UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 916

Macromolecular Matchmaking

Mechanisms and Biology of Bacterial Small RNAs

ERIK HOLMQVIST

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2012

ISSN 1651-6214
ISBN 978-91-554-8331-9
urn:nbn:se:uu:diva-171642

Dissertation presented at Uppsala University to be publicly examined in Sal B42, Biomedicinskt Centrum, Husargatan 3, Uppsala, Wednesday, May 16, 2012 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Holmqvist, E. 2012. Macromolecular Matchmaking: Mechanisms and Biology of Bacterial Small RNAs. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 916. 66 pp. Uppsala. ISBN 978-91-554-8331-9.

Cells sense the properties of the surrounding environment and convert this information into changes in gene expression. Bacteria are, in contrast to many multi-cellular eukaryotes, remarkable in their ability to cope with rapid environmental changes and to endure harsh and extreme milieus. Previously, control of gene expression was thought to be carried out exclusively by proteins. However, it is now clear that small regulatory RNAs (sRNA) also carry out gene regulatory functions. Bacteria such as *E. coli* harbor a large class of sRNAs that bind to mRNAs to alter translation and/or mRNA stability.

By identifying mRNAs that are targeted by sRNAs, my studies have broadened the understanding of the mechanisms that underlie sRNA-dependent gene regulation, and have shed light on the impact that this type of regulation has on bacterial physiology. Control of gene expression often relies on the interplay of many regulators. This interplay is exemplified by our discovery of mutual regulation between the sRNA MicF and the globally acting transcription factor Lrp. Through double negative feedback, these two regulators respond to nutrient availability in the environment which results in reprogramming of downstream gene expression. We have also shown that both the transcription factor CsgD, and the anti-sigma factor FlgM, are repressed by the two sRNAs OmrA and OmrB, suggesting that these sRNAs are important players in the complex regulation that allow bacteria to switch between motility and sessility. Bacterial populations of genetically identical individuals show phenotypic variations when switching to the sessile state due to bistability in gene expression. While bistability has previously been demonstrated to arise from stochastic fluctuations in transcription, our results suggest that bistability possibly may arise from sRNA-dependent regulatory events also on the post-transcriptional level.

Keywords: Small RNA, sRNA, non-coding RNA, antisense, gene regulation, post-transcriptional regulation, translational control, outer membrane protein, OmpA, MicA, biofilm, flagella, curli, bistability, Lrp, MicF, CsgD, FlgM, OmrA, OmrB

Erik Holmqvist, Uppsala University, Department of Cell and Molecular Biology, Microbiology, Box 596, SE-751 24 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-554-8331-9
urn:nbn:se:uu:diva-171642 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-171642)

Till min familj

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

I Udekwu KI, Darfeuille F, Vogel J, Reimegård J, **Holmqvist E**, Wagner EGH (2005) Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. *Genes Dev*, 19:2355-66

II **Holmqvist E**, Reimegård J, Sterk M, Grantcharova N, Römling U, Wagner EGH (2010) Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO J*, 29:1840-50

III **Holmqvist E**, Unoson C, Reimegård J, Wagner EGH (2012) A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. *Mol Microbiol*, doi:10.1111/j.1365-2958.2012.07994.x.

IV **Holmqvist E**, Sterk M, Wagner EGH (2012) Regulation of bacterial motility and bistable gene expression by base-pairing sRNAs. *Manuscript*

Published papers not included in the thesis:

Baranowska I, Jäderlund KH, Nennesmo I, **Holmqvist E**, Heidrich N, Larsson NG, Andersson G, Wagner EGH, Hedhammar A, Wibom R, Andersson L (2009) Sensory ataxic neuropathy in golden retriever dogs is caused by a deletion in the mitochondrial tRNATyr gene. *PLoS Genetics* 5(5):e1000499

Contents

Abbreviations

DNA	deoxyribonucleic acid
RNA	ribonucleic acid
mRNA	messenger RNA
ORF	open reading frame
TF	transcription factor
miRNA	microRNA
sRNA	small regulatory RNA
RBS	ribosome binding site
FACS	fluorescence-activated cell sorting
PCR	polymerase chain reaction
GTP	guanosine triphosphate
c-di-GMP	cyclic di-3',5'-guanylate
csg	curli specific gene
Omp	outer membrane protein
Mic	mRNA-interfering-complementary

Introduction

Preface

All living cells sense the intracellular and extracellular environments and convert this information into changes in gene expression. Subsequently, this leads to changes in cell physiology, metabolism, behavior and development. In contrast to most multicellular organisms, bacteria have evolved a tremendous ability to cope with harsh environmental conditions such as extremely high or low temperature, high salt concentrations, high or low pH, just to mention a few examples. Many bacteria are symbionts that carry out functions beneficial for their hosts, and the human microbiome consists of many bacterial species important for human health. However, some bacterial species are pathogens that threaten the lives of humans, animals and plants. An understanding of how bacteria control gene expression is therefore of high relevance for healthcare.

