP* operon encoding the enzymes of the glycine cleavage pathway are under the control of the GcvA regulatory protein strongly suggests the existence of a physiological link between GcvB activity and the glycine cleavage system. However, this link is not obvious given the *gcvTHP* operon is maximally transcribed under conditions - glycine-supplemented minimal-glucose medium - in which the *gcvB* gene is expressed at very low level. The situation is likely reversed in LB medium where *gcvB* is maximally transcribed. Although we have been unable to find data on *gcvTHP* expression in LB in the literature, such levels are likely to be quite low since *gcvTHP* activation by GcvA is completely dependent on Lrp and the levels of this protein are drastically reduced in LB (Newman & Lin, 1995). Interestingly, GcvB activity contributes to the Lrp decrease (Sharma et al., 2011) raising the intriguing possibility that a role of GcvB is to act as an indirect negative regulator of *gcvTHP* expression. This would make physiological sense since the need for one-carbon units (products of the glycine cleavage pathway) is greatly reduced in rich medium as compared to minimal medium (see (Calvo & Matthews, 1994) and Discussion of article in section II.2).

Hoping to gather insight onto the physiological role of GcvB, in the course of this thesis, I randomly mutagenized a strain carrying a *lacZ* transcriptional fusion to the *gcvB* promoter with a tetracycline-resistance transposon and screened for insertion mutants showing increased or decreased β -galactosidase activity. I found a number of such mutants. Unfortunately, their initial characterization gave inconclusive and confusing results that discouraged further analysis as well as their description in this report.

III. 3. The intriguing HdfR(YifA)-YifE regulatory locus

In the course of this work, I found that inactivation of either *hdfR* (*yifA*) or *yifE*, two adjacent genes in opposite orientation results in a significant increase in *yifK* expression (see section II.3). These findings suggest that the products of these genes directly or indirectly inhibit *yifK* expression. The HdfR protein is a poorly known transcriptional regulator previously shown to repress the flagellar master operon *flhDC* (acting in concert with H-NS) in *E. coli* (Ko & Park, 2000) and to activate the fimbrial operon *std* in *Salmonella* (Jakomin et al., 2008). It therefore seems possible that HdfR represses *yifK* transcription by binding to the *yifK* promoter. The independent identification of several mutations causing increased *yifK* expression, which map in the region upstream from the -35 box of the *yifK* promoter suggest that this could be the site of HdfR binding.

In contrast, nothing can be said of YifE since no reports on this protein exist and its sequence, although highly conserved, does not show similarity to any known protein in databases. The YifE implication in *yifK* regulation is therefore intriguing and encourages further investigation. In particular, it will be important to determine if YifE acts on *yifK* directly, perhaps as part of a complex with HdfR, or indirectly; for example if required for *hdfR* expression. An alternative possibility is that YifE is the sole regulator of *yifK* and that the role of HdfR is to activate *yifE* expression. The construction of *lacZ* fusions to either of these genes and the measurement of β -galactosidase activity in cells carrying the *hdfR-lacZ* fusion combined with wild-type or mutant *yifE*, or,

viceversa, carrying *yifE-lacZ* combined with wild-type or mutant *hdfR*, should help distinguishing between these hypotheses. Finally, the construction of epitope-tagged derivatives of both proteins can allow testing the possible interaction of the two proteins by co-immuno precipitation techniques as already done in the laboratory to study repressor-antirepressor interactions (Lemire et al., 2011). These experiments should not take too long to be completed and the results could lead to an additional publication from my thesis work.

IV. MATERIALS AND METHODS

All strains used in this study are derivatives of *S. enterica* serovar Typhimurium. Strains that were used for the article in section II.2 are listed in Table 1 in that section (page). Additional strains are listed in Table 2. The bacteria were cultured in LB broth (Bertani, 2004) solidified by the addition of 1.5% Difco agar when needed. When appropriate, the LB medium was supplemented with 0.2% arabinose. Antibiotics were included at the following final concentrations: chloramphenicol, $10 \mu\text{g mL}^{-1}$; kanamycin monosulfate, $50 \mu\text{g mL}^{-1}$; sodium ampicillin, $75 \mu\text{g mL}^{-1}$; spectinomycin dihydrochloride, $80 \mu\text{g mL}^{-1}$; and tetracycline hydrochloride, $25 \mu\text{g mL}^{-1}$. LB plates containing $40 \mu\text{g mL}^{-1}$ 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-Gal) (Sigma-Aldrich) or MacConkey agar plates containing 1% lactose (Macconkey, 1905) were used to monitor *lacZ* expression in bacterial colonies. Liquid cultures were grown in New Brunswick gyratory shakers and growth was monitored by measuring the optical density at 600 nm with a Shimadzu UV-mini 1240 spectrophotometer.

Genetic Techniques

Generalized Transduction

P22 is a temperate phage that infects *Salmonella* by binding to the O-antigen, part of the lipopolysaccharide on the outer membrane. After infection, P22 circularizes by recombination between terminal redundancies at each end of the phage DNA. During lytic growth, the circular genome of P22 initially undergoes several rounds of θ (tetha) replication, then changes to rolling-circle replication. Rolling circle replication produces long concatemers of double stranded P22 DNA. These concatemers are packaged into phage heads by a "headful" mechanism: packaging is initiated at a specific sequence on the DNA called a *pac* site, then a nuclease moves down the concatemer, cutting every 48 Kb (Casjens and Hayden, 1988). Since the P22 genome is only 44 Kb, this yields the terminal redundancy at the ends of P22 (Susskind

and Botstein, 197). This linear double stranded DNA is packaged into new phage particles. When the cell lyses, it releases 50-100 new phage.

There are sequences in the *Salmonella* chromosome that are homologous to the P22 *pac* site. When the phage infects bacteria, occasionally the P22 nuclease cuts one of these chromosomal sites and packages 48 Kb (Casjens & Hayden, 1988) fragments of chromosomal DNA into P22 phage heads. The P22 particles carrying chromosomal DNA (transducing particles) can inject this DNA into a new host. The DNA can then recombine into the chromosome by homologous recombination. In order to increase the efficiency of the process, most laboratories doing P22 transduction use P22 HT105/1 int-201, a P22 mutant that is very useful for generalized transduction (Schmieger, 1972). This phage has a high transducing (HT) frequency due to a nuclease with less specificity for the *pac* sequence. About 50% of the P22 HT phage heads carry random transducing fragments of chromosomal DNA ((Schmieger, 1972)). The *int* mutation prevents formation of stable lysogens.

Transduction Procedure

Typically, P22 lysates were used at a 1:50 dilution, mixed with aliquots from overnight cultures of recipient bacteria in a 1:2 ratio, and incubated for 30 min at 37°C prior to being plated on selective media. Transductant colonies were purified by two sequential passages on selective plates and verified to be free of phage by streaking on Evans Blue Uranine plates(Bochner, 1984).

Plasmids and DNA templates.

Plasmids used as PCR templates for λ Red-mediated recombination included pKD3, pKD4 and pKD13(Datsenko & Wanner, 2000). Plasmid pCP20, carrying Flp recombinase(Cherepanov & Wackernagel, 1995), was used as needed, to remove antibiotic resistance cassettes flanked by FRT sites or to promote integration of plasmid pCE40. The latter was used to create chromosomal *lacZ* fusions to genes of interest(Ellermeier et al., 2002). Chromosomal DNA from strain MA3397 (Figuroa-Bossi et al., 1997) was

used as template for amplification of the *tetAR* module in the first step of the scarless recombineering procedure (see below).

Chromosomal engineering: λ Red recombination method

Chromosomal engineering was carried out by the λ Red recombination method according to the procedure of Datsenko and Wanner (Datsenko & Wanner, 2000). Donor DNA fragments were generated by PCR using plasmid or chromosomal DNA templates. Oligonucleotides used as primers in these experiments are presented in Table S1 (page 76) and in Table 3. Amplified fragments were electroporated into appropriate strains harboring the conditionally replicating plasmid pKD46, which carries a λ *red* operon under the control of the P^{BAD} promoter (Datsenko & Wanner, 2000). Bacteria carrying pKD46 were grown at 30°C in the presence of ampicillin and exposed to arabinose (10 mM) for 3 hours prior to preparation of electrocompetent cells. Electroporation was carried out using a Bio-Rad MicroPulser under the conditions specified by the manufacturer. Recombinant colonies were selected on LB plates containing the appropriate antibiotic. Constructs were verified by PCR and/or DNA sequencing. When needed, the antibiotic resistance cassette was excised by transforming strains with plasmid pCP20, which encodes the Flp recombinase (Cherepanov & Wackernagel, 1995). Scarless recombination experiments were performed as described in the Materials and Methods of the Article in section II.2.

Variations on λ Red

Construction of dppA-lacZY fusions

The procedure involved the insertion of the *aph* (kan^R) module of plasmid pKD13 in the *dppA* gene, and Flp recombinase-mediated conversion of the insert to a *lacZY* fusion with plasmid pCE40 as described (Ellermeier et al., 2002). Primer extension sequences were chosen so as to place the fusion boundary within signal peptide-encoding segments. Disruption of the signal sequence was intended to prevent deleterious effects that might result from translocation of the β -galactosidase moiety of the hybrid proteins across cell

membranes. Two fusions were constructed with the *lacZY* sequence at slightly different positions within *dppA*: *dppA147::lacZY* carrying *lacZ* fused to the 16th codon of *dppA* (primers ppK77/ppK79; Table 3) and *dppA148::lacZY* carrying *lacZ* fused to the 9th codon of *dppA* (ppK78/ppK79; Table 3). The two fusions behaved similarly and were used interchangeably throughout this study.

Random PCR mutagenesis

Chromosomal DNAs encompassing the *gcvB* gene, or the promoter proximal region of the *yifK-lacZ* fusion, were amplified by PCR under error-prone conditions. In both cases the amplified region included a suitably positioned *cat* marker for selection (Fig. 25 and 26).

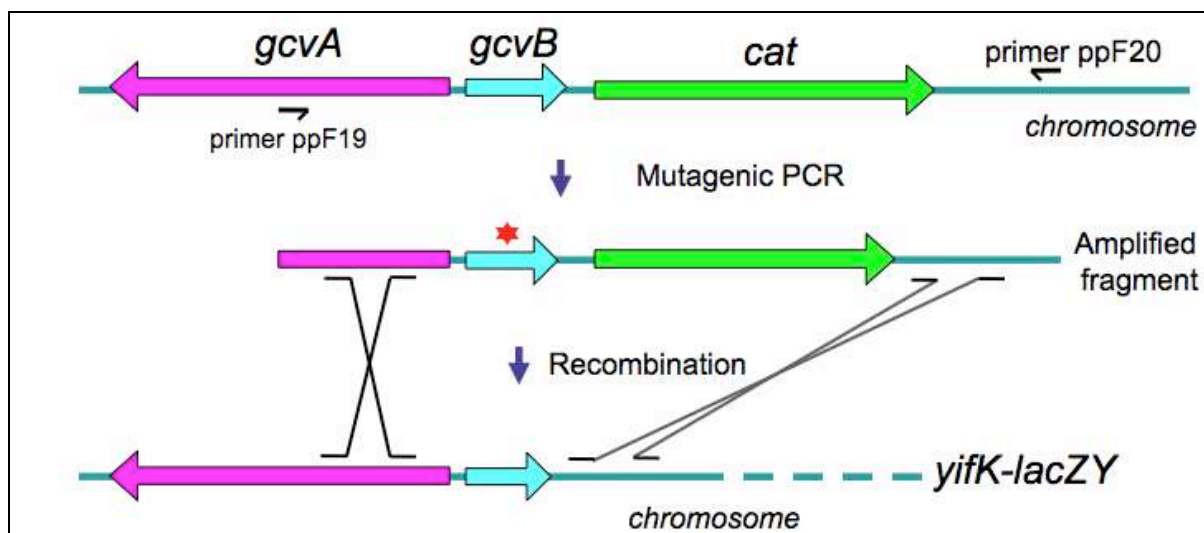
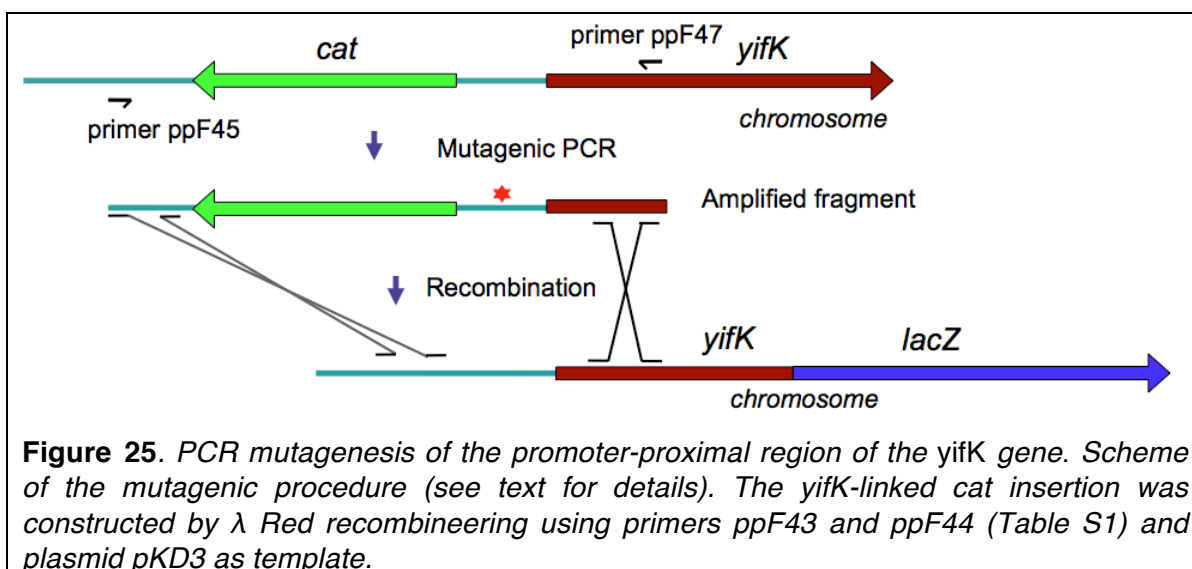


Figure 25. PCR mutagenesis of the *gcvB* gene. Scheme of the mutagenic procedure (see text for details). The *gcvB*-linked *cat* insertion was constructed by λ Red recombineering using primers ppF17 and ppF18 (Table S1) and plasmid pKD3 as template.



Briefly, amplification was carried out with Taq polymerase (New England Biolabs), in the absence of Pfu polymerase, in 10mM Tris-HCl pH 9.0, 50 mM KCl, 1.5 mM MgCl₂ 0.1% Triton X-100, 0.2 mg/ml BSA (Qbiogene). The amplification program included 10 cycles of differential temperature decrease (0.1°C/sec) from 95°C to 64°C (annealing temperature) followed by 15 cycles with annealing at 62°C (no ramp step). The elongation step was set at 72°C for 45 seconds in all cycles. For *gcvB* mutagenesis, the template was chromosomal DNA from strain MA10307 and the primers ppF19 and ppF20 (Table S1, page 76). For mutagenesis of *yifK* 5'UTR, template was chromosomal DNA from strain MA10308 and the primers were ppF45 and ppF47. In both amplification experiments, the priming sites were chosen so as to have approximately 0.5 Kb at both ends of the DNA products providing homology for recombination. The fragments obtained (1969 bp and 2034 bp, respectively) were used to transform strain MA10280 (*yifK*::MudK / pKD46) and Cm^R recombinants selected on MacConkey-lactose plate supplemented with chloramphenicol. Deep red colonies were picked and purified by streaking on selective plates. The mutagenized region was PCR-amplified ppF19/pp595, *GcvB*; ppF45/ppF62 *yifK* 5' UTR) with a 3:1 Taq/Pfu polymerase mixture and the resulting fragment subjected to DNA sequence analysis.

“Scarless” DNA recombineering.

Scarless modification of chromosomal DNA sequences at the single base-pair level was achieved with a two-step recombineering procedure adapted from (Karlinsky, 2007). The procedure involves: *i*) inserting a *tetAR* module (produced by PCR) at the chromosomal site to be modified and *ii*) replacement of the *tetAR* module by a DNA fragment carrying the desired change through positive selection tetracycline-sensitive recombinants (Bochner et al., 1980). Typically, the DNA fragment in the second step was also obtained by PCR using oligonucleotides with complementary sequences at their 3' ends priming DNA synthesis on each other (“reciprocal priming”)

β-Galactosidase Assay (adapted from Miller (Miller, 1992))

For exponential growth measurements, an overnight culture was diluted 1:100 and grown until O.D.600 around 0.4. From this same 1 ml culture, 0.1 ml (V_C) were taken and placed in glass tube with Z buffer to make up 1 ml final. One drop of Toluene was added with a Pasteur pipette and each tube was immediately vortexed for 30 seconds. Toluene was evaporated at 42°C for 2h. Tubes were placed in a water bath at 28°C for 5min. Reaction was started by adding 0.2ml *ortho*-Nitrophenyl-β-galactoside (ONPG, 4mg/ml, freshly made). Start time (t_0) for each tube was recorded. After sufficient yellow color developed, the reaction was stopped with 0.5 ml of 1M Na₂CO₃ and final time (t_f) recorded. In order to avoid the contribution of bacteria to the O.D.420, the bacteria were pelleted and the supernatant used to read OD420. β-galactosidase activity was calculated as as:

$$\frac{1000 \times \text{O.D.420}}{(t_f - t_0) \times V_C \times \text{O.D.600}}$$

Molecular Biology and RNA structural techniques***5'-end mapping of yifK mRNA by primer extension.***

The reverse transcriptase reaction (Superscript II from Invitrogen) was carried out using 5 μ g of RNA from strains MA8020, MA11781, MA11782 and MA11783 with primer ppF49 labeled at the 5' end with 32 P. The same primer was used to generate the DNA sequence ladder. The sequencing template was a DNA fragment obtained by PCR amplification of chromosomal DNA from strain MA11780 (Table S1) with primers ppF84 and ppF85. (For primer sequences see Table 3). Reactions were performed with the fmol DNA Cycle Sequencing System from Promega, according to the manufacturer's protocol.

RNA extraction and Northern analysis

RNA was prepared by the acid-hot-phenol method from exponentially growing cells (OD_{600} of 0.35) as previously described (Figuroa *et al.*, 1991). Total RNA was estimated from the value of the OD at 260 nm. For Northern analysis, 7.5 mg of total RNA were separated under denaturing conditions either by 8M urea-8% polyacrylamide in TBE (Tris-borate-EDTA pH 8.3) buffer or by 2.2 M formaldehyde-1.3% agarose gel electrophoresis in MOPS ([N-morpholino] propanesulfonic acid-sodium acetate-EDTA pH7.0) buffer. For acrylamide gels, transfer of the RNA onto Hybond-N⁺ membrane (Amersham), was performed with a semidry electrotransfer apparatus (Transblot SD; BioRad); for agarose gels, transfer to the same support was done using a vacuum blotter (Boekel/Appligene) after mild denaturation treatment in 50 mM NaOH, 10 mM NaCl. RNA was crosslinked to membrane by UV irradiation (Statagene UV Stratalinker 2400). Membranes were hybridized to probes as follows: 5 pmol of oligonucleotides ppF16, ppG42, ppH27 and pp813 (Table S1) were 5' end-labelled using 10 U of T4 polynucleotide kinase (New England Biolabs) and 30 μ Ci of [γ - 32 P]ATP (3,000 mCi mmol⁻¹, Amersham). Unincorporated radioactivity was eliminated by passage through Micro-Bio Spin 6 chromatography columns (BioRad). Hybridization was carried out in Ambion Oligonucleotide Hybridization Buffer at 45-50°C following Ambion's protocol. RNA was analysed by Phosphorimaging using ImageQuant software.

In vitro RNA synthesis

Templates for *in vitro* transcription with T7 polymerase were generated by PCR from genomic DNA. To produce the *yifK* 5' UTR template, DNA was amplified with primers ppl22 and ppl23 from strain MA8020. To make the wild-type *GcvB* template, primers were ppJ01 and ppJ03. *In vitro* transcription was performed with the MegashortscriptT7 kit (Ambion AM1354) according to the manufacturer's protocol. Transcription products were an 137 nucleotide RNA corresponding to the first 134 residues of *yifK* mRNA plus a 3-nucleotide extension from the T7 promoter (GGG) at the 5' end, and full-length *GcvB* with an extra G at its 5' end. After incubation for 2 hrs at 37°C, DNase was added and incubation continued for further 15 min. The samples were treated with phenol and the RNA precipitated at -20°C overnight with sodium acetate-ethanol-glycogen. RNA was recovered by centrifugation, resuspended in water and quantified by nanodrop reading.