Flow of genetic information

In almost all organisms, genetic information is stored in double-stranded deoxyribonucleic acid (DNA) (Avery *et al.*, 1944; Hershey and Chase, 1952). During transcription, defined regions of DNA called transcription units act as templates for the generation of single-stranded ribonucleic acid (RNA). Messenger RNAs (mRNAs) contain open reading frames (ORFs) which are templates for protein synthesis carried out by the multimolecular complex known as the ribosome. However, many RNAs do not contain ORFs, and are therefore called non-coding RNAs. Instead of being templates for translation, non-coding RNAs harbor enzymatic, structural, or regulatory activity, the latter of which is the focus of this thesis.

Gene regulation

The amount of a specific protein in a cell at any given time is determined by the rate by which it is produced, and the rate by which it is degraded. In all kingdoms of life, production and degradation of proteins are highly controlled processes. Since the protein production rate is dependent on the

amount of the corresponding mRNA, cellular protein levels can be set by varying mRNA production and degradation rates. Control of mRNA production is called transcriptional regulation. In bacteria, transcriptional regulation is generally carried out by proteins called transcription factors (TFs). Post-transcriptional regulation refers to processes that control the rate by which mRNAs are translated into proteins, and the rate of mRNA degradation. Control of protein synthesis is achieved by proteins, RNAs or metabolites that specifically bind to mRNAs to increase or decrease the rate of translation, or by abiotic factors such as temperature and pH. Degradation of mRNAs is performed by enzymes called RNases. Processes that control the activity or stability of proteins are referred to as post-translational regulation, but these will not be covered in this thesis. The multiple levels of gene regulation are schematically outlined in Figure 1.

Gene regulation by transcription factors

TFs are proteins that bind to specific DNA sequences in the vicinity of a transcription unit to enhance or repress RNA polymerase activity and hence alter the transcription initiation rate. The first demonstration of transcription factor regulation was through the discovery of the *lac* operon in *Escherichia coli* (*E. coli*), which since then has served as the classical example of this kind of gene control (Jacob and Monod, 1961).

Control of gene expression is often a response to changes in the internal or external environment. In bacteria, this is exemplified by so called two-component systems, in which histidine kinases and response regulators transduce external signals to cause a response in gene expression. Upon sensing a stimulus in the environment, the membrane-bound histidine kinase autophosphorylates, and then catalyzes the transfer of a phosphate group to the response regulator (a TF), which in turn binds to specific DNA recognition motifs to regulate transcription. Three important *E. coli* TFs relevant for this thesis are CsgD, Lrp and OmpR. CsgD is a major regulator of biofilm formation that controls the expression of the extracellular matrix components curli and cellulose. Lrp is a global TF that, in response to the nutritional status, controls several hundred genes in *E. coli*. The acronym Lrp stands for leucine responsive regulatory protein since cellular leucine can change the regulatory output of Lrp action at target promoters. OmpR is the response regulator of the EnvZ-OmpR two-component system, whose activity is controlled by the osmolarity in the surrounding environment.

Regulation of RNA stability

Protein production is one of the energetically most expensive processes in the cell, and its effectiveness must therefore be optimal. In order to produce only the proteins that are needed in a given situation, and in the appropriate