5'-end mapping of yifK mRNA by primer extension.

The reverse transcriptase reaction (Superscript II from Invitrogen) was carried out using 5 μ g of RNA with primer ppF49 labeled at the 5' end with ³²P. The same primer was used to generate the DNA sequence ladder. The sequencing template was a DNA fragment obtained by PCR amplification of chromosomal DNA with primers ppl22 and ppF62 (Table S1). Reactions were performed with the fmol DNA Cycle Sequencing System from Promega, according to the manufacturer's protocol.

In vitro translation

In vitro coupled transcription/translation was performed using New England Biolabs' PURExpress *In vitro* Protein Synthesis kit (NEB #E6800) according to the manufacturer instructions, as described in the article in section II.2 of this thesis manuscript. DNA templates were constructed using the backbone of plasmid DHFR provided with the kit. To this end, plasmid DNA was digested with XbaI and PstI. PCR amplification from chromosome of strains MA11798

and MA11799 with primers ppH29 and ppH30, yielded fragments carrying a XbaI and a PstI site at the ends. These fragments, as well as DHFR plasmid DNA, were digested with XbaI and PstI. Following appropriate ligation, two plasmids carrying wt or mutant yifK 5' UTR-*cat*-3XFLAG chimeric constructs under the control of the T7 promoter were obtained and used for in vitro transcription/translation.

Table 2. *Salmonella enterica* strains used in this work (complements Table 1, page 61).

Strain ^a	Genotype	Source or reference
4YA	<i>yifK</i> [A48U,C49G,A50U,C51G] <i>yifK87</i> ::MudK <i>wecG144</i> :: <i>cat</i>	this work
YQ44	<i>yifK87</i> ::MudK <i>gcvB</i> [G83C,U84A,G85C,U86A] <i>ygdI143</i> :: <i>cat</i>	this work
YQ47	<i>yifK</i> [A48U,C49G,A50U,C51G] <i>yifK87</i> ::MudK <i>wecG144</i> :: <i>cat</i> <i>ygdI143</i> :: <i>cat</i> <i>gcvB</i> [G83C,U84A,G85C,U86A]	this work
YQ328	<i>dppA147</i> :: <i>lacZY-aph</i> (Kan ^R)	this work
YQ337	<i>dppA147</i> :: <i>lacZY-aph</i> (Kan ^R) <i>gcvB</i> [U86A,U87C,U88A] <i>ygdI143</i> :: <i>cat</i>	this work
YQ340	<i>dppA147</i> :: <i>lacZY-aph</i> (Kan ^R) Δ <i>gcvB141</i> :: <i>cat</i>	this work
YQ111	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>gcvB141</i> :: <i>scar</i>	this work
YQ117	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>gcvB141</i> :: <i>scar</i> <i>lrp-42</i> ::Tn10dCm	this work
YQ233	<i>yifK87</i> ::MudK Δ <i>gcvB142</i> :: <i>aadA</i> (Spec ^R) <i>lrp-42</i> ::Tn10dCm	this work
YQ235	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>hdfR</i> :: <i>tetAR</i>	this work
YQ236	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>yifE</i> :: <i>tetAR</i>	this work
YQ820	<i>yifK87</i> ::MudK Δ <i>gcvB141</i> :: <i>cat</i> Δ <i>hdfR</i> :: <i>tetAR</i>	this work
YQ821	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>yifE</i> :: <i>tetAR</i>	this work
YQ822	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>gcvB141</i> :: <i>cat</i> Δ <i>hdfR</i> :: <i>tetAR</i> <i>lrp-42</i> ::Tn10dCm	this work
YQ823	<i>yifK87</i> ::MudK Δ <i>gcvB142</i> :: <i>aadA</i> (Spec ^R) Δ <i>hdfR</i> :: <i>tetAR</i> <i>lrp-42</i> ::Tn10dCm	this work
YQ824	<i>yifK87</i> ::MudK Δ <i>gcvB141</i> :: <i>cat</i> Δ <i>yifE</i> :: <i>tetAR</i>	this work
YQ825	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>gcvB141</i> :: <i>cat</i> Δ <i>yifE</i> :: <i>tetAR</i>	this work
YQ826	<i>yifK</i> [C49U,A50C] <i>yifK87</i> ::MudK Δ <i>gcvB141</i> :: <i>cat</i> Δ <i>yifE</i> :: <i>tetAR</i> <i>lrp-42</i> ::Tn10dCm	this work
YQ826	<i>yifK87</i> ::MudK Δ <i>gcvB142</i> :: <i>aadA</i> (Spec ^R) Δ <i>yifE</i> :: <i>tetAR</i> <i>lrp-42</i> ::Tn10dCm	this work
MA8024	<i>aphA</i> ::MudK	this work

Table 3. Oligonucleotides used in this work (complements Table S1, page 76).

name	Sequence (5'-3')
pp813	GCGGAGGCTAGGGAGAGAGG
pp595	ACCAGCTCACCGTCTTTC
pp615	GCGGGCAAGAATGTGAAT
ppF16	GGCAATCAATTCAATATGACGAGCTTCCAGCCCACGC
ppF19	TGCTCGCGTAAAGCTCAAAT
ppF20	CGCAAACCTGGCTACGATGAA
ppF45	CGTGGTGC GTTCAATACGTA
ppF47	CGCCATCCACATAAACCCAGT
ppF62	CGCCCATTTCAGCGTACT
ppF84	CGCCCGGTAGTGATCTTATT
ppF85	CGAGGGCAATCAATTCAATA
ppF90	CTGTAGCTCCGGTTTTTCTCTGCCATAAATCCTCATGTCCACATTGCAACGGTTCGGGTAGTC
ppF91	CTGATGGGCTTTTGGCTTACGGTTGTGATGTTGTGTTGTTCCACATTGCAATTGGTCTGCGATTA
ppG42	AGGGCAGCGCTCTATCCAGCTGAGCTACGGGCGCTTAG
ppH27	TGAAACTTTTTACGCGGTATTAACCACCGCAGCTCAAGCACCTAAATAAA
ppi22	GGTAATACGACTCACTATAGGGCAATCGATGTACCCATAACAATAACCG
ppi23	CCCGAGGGCAATCAATTCAA
ppJ01	GAAATTAATACGACTCACTATAGGACTTCCTGAGCCGGAACG
ppJ03	AGACAGGGTAAATGTACAGGAAGTG
ppJ12	TGAGCTTTCGTATCAGGCAACTGGAGAATCAACTGGGCGTTTAAGACCCACTTTCACATT
ppJ13	TTTTGCAGCCAAATAGCGTATAGTGGCCGTGAAAGCGTTGCGCTAAGCACTTGTCTCCT
ppJ14	AATCATCAGCTTCGGTGTAGTCTTCTGCGCCTTCAACTTGCTAAGCACTTGTCTCCTG
ppJ15	TTCACGATTAAGAGGCACAACCTGCTTGAGCGTCATGGTCTTAAGACCCACTTTCACATT
ppJ16	GGGCGATAGACAGGTTGTTAATTC
ppJ17	CGCCTGCCACGTATTCATCA
ppK77	TATTTCTTGAAGAAGTCAGGGATGCTGAAGCTTGGTTTGAAGATCCGTCGACCTGCAGTTC
ppK78	ATTGGAGCAGAAGAATGAGTATTTCTTGAAGAAGTCAGGGATCCGTCGACCTGCAGTTC
ppK79	TTTATAGAGCTCAATACGCTTGTGTTGTTGTCATCGGTGCGATGTGTAGGCTGGAGCTGCTT
ppM29	AATTTATATTTAACTATTCTAGACAATCGATGTACCCATAACAATAACC
ppM30	ATTTATATTTCCGGCCGCTCGAGCTGCAGTGGCGCTTACTATTTATCGTCG

Table 4. Relevant alleles constructed in this work (complements Table S2, page 78).

Primer pair	Template ^a	Allele	Description
ppF18 / ppF91	MA11779	<i>gcvB</i> [83-86 CACA]	Replacement of GcvB segment between +86 and +88 (GUGU to CACA)
ppF43 / ppF90	MA11780	<i>yifK</i> [48-51 UGUG]	Replacement of yifk segment between +48 and +51 (CACA to GUGU)
ppJ12 / ppJ13	pNK2883	$\Delta hdfR::tetAR$	<i>tetAR</i> module replaces <i>hdfR</i> gene
ppJ14 / ppJ15	pNK2883	$\Delta yifE::tetAR$	<i>tetAR</i> module replaces <i>yifE</i> gene
ppK77 / ppK79	pKD13	<i>dppA147::aph</i>	<i>aph</i> gene (Kan ^R) replaces <i>dppA</i> (starting at codon 16)
ppK78 / ppK79	pKD13	<i>dppA148::aph</i>	<i>aph</i> gene (Kan ^R) replaces <i>dppA</i> (starting at codon 9)

^aTemplate DNA was from chromosome (“MA” numbers identify the *Salmonella* strains used as source; described in the text) or from plasmids pNK2883.(Kleckner *et al.*, 1991) or pKD3 (Datsenko & Wanner, 2000).

V. BIBLIOGRAPHY

- Achsel, T., H. Brahm, B. Kastner, A. Bachi, M. Wilm & R. Luhrmann, (1999) A doughnut-shaped heteromer of human Sm-like proteins binds to the 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *The EMBO journal* **18**: 5789-5802.
- Arluisson, V., P. Derreumaux, F. Allemand, M. Folichon, E. Hajnsdorf & P. Regnier, (2002) Structural Modelling of the Sm-like Protein Hfq from Escherichia coli. *Journal of molecular biology* **320**: 705-712.
- Baek, C.H., H.Y. Kang, K.L. Roland & R. Curtiss, 3rd, (2011) Lrp acts as both a positive and negative regulator for type 1 fimbriae production in Salmonella enterica serovar Typhimurium. *PLoS one* **6**: e26896.
- Baek, C.H., S. Wang, K.L. Roland & R. Curtiss, 3rd, (2009) Leucine-responsive regulatory protein (Lrp) acts as a virulence repressor in Salmonella enterica serovar Typhimurium. *Journal of bacteriology* **191**: 1278-1292.
- Baker, J.L., N. Sudarsan, Z. Weinberg, A. Roth, R.B. Stockbridge & R.R. Breaker, (2012) Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* **335**: 233-235.
- Balbontin, R., F. Fiorini, N. Figueroa-Bossi, J. Casadesus & L. Bossi, (2010) Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in Salmonella enterica. *Molecular microbiology* **78**: 380-394.
- Beisel, C.L. & G. Storz, (2011) The base-pairing RNA spot 42 participates in a multioutput feedforward loop to help enact catabolite repression in Escherichia coli. *Mol Cell* **41**: 286-297.
- Beltran, P., J.M. Musser, R. Helmuth, J.J. Farmer, 3rd, W.M. Frerichs, I.K. Wachsmuth, K. Ferris, A.C. McWhorter, J.G. Wells, A. Cravioto & et al., (1988) Toward a population genetic analysis of Salmonella: genetic diversity and relationships among strains of serotypes S. choleraesuis, S. derby, S. dublin, S. enteritidis, S. heidelberg, S. infantis, S. newport, and S. typhimurium. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 7753-7757.
- Bertani, G., (2004) Lysogeny at mid-twentieth century: P1, P2, and other experimental systems. *Journal of bacteriology* **186**: 595-600.
- Blattner, F.R., G. Plunkett, 3rd, C.A. Bloch, N.T. Perna, V. Burland, M. Riley, J. Collado-Vides, J.D. Glasner, C.K. Rode, G.F. Mayhew, J. Gregor, N.W. Davis, H.A. Kirkpatrick, M.A. Goeden, D.J. Rose, B. Mau & Y. Shao, (1997) The complete genome sequence of Escherichia coli K-12. *Science* **277**: 1453-1462.
- Bochner, B.R., (1984) Curing bacterial cells of lysogenic viruses by using UCB indicator plates. *Biotechniques* **2**: 234-240.
- Bochner, B.R., H.C. Huang, G.L. Schieven & B.N. Ames, (1980) Positive selection for loss of tetracycline resistance. *Journal of bacteriology* **143**: 926-933.
- Boni, I.V., D.M. Isaeva, M.L. Musychenko & N.V. Tzareva, (1991) Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic acids research* **19**: 155-162.
- Bossi, L. & N. Figueroa-Bossi, (2007) A small RNA downregulates LamB maltoporin in Salmonella. *Molecular microbiology* **65**: 799-810.

- Bossi, L., D. Maloriol & N. Figueroa-Bossi, (2008) Porin biogenesis activates the sigma(E) response in Salmonella hfq mutants. *Biochimie* **90**: 1539-1544.
- Bossi, L., A. Schwartz, B. Guillemardet, M. Boudvillain & N. Figueroa-Bossi, (2012) A role for Rho-dependent polarity in gene regulation by a noncoding small RNA. *Genes & development* **26**: 1864-1873.
- Bossi, L. & D.M. Smith, (1984) Conformational change in the DNA associated with an unusual promoter mutation in a tRNA operon of Salmonella. *Cell* **39**: 643-652.
- Boudvillain, M., N. Figueroa-Bossi & L. Bossi, (2013) Terminator still moving forward: expanding roles for Rho factor. *Current opinion in microbiology* **16**: 118-124.
- Bourne, H.R., D.A. Sanders & F. McCormick, (1991) The GTPase superfamily: conserved structure and molecular mechanism. *Nature* **349**: 117-127.
- Bouvier, M., C.M. Sharma, F. Mika, K.H. Nierhaus & J. Vogel, (2008) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Molecular cell* **32**: 827-837.
- Brantl, S., (2002) Antisense-RNA regulation and RNA interference. *Biochimica et biophysica acta* **1575**: 15-25.
- Brennan, R.G. & T.M. Link, (2007) Hfq structure, function and ligand binding. *Curr Opin Microbiol* **10**: 125-133.
- Brenner, F.W., R.G. Villar, F.J. Angulo, R. Tauxe & B. Swaminathan, (2000) Salmonella nomenclature. *Journal of clinical microbiology* **38**: 2465-2467.
- Brink, M.F., M.P. Verbeet & H.A. de Boer, (1995) Specialized ribosomes: highly specific translation in vivo of a single targetted mRNA species. *Gene* **156**: 215-222.
- Brinkman, A.B., T.J. Ettema, W.M. de Vos & J. van der Oost, (2003) The Lrp family of transcriptional regulators. *Molecular microbiology* **48**: 287-294.
- Brouns, S.J., M.M. Jore, M. Lundgren, E.R. Westra, R.J. Slijkhuis, A.P. Snijders, M.J. Dickman, K.S. Makarova, E.V. Koonin & J. van der Oost, (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**: 960-964.
- Calvo, J.M. & R.G. Matthews, (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. *Microbiological reviews* **58**: 466-490.
- Camacho, E.M. & J. Casadesus, (2002) Conjugal transfer of the virulence plasmid of Salmonella enterica is regulated by the leucine-responsive regulatory protein and DNA adenine methylation. *Mol Microbiol* **44**: 1589-1598.
- Camacho, E.M. & J. Casadesus, (2005) Regulation of traJ transcription in the Salmonella virulence plasmid by strand-specific DNA adenine hemimethylation. *Molecular microbiology* **57**: 1700-1718.
- Canonaco, M.A., R.A. Calogero & C.O. Gualerzi, (1986) Mechanism of translational initiation in prokaryotes. Evidence for a direct effect of IF2 on the activity of the 30 S ribosomal subunit. *FEBS letters* **207**: 198-204.
- Caron, M.P., D.A. Lafontaine & E. Masse, (2010) Small RNA-mediated regulation at the level of transcript stability. *RNA biology* **7**: 140-144.
- Carte, J., R. Wang, H. Li, R.M. Terns & M.P. Terns, (2008) Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes & development* **22**: 3489-3496.
- Casjens, S. & M. Hayden, (1988) Analysis in vivo of the bacteriophage P22 headful nuclease. *Journal of molecular biology* **199**: 467-474.

- Chen, H., M. Bjerknes, R. Kumar & E. Jay, (1994) Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of Escherichia coli mRNAs. *Nucleic acids research* **22**: 4953-4957.
- Chen, S. & J.M. Calvo, (2002) Leucine-induced dissociation of Escherichia coli Lrp hexadecamers to octamers. *Journal of molecular biology* **318**: 1031-1042.
- Cherepanov, P.P. & W. Wackernagel, (1995) Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**: 9-14.
- Coornaert, A., C. Chiaruttini, M. Springer & M. Guillier, (2013) Post-transcriptional control of the Escherichia coli PhoQ-PhoP two-component system by multiple sRNAs involves a novel pairing region of GcvB. *PLoS genetics* **9**: e1003156.
- Cromie, M.J., Y. Shi, T. Latifi & E.A. Groisman, (2006) An RNA sensor for intracellular Mg(2+). *Cell* **125**: 71-84.
- Crosa, J.H., D.J. Brenner, W.H. Ewing & S. Falkow, (1973) Molecular relationships among the Salmonelleae. *Journal of bacteriology* **115**: 307-315.
- Cui, Y., Q. Wang, G.D. Stormo & J.M. Calvo, (1995) A consensus sequence for binding of Lrp to DNA. *Journal of bacteriology* **177**: 4872-4880.
- Darfeuille, F., C. Unoson, J. Vogel & E.G. Wagner, (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Molecular cell* **26**: 381-392.
- Datsenko, K.A. & B.L. Wanner, (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 6640-6645.
- De Lay, N., D.J. Schu & S. Gottesman, (2013) Bacterial small RNA-based negative regulation: Hfq and its accomplices. *The Journal of biological chemistry* **288**: 7996-8003.
- de los Rios, S. & J.J. Perona, (2007) Structure of the Escherichia coli leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. *Journal of molecular biology* **366**: 1589-1602.
- Deltcheva, E., K. Chylinski, C.M. Sharma, K. Gonzales, Y. Chao, Z.A. Pirzada, M.R. Eckert, J. Vogel & E. Charpentier, (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**: 602-607.
- Desnoyers, G. & E. Masse, (2012) Noncanonical repression of translation initiation through small RNA recruitment of the RNA chaperone Hfq. *Genes & development* **26**: 726-739.
- Diestra, E., B. Cayrol, V. Arluison & C. Risco, (2009) Cellular electron microscopy imaging reveals the localization of the Hfq protein close to the bacterial membrane. *PLoS one* **4**: e8301.
- Ding, Y., B.M. Davis & M.K. Waldor, (2004) Hfq is essential for Vibrio cholerae virulence and downregulates sigma expression. *Molecular microbiology* **53**: 345-354.
- Doolittle, R.F., D.F. Feng, S. Tsang, G. Cho & E. Little, (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* **271**: 470-477.
- Dreyfus, M., (1988) What constitutes the signal for the initiation of protein synthesis on Escherichia coli mRNAs? *J Mol Biol* **204**: 79-94.
- Dubey, A.K., C.S. Baker, T. Romeo & P. Babitzke, (2005) RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction. *RNA* **11**: 1579-1587.