amounts, mRNAs that are not needed must be rapidly removed. The life-times of different *E. coli* mRNAs differ substantially, with half-lives from less than a minute up to hours. In *E. coli* and many other bacteria, RNAs are mainly degraded through the action of a multi-protein complex referred to as the RNA degradosome, consisting of the endoribonuclease E (RNase E), the RNA helicase RhlB, enolase (a glycolytic enzyme), and the 3'-5' exonuclease polynucleotide phophorylase (PNPase) (reviewed in (Carpousis, 2007)). RNase E contains an N-terminal domain with endoribonuclease activity and a C-terminal "scaffold" domain at which RhlB, enolase and PNPase associates to form the degradosome. Other proteins are also known to interact with the C-terminal domain of RNase E, for example the RNA-binding protein Hfq. Generally, mRNA degradation is initiated by the action of the pyrophosphohydrolase RppH, which converts the triphosphate at mRNA 5'-ends to monophosphate (Celesnik *et al.*, 2007; Deana *et al.*, 2008). Subsequently, RNase E associates with monophophorylated 5'-ends to cleave mRNAs at internal sites, and the resulting fragments are finally degraded through 3'-5' exoribonucleases (an enzyme with 5'-3' exoribonuclease activity has so far not been found in *E. coli*). For mRNAs, the rate of degradation is often inversely correlated with the rate of translation (Arnold *et al.*, 1998), probably because high ribosome occupancy during active translation can protect mRNAs from RNases (reviewed in (Kaberdin and Bläsi, 2006)). Since substrate binding and activity of RNases are sequence and/or structure-dependent, the stability of a specific RNA depends on its properties. Secondary structure changes due to temperature shifts may alter the translation efficiency and hence the stability (Morita *et al.*, 1999; Morita *et al.*, 1999; Afonyushkin *et al.*, 2003; Giuliodori *et al.*, 2010). mRNA stability is also affected by the action of base-pairing small RNAs and small ligands (see below).

RNA-mediated post-transcriptional gene regulation

The first examples of RNA molecules with gene regulatory functions were discovered more than 30 years ago. However, the prevalence and impact of this kind of regulation was not appreciated until much more recently. Today we know that RNA-mediated gene regulation is found in all kingdoms of life, perhaps even in all species. The first findings showed that small plasmid-encoded antisense RNAs were involved in the control of plasmid replication (Stougaard *et al.*, 1981; Tomizawa *et al.*, 1981). Shortly thereafter, a chromosomally encoded small antisense RNA in *E. coli*, later named MicF, was shown to inhibit translation and to induce degradation of the *ompF* mRNA (Mizuno *et al.*, 1984). At the beginning of this century, the *E. coli* genome was systematically searched for putative regulatory RNAs, revealing a large number of small RNAs of, at the time, unknown function (Argaman *et al.*, 2001; Wassarman *et al.*, 2001). This finding suggested that gene regu-

lation by small RNAs might be a much more general phenomenon than previously anticipated. The subgroup of bacterial small RNAs that act as regulators of gene expression are now often referred to as sRNAs. In eukaryotes, microRNAs (miRNAs) are a large class of non-coding RNAs that, in a similar fashion to many bacterial small RNAs, use antisense basepairing to induce translational repression and/or degradation of target mRNAs. The first miRNA to be discovered was *lin-4* in *C. elegans* (Lee *et al.*, 1993). It is now established that miRNAs are central regulators of processes such as development and disease (Ambros, 2011).

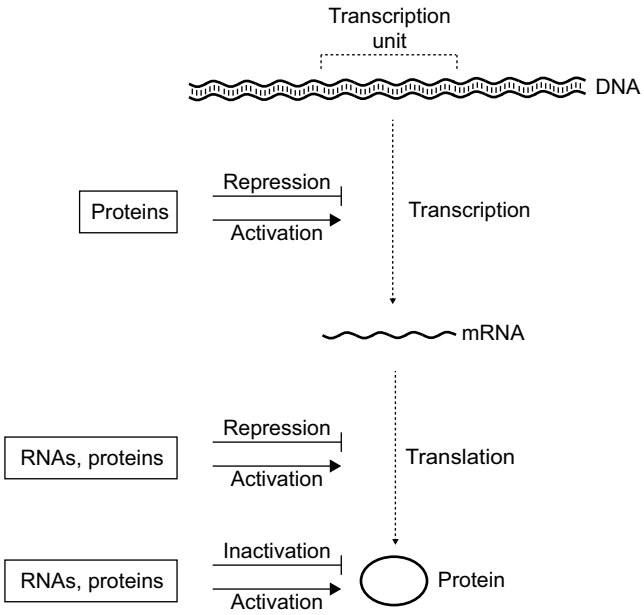


Figure 1. Three different layers of gene expression are controlled by protein or RNA regulators.