- Ellermeier, C.D., A. Janakiraman & J.M. Slauch, (2002) Construction of targeted single copy lac fusions using lambda Red and FLP-mediated site-specific recombination in bacteria. *Gene* **290**: 153-161.
- Esakova, O. & A.S. Krasilnikov, (2010) Of proteins and RNA: the RNase P/MRP family. *RNA* **16**: 1725-1747.
- Fantappie, L., M.M. Metruccio, K.L. Seib, F. Oriente, E. Cartocci, F. Ferlicca, M.M. Giuliani, V. Scarlato & I. Delany, (2009) The RNA chaperone Hfq is involved in stress response and virulence in *Neisseria meningitidis* and is a pleiotropic regulator of protein expression. *Infection and immunity* **77**: 1842-1853.
- Figueroa, N., N. Wills & L. Bossi, (1991) Common sequence determinants of the response of a prokaryotic promoter to DNA bending and supercoiling. *The EMBO journal* **10**: 941-949.
- Figueroa-Bossi, N., E. Coissac, P. Netter & L. Bossi, (1997) Unsuspected prophage-like elements in *Salmonella typhimurium*. *Molecular microbiology* **25**: 161-173.
- Figueroa-Bossi, N., S. Lemire, D. Maloriol, R. Balbontin, J. Casadesus & L. Bossi, (2006a) Loss of Hfq activates the sigmaE-dependent envelope stress response in *Salmonella enterica*. *Molecular microbiology* **62**: 838-852.
- Figueroa-Bossi, N., S. Lemire, D. Maloriol, R. Balbontín, J. Casadesús & L. Bossi, (2006b) Loss of Hfq activates the sigmaE-dependent envelope stress response in *Salmonella enterica*. *Mol Microbiol* **62**: 838-852.
- Figueroa-Bossi, N., M. Valentini, L. Malleret, F. Fiorini & L. Bossi, (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev* **23**: 2004-2015.
- Folichon, M., V. Arluison, O. Pellegrini, E. Huntzinger, P. Regnier & E. Hajnsdorf, (2003) The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. *Nucleic acids research* **31**: 7302-7310.
- Franze de Fernandez, M.T., L. Eoyang & J.T. August, (1968) Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219**: 588-590.
- Franze de Fernandez, M.T., W.S. Hayward & J.T. August, (1972) Bacterial proteins required for replication of phage Q ribonucleic acid. Purification and properties of host factor I, a ribonucleic acid-binding protein. *The Journal of biological chemistry* **247**: 824-831.
- Galanis, E., D.M. Lo Fo Wong, M.E. Patrick, N. Binsztein, A. Cieslik, T. Chalermchikit, A. Aidara-Kane, A. Ellis, F.J. Angulo & H.C. Wegener, (2006) Web-based surveillance and global *Salmonella* distribution, 2000-2002. *Emerging infectious diseases* **12**: 381-388.
- Garneau, J.E., M.E. Dupuis, M. Villion, D.A. Romero, R. Barrangou, P. Boyaval, C. Fremaux, P. Horvath, A.H. Magadan & S. Moineau, (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**: 67-71.
- Geissmann, T.A. & D. Touati, (2004) Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *The EMBO journal* **23**: 396-405.
- Gelfand, M.S., A.A. Mironov, J. Jomantas, Y.I. Kozlov & D.A. Perumov, (1999) A conserved RNA structure element involved in the regulation of bacterial riboflavin synthesis genes. *Trends in genetics : TIG* **15**: 439-442.
- Ghrist, A.C. & G.V. Stauffer, (1995) Characterization of the *Escherichia coli* *gcvR* gene encoding a negative regulator of *gcv* expression. *Journal of bacteriology* **177**: 4980-4984.

- Gogol, E.B., V.A. Rhodius, K. Papenfort, J. Vogel & C.A. Gross, (2011) Small RNAs endow a transcriptional activator with essential repressor functions for single-tier control of a global stress regulon. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 12875-12880.
- Gold, L., D. Brown, Y. He, T. Shtatland, B.S. Singer & Y. Wu, (1997) From oligonucleotide shapes to genomic SELEX: novel biological regulatory loops. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 59-64.
- Gorke, B. & J. Vogel, (2008) Noncoding RNA control of the making and breaking of sugars. *Genes & development* **22**: 2914-2925.
- Gottesman, S., (2004) The small RNA regulators of Escherichia coli: roles and mechanisms*. *Annual review of microbiology* **58**: 303-328.
- Gottesman, S., (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends in genetics : TIG* **21**: 399-404.
- Gottesman, S. & G. Storz, (2011a) Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harbor perspectives in biology* **3**.
- Gottesman, S. & G. Storz, (2011b) Bacterial Small RNA Regulators: Versatile Roles and Rapidly Evolving Variations. *Cold Spring Harb Perspect Biol* doi: **10.1101/cshperspect.a003798**.
- Grigoriadou, C., S. Marzi, S. Kirillov, C.O. Gualerzi & B.S. Cooperman, (2007) A quantitative kinetic scheme for 70 S translation initiation complex formation. *Journal of molecular biology* **373**: 562-572.
- Groisman, E.A. & H. Ochman, (1997) How Salmonella became a pathogen. *Trends in microbiology* **5**: 343-349.
- Grose, J.H., U. Bergthorsson, Y. Xu, J. Sternecker, B. Khodaverdian & J.R. Roth, (2005) Assimilation of nicotinamide mononucleotide requires periplasmic AphA phosphatase in Salmonella enterica. *Journal of bacteriology* **187**: 4521-4530.
- Gruber, T.M. & C.A. Gross, (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual review of microbiology* **57**: 441-466.
- Grunberg-Manago, M., P. Dessen, D. Pantaloni, T. Godefroy-Colburn, A.D. Wolfe & J. Dondon, (1975) Light-scattering studies showing the effect of initiation factors on the reversible dissociation of Escherichia coli ribosomes. *Journal of molecular biology* **94**: 461-478.
- Gu, G., J. Hu, J.M. Cevallos-Cevallos, S.M. Richardson, J.A. Bartz & A.H. van Bruggen, (2011) Internal colonization of Salmonella enterica serovar Typhimurium in tomato plants. *PloS one* **6**: e27340.
- Gualerzi, C., G. Risuleo & C.L. Pon, (1977) Initial rate kinetic analysis of the mechanism of initiation complex formation and the role of initiation factor IF-3. *Biochemistry* **16**: 1684-1689.
- Gualerzi, C.O. & C.L. Pon, (1990) Initiation of mRNA translation in prokaryotes. *Biochemistry* **29**: 5881-5889.
- Guibourdenche, M., P. Roggentin, M. Mikoleit, P.I. Fields, J. Bockemuhl, P.A. Grimont & F.X. Weill, (2010) Supplement 2003-2007 (No. 47) to the White-Kauffmann-Le Minor scheme. *Research in microbiology* **161**: 26-29.
- Hajnsdorf, E. & P. Regnier, (2000) Host factor Hfq of Escherichia coli stimulates elongation of poly(A) tails by poly(A) polymerase I. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 1501-1505.

- Hale, C.R., P. Zhao, S. Olson, M.O. Duff, B.R. Graveley, L. Wells, R.M. Terns & M.P. Terns, (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**: 945-956.
- Hammer, B.K. & B.L. Bassler, (2007) Regulatory small RNAs circumvent the conventional quorum sensing pathway in pandemic *Vibrio cholerae*. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 11145-11149.
- Hantke, K., (2001) Iron and metal regulation in bacteria. *Current opinion in microbiology* **4**: 172-177.
- Hapfelmeier, S., K. Ehrbar, B. Stecher, M. Barthel, M. Kremer & W.D. Hardt, (2004) Role of the *Salmonella* pathogenicity island 1 effector proteins SipA, SopB, SopE, and SopE2 in *Salmonella enterica* subspecies 1 serovar Typhimurium colitis in streptomycin-pretreated mice. *Infection and immunity* **72**: 795-809.
- Hartl, D.L. & H. Ochman, (1996) Inverse polymerase chain reaction. *Methods Mol Biol* **58**: 293-301.
- Haurwitz, R.E., M. Jinek, B. Wiedenheft, K. Zhou & J.A. Doudna, (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**: 1355-1358.
- Hensel, M., (2000) *Salmonella* pathogenicity island 2. *Molecular microbiology* **36**: 1015-1023.
- Hensel, M., J.E. Shea, A.J. Baumler, C. Gleeson, F. Blattner & D.W. Holden, (1997) Analysis of the boundaries of *Salmonella* pathogenicity island 2 and the corresponding chromosomal region of *Escherichia coli* K-12. *Journal of bacteriology* **179**: 1105-1111.
- Hershberg, R., S. Altuvia & H. Margalit, (2003) A survey of small RNA-encoding genes in *Escherichia coli*. *Nucleic acids research* **31**: 1813-1820.
- Hook-Barnard, I.G., T.J. Brickman & M.A. McIntosh, (2007) Identification of an AU-rich translational enhancer within the *Escherichia coli* *fepB* leader RNA. *Journal of bacteriology* **189**: 4028-4037.
- Hughes, K.T. & J.R. Roth, (1988) Transitory cis complementation: a method for providing transposition functions to defective transposons. *Genetics* **119**: 9-12.
- Hui, A. & H.A. de Boer, (1987) Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 4762-4766.
- Ishihama, A., (2012) Prokaryotic genome regulation: a revolutionary paradigm. *Proceedings of the Japan Academy. Series B, Physical and biological sciences* **88**: 485-508.
- Ishino, Y., H. Shinagawa, K. Makino, M. Amemura & A. Nakata, (1987) Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of bacteriology* **169**: 5429-5433.
- Jacob, W.F., M. Santer & A.E. Dahlberg, (1987) A single base change in the Shine-Dalgarno region of 16S rRNA of *Escherichia coli* affects translation of many proteins. *Proceedings of the National Academy of Sciences of the United States of America* **84**: 4757-4761.
- Jacques, J.F., S. Jang, K. Prevost, G. Desnoyers, M. Desmarais, J. Imlay & E. Masse, (2006) RyhB small RNA modulates the free intracellular iron pool and is essential for normal growth during iron limitation in *Escherichia coli*. *Mol Microbiol* **62**: 1181-1190.

- Jakomin, M., D. Chessa, A.J. Baumler & J. Casadesus, (2008) Regulation of the *Salmonella enterica* std fimbrial operon by DNA adenine methylation, SeqA, and HdfR. *Journal of bacteriology* **190**: 7406-7413.
- Johansen, J., A.A. Rasmussen, M. Overgaard & P. Valentin-Hansen, (2006) Conserved small non-coding RNAs that belong to the sigmaE regulon: role in down-regulation of outer membrane proteins. *Journal of molecular biology* **364**: 1-8.
- Jourdan, A.D. & G.V. Stauffer, (1999) Genetic analysis of the GcvA binding site in the *gcvA* control region. *Microbiology* **145 (Pt 8)**: 2153-2162.
- Jousselin, A., L. Metzinger & B. Felden, (2009) On the facultative requirement of the bacterial RNA chaperone, Hfq. *Trends in microbiology* **17**: 399-405.
- Karlinsey, J.E., (2007) lambda-Red genetic engineering in *Salmonella enterica* serovar Typhimurium. *Methods in enzymology* **421**: 199-209.
- Kawamoto, H., Y. Koide, T. Morita & H. Aiba, (2006) Base-pairing requirement for RNA silencing by a bacterial small RNA and acceleration of duplex formation by Hfq. *Molecular microbiology* **61**: 1013-1022.
- Kawano, M., A.A. Reynolds, J. Miranda-Rios & G. Storz, (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic acids research* **33**: 1040-1050.
- Kawashima, T., H. Aramaki, T. Oyamada, K. Makino, M. Yamada, H. Okamura, K. Yokoyama, S.A. Ishijima & M. Suzuki, (2008) Transcription regulation by feast/famine regulatory proteins, FFRPs, in archaea and eubacteria. *Biol Pharm Bull* **31**: 173-186.
- Kleckner, N., J. Bender & S. Gottesman, (1991) Uses of transposons with emphasis on Tn10. *Methods in enzymology* **204**: 139-180.
- Ko, M. & C. Park, (2000) H-NS-Dependent regulation of flagellar synthesis is mediated by a LysR family protein. *Journal of bacteriology* **182**: 4670-4672.
- Komarova, A.V., L.S. Tchufistova, L.V. Aseev & I.V. Boni, (2005a) [An *Escherichia coli* strain producing a leaderless mRNA from the chromosomal lac promoter]. *Bioorganicheskaiia khimiia* **31**: 557-560.
- Komarova, A.V., L.S. Tchufistova, M. Dreyfus & I.V. Boni, (2005b) AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli*. *Journal of bacteriology* **187**: 1344-1349.
- Komarova, A.V., L.S. Tchufistova, E.V. Supina & I.V. Boni, (2002) Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. *RNA* **8**: 1137-1147.
- Kunin, V., R. Sorek & P. Hugenholtz, (2007) Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome biology* **8**: R61.
- Kunst, F., N. Ogasawara, I. Moszer, A.M. Albertini, G. Alloni, V. Azevedo, M.G. Bertero, P. Bessieres, A. Bolotin, S. Borchert, R. Borriss, L. Boursier, A. Brans, M. Braun, S.C. Brignell, S. Bron, S. Brouillet, C.V. Bruschi, B. Caldwell, V. Capuano, N.M. Carter, S.K. Choi, J.J. Codani, I.F. Connerton, A. Danchin & et al., (1997) The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* **390**: 249-256.
- Kyrpides, N.C. & C.R. Woese, (1998) Universally conserved translation initiation factors. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 224-228.
- La Teana, A., C.O. Gualerzi & R. Brimacombe, (1995) From stand-by to decoding site. Adjustment of the mRNA on the 30S ribosomal subunit under the influence of the initiation factors. *RNA* **1**: 772-782.

- Laursen, B.S., H.P. Sorensen, K.K. Mortensen & H.U. Sperling-Petersen, (2005) Initiation of protein synthesis in bacteria. *Microbiology and molecular biology reviews* : *MMBR* **69**: 101-123.
- Le Hello, S., A. Brisabois, M. Accou-Demartin, A. Josse, M. Marault, S. Francart, N.J. Da Silva & F.X. Weill, (2012) Foodborne outbreak and nonmotile *Salmonella enterica* variant, France. *Emerging infectious diseases* **18**: 132-134.
- Le Minor, L. & J. Bockemuhl, (1984) [Supplement No. XXVII (1983) to the Kauffmann-White scheme]. *Annales de microbiologie* **135B**: 45-51.
- Le Minor, L., M.Y. Popoff, B. Laurent & D. Hermant, (1986) [Characterization of a 7th subspecies of *Salmonella*: *S. choleraesuis* subsp. *indica* subsp. nov.]. *Annales de l'Institut Pasteur. Microbiology* **137B**: 211-217.
- Le Minor, L., M. Veron & M. Popoff, (1982) [The taxonomy of *Salmonella*]. *Annales de microbiologie* **133**: 223-243.
- Lee, D.J., S.D. Minchin & S.J. Busby, (2012) Activating transcription in bacteria. *Annual review of microbiology* **66**: 125-152.
- Lemire, S., N. Figueroa-Bossi & L. Bossi, (2011) Bacteriophage crosstalk: coordination of prophage induction by trans-acting antirepressors. *PLoS genetics* **7**: e1002149.
- Lenz, D.H., K.C. Mok, B.N. Lilley, R.V. Kulkarni, N.S. Wingreen & B.L. Bassler, (2004) The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. *Cell* **118**: 69-82.
- Lilleengen, K., (1948) Typing of *Salmonella typhimurium* by means of bacteriophage. *Acta Pathol. Microbiol. Scand.* **77**: 2-125.
- Lin, R., R. D'Ari & E.B. Newman, (1992) Lambda placMu insertions in genes of the leucine regulon: extension of the regulon to genes not regulated by leucine. *Journal of bacteriology* **174**: 1948-1955.
- Link, T.M., P. Valentin-Hansen & R.G. Brennan, (2009) Structure of *Escherichia coli* Hfq bound to polyribadenylate RNA. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 19292-19297.
- Lobocka, M.B., D.J. Rose, G. Plunkett, 3rd, M. Rusin, A. Samojedny, H. Lehnerr, M.B. Yarmolinsky & F.R. Blattner, (2004) Genome of bacteriophage P1. *Journal of bacteriology* **186**: 7032-7068.
- Lybecker, M.C., C.A. Abel, A.L. Feig & D.S. Samuels, (2010) Identification and function of the RNA chaperone Hfq in the Lyme disease spirochete *Borrelia burgdorferi*. *Molecular microbiology* **78**: 622-635.
- Macconkey, A., (1905) Lactose-Fermenting Bacteria in Faeces. *The Journal of hygiene* **5**: 333-379.
- Majdalani, N., C.K. Vanderpool & S. Gottesman, (2005) Bacterial small RNA regulators. *Critical reviews in biochemistry and molecular biology* **40**: 93-113.
- Majowicz, S.E., J. Musto, E. Scallan, F.J. Angulo, M. Kirk, S.J. O'Brien, T.F. Jones, A. Fazil & R.M. Hoekstra, (2010) The global burden of nontyphoidal *Salmonella* gastroenteritis. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **50**: 882-889.
- Maloy, S.R. & K.T. Hughes, (2007) Strain collections and genetic nomenclature. *Methods in enzymology* **421**: 3-8.
- Maloy, S.R. & J.R. Roth, (1983) Regulation of proline utilization in *Salmonella typhimurium*: characterization of put::Mu d(Ap, lac) operon fusions. *Journal of bacteriology* **154**: 561-568.

- Marshall, D.G., B.J. Sheehan & C.J. Dorman, (1999) A role for the leucine-responsive regulatory protein and integration host factor in the regulation of the Salmonella plasmid virulence (spv) locus in Salmonella typhimurium. *Molecular microbiology* **34**: 134-145.
- Martin-Farmer, J. & G.R. Janssen, (1999) A downstream CA repeat sequence increases translation from leadered and unleadered mRNA in Escherichia coli. *Molecular microbiology* **31**: 1025-1038.
- Masse, E., F.E. Escorcía & S. Gottesman, (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes & development* **17**: 2374-2383.
- Masse, E. & S. Gottesman, (2002) A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proc Natl Acad Sci U S A* **99**: 4620-4625.
- Masse, E., H. Salvail, G. Desnoyers & M. Arguin, (2007) Small RNAs controlling iron metabolism. *Current opinion in microbiology* **10**: 140-145.
- Masse, E., C.K. Vanderpool & S. Gottesman, (2005) Effect of RyhB small RNA on global iron use in Escherichia coli. *Journal of bacteriology* **187**: 6962-6971.
- McCarthy, J.E., H.U. Schairer & W. Sebald, (1985) Translational initiation frequency of atp genes from Escherichia coli: identification of an intercistronic sequence that enhances translation. *Embo J* **4**: 519-526.
- McClelland, M., K.E. Sanderson, J. Spieth, S.W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F. Du, S. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Nhan, R. Waterston & R.K. Wilson, (2001) Complete genome sequence of Salmonella enterica serovar Typhimurium LT2. *Nature* **413**: 852-856.
- Meibom, K.L., A.L. Forslund, K. Kuoppa, K. Alkhuder, I. Dubail, M. Dupuis, A. Forsberg & A. Charbit, (2009) Hfq, a novel pleiotropic regulator of virulence-associated genes in Francisella tularensis. *Infection and immunity* **77**: 1866-1880.
- Mikulecky, P.J., M.K. Kaw, C.C. Brescia, J.C. Takach, D.D. Sledjeski & A.L. Feig, (2004) Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. *Nature structural & molecular biology* **11**: 1206-1214.
- Miller, J.H., (1992) *A Short Course in Bacterial Genetics. A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor New York.
- Milon, P., C. Maracci, L. Filonava, C.O. Gualerzi & M.V. Rodnina, (2012) Real-time assembly landscape of bacterial 30S translation initiation complex. *Nature structural & molecular biology* **19**: 609-615.
- Miranda, G., D. Schuppli, I. Barrera, C. Hausherr, J.M. Sogo & H. Weber, (1997) Recognition of bacteriophage Qbeta plus strand RNA as a template by Qbeta replicase: role of RNA interactions mediated by ribosomal proteins S1 and host factor. *Journal of molecular biology* **267**: 1089-1103.
- Miranda-Rios, J., M. Navarro & M. Soberon, (2001) A conserved RNA structure (thi box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 9736-9741.
- Mironov, A.S., I. Gusarov, R. Rafikov, L.E. Lopez, K. Shatalin, R.A. Kreneva, D.A. Perumov & E. Nudler, (2002) Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111**: 747-756.

- Mizuno, T., M.Y. Chou & M. Inouye, (1984) A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proceedings of the National Academy of Sciences of the United States of America* **81**: 1966-1970.
- Moazed, D., R.R. Samaha, C. Gualerzi & H.F. Noller, (1995) Specific protection of 16 S rRNA by translational initiation factors. *Journal of molecular biology* **248**: 207-210.
- Modi, S.R., D.M. Camacho, M.A. Kohanski, G.C. Walker & J.J. Collins, Functional characterization of bacterial sRNAs using a network biology approach. *Proc Natl Acad Sci U S A* **108**: 15522-15527.
- Modi, S.R., D.M. Camacho, M.A. Kohanski, G.C. Walker & J.J. Collins, (2011) Functional characterization of bacterial sRNAs using a network biology approach. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 15522-15527.
- Mohanty, B.K., V.F. Maples & S.R. Kushner, (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Molecular microbiology* **54**: 905-920.
- Moll, I., S. Grill, A. Grundling & U. Blasi, (2002) Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in Escherichia coli. *Molecular microbiology* **44**: 1387-1396.
- Moller, T., T. Franch, P. Hojrup, D.R. Keene, H.P. Bachinger, R.G. Brennan & P. Valentin-Hansen, (2002) Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. *Molecular cell* **9**: 23-30.
- Morita, T., K. Maki & H. Aiba, (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes & development* **19**: 2176-2186.
- Muffler, A., D.D. Traulsen, D. Fischer, R. Lange & R. Hengge-Aronis, (1997) The RNA-binding protein HF-I plays a global regulatory role which is largely, but not exclusively, due to its role in expression of the sigmaS subunit of RNA polymerase in Escherichia coli. *Journal of bacteriology* **179**: 297-300.
- Murphy, K.C., K.G. Campellone & A.R. Poteete, (2000) PCR-mediated gene replacement in Escherichia coli. *Gene* **246**: 321-330.
- Myasnikov, A.G., A. Simonetti, S. Marzi & B.P. Klaholz, (2009) Structure-function insights into prokaryotic and eukaryotic translation initiation. *Current opinion in structural biology* **19**: 300-309.
- Nahvi, A., N. Sudarsan, M.S. Ebert, X. Zou, K.L. Brown & R.R. Breaker, (2002) Genetic control by a metabolite binding mRNA. *Chemistry & biology* **9**: 1043.
- Nakagawa, S., Y. Niimura, K. Miura & T. Gojobori, (2010) Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 6382-6387.
- Nakata, A., M. Amemura & K. Makino, (1989) Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 chromosome. *Journal of bacteriology* **171**: 3553-3556.
- Newman, E.B. & R. Lin, (1995) Leucine-responsive regulatory protein: a global regulator of gene expression in E. coli. *Annual review of microbiology* **49**: 747-775.
- O'Connor, M. & A.E. Dahlberg, (2001) Enhancement of translation by the epsilon element is independent of the sequence of the 460 region of 16S rRNA. *Nucleic acids research* **29**: 1420-1425.