Many bacterial sRNAs and all eukaryotic miRNAs regulate expression of mRNAs that are transcribed from a different genetic locus than the regulator itself, and are therefore referred to as *trans*-encoded regulators (Figure 2). There are also examples of *cis*-encoded sRNAs. They are transcribed, in opposite orientation, from the same locus that encodes their target RNAs, thus generating fully complementary RNAs (Figure 2). Since both of these types of regulators control the activity of separate RNA molecules, they are *trans*-acting (Figure 2). However, there are also *cis*-encoded *and cis*-acting RNA elements (Figure 2), for example riboswitches and RNA thermometers. A riboswitch is a 5'-segment of an mRNA that senses the concentration of a

small ligand, and upon ligand binding alters the activity of the mRNA by either controlling transcription elongation or translation (Mironov *et al.*, 2002; Nahvi *et al.*, 2002; Winkler *et al.*, 2002; Winkler *et al.*, 2002; Tucker and Breaker, 2005). RNA thermometers are RNA secondary structure elements which control mRNA activity with respect to temperature (Altuvia *et al.*, 1989; Storz, 1999; Johansson *et al.*, 2002).

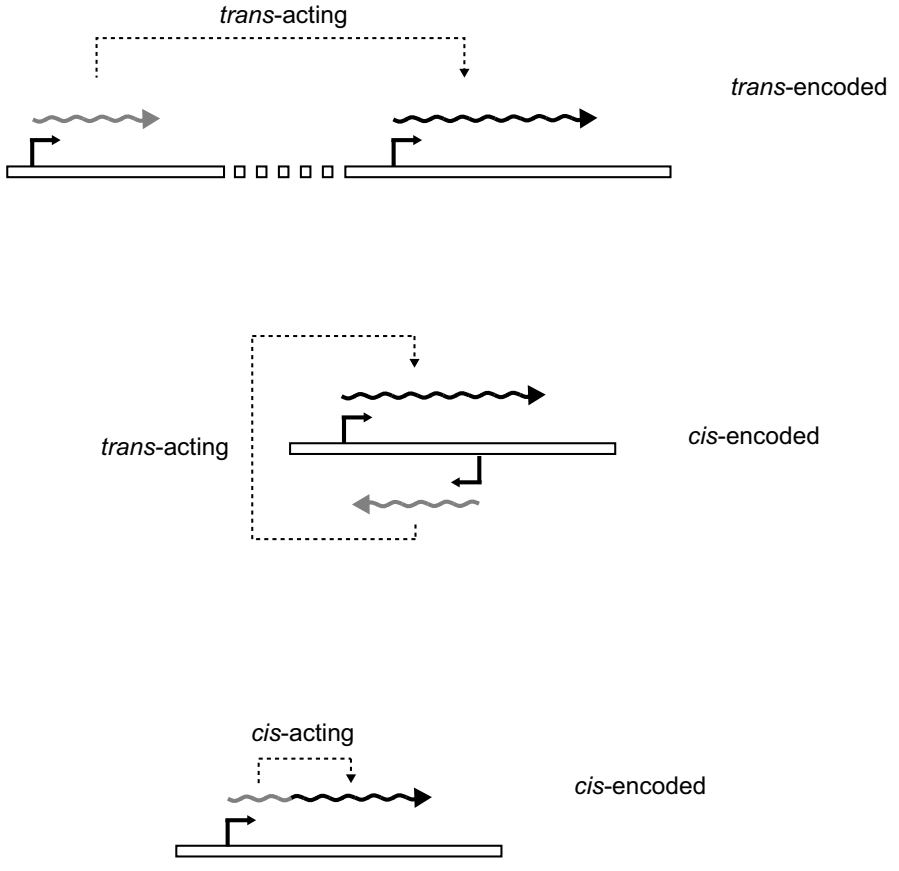


Figure 2. Regulatory RNAs can act on other RNA molecules (*trans*-acting) or on themselves (*cis*-acting), and can be expressed from a separate locus (*trans*-encoded) or the same locus as the target RNA (*cis*-encoded). Regulatory RNA, mRNA, and DNA are shown in grey arrows, black arrows, and white boxes, respectively. Bent black arrows indicate transcription start sites. Dotted lines with arrows indicate directionality in regulation.