- Ochman, H. & A.C. Wilson, (1987) Evolution in bacteria: evidence for a universal substitution rate in cellular genomes. *Journal of molecular evolution* **26**: 74-86.
- Opdyke, J.A., J.G. Kang & G. Storz, (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *Journal of bacteriology* **186**: 6698-6705.
- Otaka, H., H. Ishikawa, T. Morita & H. Aiba, (2011) PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 13059-13064.
- Papenfort, K., V. Pfeiffer, F. Mika, S. Lucchini, J.C. Hinton & J. Vogel, (2006) SigmaE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global omp mRNA decay. *Molecular microbiology* **62**: 1674-1688.
- Papenfort, K., Y. Sun, M. Miyakoshi, C.K. Vanderpool & J. Vogel, (2013) Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. *Cell* **153**: 426-437.
- Patrick A.D. Grimont, F.-X.W., (2007) ANTIGENIC FORMULAE OF THE SALMONELLA SEROVARS. *WHO Collaborating Centre for Reference and Research on Salmonella*.
- Peters, J.M., A.D. Vangeloff & R. Landick, (2011) Bacterial transcription terminators: the RNA 3'-end chronicles. *Journal of molecular biology* **412**: 793-813.
- Pfeiffer, V., K. Papenfort, S. Lucchini, J.C. Hinton & J. Vogel, (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nature structural & molecular biology* **16**: 840-846.
- Pon, C.L. & C.O. Gualerzi, (1984) Mechanism of protein biosynthesis in prokaryotic cells. Effect of initiation factor IF1 on the initial rate of 30 S initiation complex formation. *FEBS letters* **175**: 203-207.
- Pon, C.L., B. Wittmann-Liebold & C. Gualerzi, (1979) Structure--function relationships in *Escherichia coli* initiation factors. II. Elucidation of the primary structure of initiation factor IF-1. *FEBS letters* **101**: 157-160.
- Pourcel, C., G. Salvignol & G. Vergnaud, (2005) CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**: 653-663.
- Prevost, K., G. Desnoyers, J.F. Jacques, F. Lavoie & E. Masse, (2011) Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes & development* **25**: 385-396.
- Prevost, K., H. Salvail, G. Desnoyers, J.F. Jacques, E. Phaneuf & E. Masse, (2007) The small RNA RyhB activates the translation of shiA mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis. *Molecular microbiology* **64**: 1260-1273.
- Pulvermacher, S.C., L.T. Stauffer & G.V. Stauffer, (2008) The role of the small regulatory RNA GcvB in GcvB/mRNA posttranscriptional regulation of oppA and dppA in *Escherichia coli*. *FEMS microbiology letters* **281**: 42-50.
- Pulvermacher, S.C., L.T. Stauffer & G.V. Stauffer, (2009a) Role of the *Escherichia coli* Hfq protein in GcvB regulation of oppA and dppA mRNAs. *Microbiology* **155**: 115-123.
- Pulvermacher, S.C., L.T. Stauffer & G.V. Stauffer, (2009b) Role of the sRNA GcvB in regulation of cycA in *Escherichia coli*. *Microbiology* **155**: 106-114.

- Pulvermacher, S.C., L.T. Stauffer & G.V. Stauffer, (2009c) The small RNA GcvB regulates *ssfT* mRNA expression in *Escherichia coli*. *Journal of bacteriology* **191**: 238-248.
- Ramirez-Pena, E., J. Trevino, Z. Liu, N. Perez & P. Sumby, (2010) The group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the *ska* mRNA transcript. *Molecular microbiology* **78**: 1332-1347.
- Rappleye, C.A. & J.R. Roth, (1997) A Tn10 derivative (T-POP) for isolation of insertions with conditional (tetracycline-dependent) phenotypes. *Journal of bacteriology* **179**: 5827-5834.
- Rasmussen, A.A., M. Eriksen, K. Gilany, C. Udesen, T. Franch, C. Petersen & P. Valentin-Hansen, (2005) Regulation of *ompA* mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Molecular microbiology* **58**: 1421-1429.
- Reeves, M.W., G.M. Ewins, A.A. Heiba, B.D. Plikaytis & J.J. Farmer, 3rd, (1989) Clonal nature of *Salmonella typhi* and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis, and proposal of *Salmonella bongori* comb. nov. *Journal of clinical microbiology* **27**: 313-320.
- Repoila, F., N. Majdalani & S. Gottesman, (2003) Small non-coding RNAs, coordinators of adaptation processes in *Escherichia coli*: the RpoS paradigm. *Molecular microbiology* **48**: 855-861.
- Ribes, V., K. Romisch, A. Giner, B. Dobberstein & D. Tollervy, (1990) *E. coli* 4.5S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. *Cell* **63**: 591-600.
- Richardson, J.P., (2003) Loading Rho to terminate transcription. *Cell* **114**: 157-159.
- Ringquist, S., T. Jones, E.E. Snyder, T. Gibson, I. Boni & L. Gold, (1995) High-affinity RNA ligands to *Escherichia coli* ribosomes and ribosomal protein S1: comparison of natural and unnatural binding sites. *Biochemistry* **34**: 3640-3648.
- Rochat, T., P. Bouloc, Q. Yang, L. Bossi & N. Figueroa-Bossi, (2012) Lack of interchangeability of Hfq-like proteins. *Biochimie* **94**: 1554-1559.
- Romeo, T., (1998) Global regulation by the small RNA-binding protein CsrA and the non-coding RNA molecule CsrB. *Molecular microbiology* **29**: 1321-1330.
- Sacerdot, C., C. Chiaruttini, K. Engst, M. Graffe, M. Milet, N. Mathy, J. Dondon & M. Springer, (1996) The role of the AUU initiation codon in the negative feedback regulation of the gene for translation initiation factor IF3 in *Escherichia coli*. *Molecular microbiology* **21**: 331-346.
- Sacerdot, C., G. Fayat, P. Dessen, M. Springer, J.A. Plumbridge, M. Grunberg-Manago & S. Blanquet, (1982) Sequence of a 1.26-kb DNA fragment containing the structural gene for *E. coli* initiation factor IF3: presence of an AUU initiator codon. *The EMBO journal* **1**: 311-315.
- Salgado-Garrido, J., E. Bragado-Nilsson, S. Kandels-Lewis & B. Seraphin, (1999) Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *The EMBO journal* **18**: 3451-3462.
- Sanderson, K.E., A. Hessel & K.E. Rudd, (1995) Genetic map of *Salmonella typhimurium*, edition VIII. *Microbiological reviews* **59**: 241-303.
- Sashital, D.G., M. Jinek & J.A. Doudna, (2011) An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nature structural & molecular biology* **18**: 680-687.

- Sauer, E. & O. Weichenrieder, (2011) Structural basis for RNA 3'-end recognition by Hfq. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 13065-13070.
- Sauter, C., J. Basquin & D. Suck, (2003) Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli*. *Nucleic acids research* **31**: 4091-4098.
- Scallan, E., R.M. Hoekstra, F.J. Angulo, R.V. Tauxe, M.A. Widdowson, S.L. Roy, J.L. Jones & P.M. Griffin, (2011) Foodborne illness acquired in the United States--major pathogens. *Emerging infectious diseases* **17**: 7-15.
- Schmieger, H., (1972) Phage P22-mutants with increased or decreased transduction abilities. *Molecular & general genetics : MGG* **119**: 75-88.
- Schumacher, M.A., R.F. Pearson, T. Moller, P. Valentin-Hansen & R.G. Brennan, (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *The EMBO journal* **21**: 3546-3556.
- Sengupta, J., R.K. Agrawal & J. Frank, (2001) Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 11991-11996.
- Serganov, A. & E. Nudler, (2013) A decade of riboswitches. *Cell* **152**: 17-24.
- Sezonov, G., D. Joseleau-Petit & R. D'Ari, (2007) *Escherichia coli* physiology in Luria-Bertani broth. *Journal of bacteriology* **189**: 8746-8749.
- Sharma, C.M., F. Darfeuille, T.H. Plantinga & J. Vogel, (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev* **21**: 2804-2817.
- Sharma, C.M., K. Papenfort, S.R. Pernitzsch, H.J. Mollenkopf, J.C. Hinton & J. Vogel, (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* **81**: 1144-1165.
- Shea, J.E., C.R. Beuzon, C. Gleeson, R. Mundy & D.W. Holden, (1999) Influence of the *Salmonella typhimurium* pathogenicity island 2 type III secretion system on bacterial growth in the mouse. *Infection and immunity* **67**: 213-219.
- Shine, J. & L. Dalgarno, (1974) The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* **71**: 1342-1346.
- Shultzaberger, R.K., R.E. Bucheimer, K.E. Rudd & T.D. Schneider, (2001) Anatomy of *Escherichia coli* ribosome binding sites. *J Mol Biol* **313**: 215-228.
- Simonetti, A., S. Marzi, L. Jenner, A. Myasnikov, P. Romby, G. Yusupova, B.P. Klaholz & M. Yusupov, (2009) A structural view of translation initiation in bacteria. *Cellular and molecular life sciences : CMLS* **66**: 423-436.
- Sittka, A., V. Pfeiffer, K. Tedin & J. Vogel, (2007) The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium*. *Molecular microbiology* **63**: 193-217.
- Soper, T., P. Mandin, N. Majdalani, S. Gottesman & S.A. Woodson, (2010) Positive regulation by small RNAs and the role of Hfq. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 9602-9607.
- Soper, T.J. & S.A. Woodson, (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* **14**: 1907-1917.
- Sorensen, H.P., J. Hedegaard, H.U. Sperling-Petersen & K.K. Mortensen, (2001) Remarkable conservation of translation initiation factors: IF1/eIF1A and