Bacterial small RNAs

The use of the nomenclature regarding *E. coli* small RNAs and their subgroups is not always consistent, and the distinction between RNAs classified as belonging to different groups is not always sharp. Generally, small RNAs are considered to be those that are much shorter than the average size of mRNAs. The majority of the characterized small RNAs are also non-coding, that is, they lack the ability to code for protein expression. However, there are examples of small RNAs with protein coding capacity, and such small RNAs are by definition also mRNAs (e.g. (Benito *et al.*, 2000; Wadler and Vanderpool, 2007)). Many if not most sRNAs are antisense RNAs, which implies that they contain sequences that are complementary to one or several other RNAs. However, antisense RNAs come in many sizes and are not restricted to the class of small RNAs. Conceptually, a strand of DNA that is transcribed into an mRNA is called sense. Transcription from an mRNA-expressing locus, but from the opposite DNA strand (the antisense strand), generates a *cis*-encoded antisense RNA that is fully complementary to the mRNA. However, antisense RNAs can also be *trans*-encoded, which often results in limited complementarity to sense RNAs. The term sRNA is used for a variety of small RNAs, both non-coding and coding, and with different cellular functions. The sRNAs analyzed in the studies on which this thesis is built, are non-coding *trans*-encoded small RNAs that regulate gene expression by using antisense mechanisms.

sRNAs in *cis* and *trans*

At present, sRNAs have been found in all bacterial phyla investigated, and therefore seem to be a widespread phenomenon. The bacterial species most intensely studied with respect to sRNA biology, *E. coli* and *Salmonella* Typhimurium, each harbor at least 100 sRNA genes, a number that may be greatly underestimated (Raghavan *et al.*, 2011). In contrast to the eukaryotic miRNAs, bacterial sRNAs are heterogeneous in size and structure with lengths ranging between 50 and 400 nucleotides. Despite this heterogeneity, sRNAs share some common features. As is the case for many other genes, transcription of almost all sRNA genes ends at Rho-independent terminators, consisting of stable stem-loops followed by a run of uridine residues (Figure 3). Metazoan miRNA-mRNA interactions rely on so called seed regions, which are six to seven nucleotide stretches at which the miRNA-mRNA interaction initiates and that have to be perfectly complementary to target sequences for efficient regulation (Lewis *et al.*, 2005). Although seed regions obeying a strict definition have not been identified in bacteria, many bacterial *trans*-encoded sRNAs use unstructured 5'-tails for binding to complementary sequences in multiple mRNA targets (Figure 3). RybB, OmrA, OmrB, MicA and MicF are examples of such sRNAs ((Guillier and

Gottesman, 2006; Papenfort *et al.*, 2006; Guillier and Gottesman, 2008; Papenfort *et al.*, 2010; Gogol *et al.*, 2011) and Paper I-III, Figure 3). Finally, the majority of the so far characterized *trans*-encoded sRNAs, at least in enterobacteria, needs to interact with the RNA-binding protein Hfq to efficiently carry out their regulatory functions (see below).

In addition to the large class of *trans*-encoded sRNAs, recent data suggest that chromosomally *cis*-encoded sRNAs may be more common than previously anticipated (Kawano *et al.*, 2005; Han *et al.*, 2010; Lorenz *et al.*, 2010; Shinhara *et al.*, 2011). In contrast to the well-studied and now classical examples of *cis*-encoded antisense RNAs in plasmids (reviewed in (Wagner *et al.*, 2002)), the functions of most *cis*-encoded sRNAs expressed from bacterial chromosomes are poorly understood.

Other small RNAs

E. coli also harbors small RNAs that exert their functions without base-pairing to mRNAs. The 6S RNA controls sigma factor availability by mimicking an open σ^{70} promoter complex (reviewed in (Wassarman, 2007)). The tmRNA rescues stalled ribosomes and promotes degradation of aberrant mRNAs (reviewed in (Dulebohn *et al.*, 2007)). CsrB and CsrC are small RNAs that sequester the regulatory protein CsrA, a key post-transcriptional regulator involved in many physiological responses, and thereby decrease its active concentration (reviewed in (Babitzke and Romeo, 2007)). Since the CsrA protein binds to many mRNAs to control protein synthesis, CsrB and CsrC indirectly affect gene expression.