- IF2/eIF5B are universally distributed phylogenetic markers. *IUBMB life* **51**: 321-327.
- Sorensen, M.A., J. Fricke & S. Pedersen, (1998) Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in *Escherichia coli* in vivo. *Journal of molecular biology* **280**: 561-569.
- Stauffer, G.V., (1987) Biosynthesis of Serine and Glycine. In: *Escherichia coli* and *Salmonella*: Cellular and Molecular Biology. F.C. Neidhart, J.L. Ingraham, K.B. Low, B. Magasanik, M. Schaechter & H.E. Umbarger (eds). Washington, DC: American Society for Microbiology Press, pp. 412-418.
- Stauffer, L.T. & G.V. Stauffer, (1998) Roles for GcvA-binding sites 3 and 2 and the Lrp-binding region in *gcvT::lacZ* expression in *Escherichia coli*. *Microbiology* **144 (Pt 10)**: 2865-2872.
- Stauffer, L.T. & G.V. Stauffer, (1999) Role for the leucine-responsive regulatory protein (Lrp) as a structural protein in regulating the *Escherichia coli* *gcvTHP* operon. *Microbiology* **145 (Pt 3)**: 569-576.
- Stauffer, L.T. & G.V. Stauffer, (2005) GcvA interacts with both the alpha and sigma subunits of RNA polymerase to activate the *Escherichia coli* *gcvB* gene and the *gcvTHP* operon. *FEMS microbiology letters* **242**: 333-338.
- Steitz, J.A. & K. Jakes, (1975) How ribosomes select initiator regions in mRNA: base pair formation between the 3' terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **72**: 4734-4738.
- Storz, G., J. Vogel & K.M. Wassarman, (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Molecular cell* **43**: 880-891.
- Su, Q., D. Schuppli, T. Tsui Hc, M.E. Winkler & H. Weber, (1997) Strongly reduced phage Qbeta replication, but normal phage MS2 replication in an *Escherichia coli* K12 mutant with inactivated Qbeta host factor (*hfq*) gene. *Virology* **227**: 211-214.
- Sun, X. & R.M. Wartell, (2006) *Escherichia coli* Hfq binds A18 and DsrA domain II with similar 2:1 Hfq6/RNA stoichiometry using different surface sites. *Biochemistry* **45**: 4875-4887.
- Sun, X., I. Zhulin & R.M. Wartell, (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nucleic acids research* **30**: 3662-3671.
- Sussman, J.K., E.L. Simons & R.W. Simons, (1996) *Escherichia coli* translation initiation factor 3 discriminates the initiation codon in vivo. *Molecular microbiology* **21**: 347-360.
- Tani, T.H., A. Khodursky, R.M. Blumenthal, P.O. Brown & R.G. Matthews, (2002) Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 13471-13476.
- Tedin, K., A. Resch & U. Blasi, (1997) Requirements for ribosomal protein S1 for translation initiation of mRNAs with and without a 5' leader sequence. *Molecular microbiology* **25**: 189-199.
- Tharun, S., W. He, A.E. Mayes, P. Lennertz, J.D. Beggs & R. Parker, (2000) Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* **404**: 515-518.
- Thompson, K.M., V.A. Rhodius & S. Gottesman, (2007) SigmaE regulates and is regulated by a small RNA in *Escherichia coli*. *Journal of bacteriology* **189**: 4243-4256.

- Tindall, B.J., P.A. Grimont, G.M. Garrity & J.P. Euzeby, (2005) Nomenclature and taxonomy of the genus *Salmonella*. *International journal of systematic and evolutionary microbiology* **55**: 521-524.
- Travers, A. & G. Muskhelishvili, (2005) DNA supercoiling - a global transcriptional regulator for enterobacterial growth? *Nature reviews. Microbiology* **3**: 157-169.
- Tzareva, N.V., V.I. Makhno & I.V. Boni, (1994) Ribosome-messenger recognition in the absence of the Shine-Dalgarno interactions. *FEBS letters* **337**: 189-194.
- Urban, J.H. & J. Vogel, (2007) Translational control and target recognition by *Escherichia coli* small RNAs in vivo. *Nucleic acids research* **35**: 1018-1037.
- Urban, J.H. & J. Vogel, (2008) Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS biology* **6**: e64.
- Urbanowski, M.L., L.T. Stauffer & G.V. Stauffer, (2000) The *gcvB* gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*. *Mol Microbiol* **37**: 856-868.
- Uzzau, S., N. Figueroa-Bossi, S. Rubino & L. Bossi, (2001) Epitope tagging of chromosomal genes in *Salmonella*. *Proc Natl Acad Sci U S A* **98**: 15264-15269.
- Valentin-Hansen, P., M. Eriksen & C. Udesen, (2004) The bacterial Sm-like protein Hfq: a key player in RNA transactions. *Molecular microbiology* **51**: 1525-1533.
- van der Woude, M.W., B.A. Braaten & D.A. Low, (1992) Evidence for global regulatory control of pilus expression in *Escherichia coli* by Lrp and DNA methylation: model building based on analysis of *pap*. *Molecular microbiology* **6**: 2429-2435.
- Vanderpool, C.K., (2007) Physiological consequences of small RNA-mediated regulation of glucose-phosphate stress. *Current opinion in microbiology* **10**: 146-151.
- Vasil'eva Iu, M. & M.B. Garber, (2002) [The regulatory role of the Hfq protein in bacterial cells]. *Molekuliarnaia biologii* **36**: 970-977.
- Vassilieva, I.M., M.V. Rouzanov, N.V. Zelinskaya, I. Moll, U. Blasi & M.B. Garber, (2002) Cloning, purification, and crystallization of a bacterial gene expression regulator--Hfq protein from *Escherichia coli*. *Biochemistry. Biokhimiia* **67**: 1293-1297.
- Vecerek, B., I. Moll & U. Blasi, (2007) Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *The EMBO journal* **26**: 965-975.
- Vesper, O., S. Amitai, M. Belitsky, K. Byrgazov, A.C. Kaberdina, H. Engelberg-Kulka & I. Moll, (2011) Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in *Escherichia coli*. *Cell* **147**: 147-157.
- Vogel, J., (2009) A rough guide to the non-coding RNA world of *Salmonella*. *Molecular microbiology* **71**: 1-11.
- Vogel, J. & B.F. Luisi, (2011) Hfq and its constellation of RNA. *Nature reviews. Microbiology* **9**: 578-589.
- Vogel, J. & E.G. Wagner, (2007) Target identification of small noncoding RNAs in bacteria. *Current opinion in microbiology* **10**: 262-270.
- Wagner, E.G., (2009) Kill the messenger: bacterial antisense RNA promotes mRNA decay. *Nature structural & molecular biology* **16**: 804-806.
- Wagner, E.G., (2013) Cycling of RNAs on Hfq. *RNA biology* **10**: 619-626.

- Wang, Q. & J.M. Calvo, (1993) Lrp, a major regulatory protein in *Escherichia coli*, bends DNA and can organize the assembly of a higher-order nucleoprotein structure. *The EMBO journal* **12**: 2495-2501.
- Waters, L.S. & G. Storz, (2009) Regulatory RNAs in bacteria. *Cell* **136**: 615-628.
- Wiedenheft, B., G.C. Lander, K. Zhou, M.M. Jore, S.J. Brouns, J. van der Oost, J.A. Doudna & E. Nogales, (2011a) Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* **477**: 486-489.
- Wiedenheft, B., S.H. Sternberg & J.A. Doudna, (2012) RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**: 331-338.
- Wiedenheft, B., E. van Duijn, J.B. Bultema, S.P. Waghmare, K. Zhou, A. Barendregt, W. Westphal, A.J. Heck, E.J. Boekema, M.J. Dickman & J.A. Doudna, (2011b) RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proceedings of the National Academy of Sciences of the United States of America* **108**: 10092-10097.
- Willins, D.A., C.W. Ryan, J.V. Platko & J.M. Calvo, (1991) Characterization of Lrp, and *Escherichia coli* regulatory protein that mediates a global response to leucine. *The Journal of biological chemistry* **266**: 10768-10774.
- Wilson, R.L., L.T. Stauffer & G.V. Stauffer, (1993) Roles of the GcvA and PurR proteins in negative regulation of the *Escherichia coli* glycine cleavage enzyme system. *Journal of bacteriology* **175**: 5129-5134.
- Wilusz, C.J. & J. Wilusz, (2013) Lsm proteins and Hfq: Life at the 3' end. *RNA biology* **10**: 592-601.
- Winkler, W., A. Nahvi & R.R. Breaker, (2002a) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**: 952-956.
- Winkler, W.C., S. Cohen-Chalamish & R.R. Breaker, (2002b) An mRNA structure that controls gene expression by binding FMN. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 15908-15913.
- Wintermeyer, W. & C. Gualerzi, (1983) Effect of *Escherichia coli* initiation factors on the kinetics of N-AcPhe-tRNA^{Phe} binding to 30S ribosomal subunits. A fluorescence stopped-flow study. *Biochemistry* **22**: 690-694.
- Yarchuk, O., N. Jacques, J. Guillerez & M. Dreyfus, (1992) Interdependence of translation, transcription and mRNA degradation in the *lacZ* gene. *J Mol Biol* **226**: 581-596.
- Yu, D., H.M. Ellis, E.C. Lee, N.A. Jenkins, N.G. Copeland & D.L. Court, (2000) An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc Natl Acad Sci U S A* **97**: 5978-5983.
- Yusupova, G.Z., M.M. Yusupov, J.H. Cate & H.F. Noller, (2001) The path of messenger RNA through the ribosome. *Cell* **106**: 233-241.
- Zhang, A., K.M. Wassarman, J. Ortega, A.C. Steven & G. Storz, (2002) The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Molecular cell* **9**: 11-22.
- Zhang, A., K.M. Wassarman, C. Rosenow, B.C. Tjaden, G. Storz & S. Gottesman, (2003) Global analysis of small RNA and mRNA targets of Hfq. *Molecular microbiology* **50**: 1111-1124.
- Zhang, J. & M.P. Deutscher, (1992) A uridine-rich sequence required for translation of prokaryotic mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **89**: 2605-2609.