The RNA-binding protein Hfq

Hfq is a homohexameric RNA-binding protein, first identified as a host factor for replication of phage Qβ (Franze de Fernandez *et al.*, 1968). However, Hfq has many other important roles in physiology and virulence in many bacterial species, often through interactions with sRNAs. Hfq seems to be rather unspecific in terms of RNA binding partner types, since binding to sRNAs, mRNAs, tRNAs and rRNAs has been observed (Vecerek *et al.*, 2003; Zhang *et al.*, 2003; Lee and Feig, 2008; Sittka *et al.*, 2008). *In vivo* and *in vitro*, Hfq forms a homohexameric ring structure where each monomer adopts the same conformation (Schumacher *et al.*, 2002; Sauter *et al.*, 2003; Nikulin *et al.*, 2005). Hfq has two RNA-binding surfaces, called the proximal and distal faces. The proximal face has high affinity for U-rich sequences, while single-stranded ARN sequences (R=purine and N=any base) are preferentially bound on the distal face (Schumacher *et al.*, 2002; Link *et al.*, 2009). Many sRNAs are unable to regulate targets in the absence of Hfq. There are probably several reasons for this. First, since Hfq has been shown to speed up the rate of sRNA-mRNA binding, loss of regulation in

Hfq-deficient strains may be due to slow rates of duplex formation (Soper and Woodson, 2008; Fender *et al.*, 2010). Second, since many sRNAs become very unstable in the absence of Hfq, the lower steady-state levels of the sRNAs due to instability may also account for loss of regulation (Paper II and (Fender *et al.*, 2010)). The stabilization of sRNAs by Hfq can be explained by the similarity between Hfq-binding sites and RNase E cleavage sites which are both A/U-rich, and has been shown to coincide on both mRNAs and sRNAs (Moll *et al.*, 2003). Additionally, sRNA-Hfq complexes guide RNase E to target mRNAs for degradation (Morita *et al.*, 2005; Pfeiffer *et al.*, 2009).

Estimates of intracellular levels of Hfq range from 400 to 10,000 hexamers per cell (Carmichael *et al.*, 1975; Kajitani *et al.*, 1994; Ali Azam *et al.*, 1999), and each sRNA is present in ten to a thousand molecules per cell (Altuvia *et al.*, 1997; Pfeiffer *et al.*, 2007; Papenfort *et al.*, 2009). Because of the plethora of sRNA interaction partners, not to mention mRNAs and tRNAs, Hfq is probably saturated with RNA binding partners at all times in the cell. One intriguing aspect is then how RNAs associate to, and dissociate from, this protein. Measured dissociation rates for RNA-Hfq complexes *in vitro* suggest half-lives of more than one hundred minutes (Fender *et al.*, 2010). This is however inconsistent with the fast sRNA regulation observed *in vivo*, which sometimes is as fast at one to two minutes (Papenfort *et al.*, 2006). Fender *et al*. have proposed a model to solve this paradox, in which RNAs actively cycle on Hfq (Fender *et al.*, 2010). In the presence of RNA competitors, *in vitro* half-lives of RNA-Hfq complexes drop to approximately one minute (Fender *et al.*, 2010). A recent study suggests that the *in vitro* assembly of the Hfq hexamers from monomers is coupled to RNA binding, and that Hfq concentration strongly affects hexamer formation (Panja and Woodson, 2012). The observed post-transcriptional autoregulation of Hfq may thus ensure hexamer-promoting concentrations *in vivo* (Vecerek *et al.*, 2005).

Regarding physiology, *hfq* deletion strains show various defects, often through poor activity of Hfq-dependent sRNAs (Tsui *et al.*, 1994; Sittka *et al.*, 2008). The σ^S-dependent response to stress is impaired due to decreased translation of the *rpoS* mRNA, which is normally activated by sRNAs (Muffler *et al.*, 1996). The membrane stress response, mediated by σ^E, also becomes impaired in an *hfq* deletion strain because of inappropriate accumulation of outer membrane proteins, which otherwise would be repressed by sRNAs (Papenfort *et al.*, 2006). Deletion of *hfq* in *Salmonella* entails misregulation of 20% of the transcriptome, impairing many cellular functions, and results in a non-motile state (Figueroa-Bossi *et al.*, 2006; Sittka *et al.*, 2007). *hfq* mutants also display several virulence defects, such as attenuation of invasion of host cells and infection of mice, lack of secretion of virulence factors, and decreased survival in macrophages (Sittka *et al.*, 2007).

sRNAs investigated in the present study

I have studied several *E. coli* sRNAs in more detail, all of which are of the *trans*-encoded type. They are called MicA, MicF, OmrA and OmrB. The following section introduces these four sRNAs, and briefly summarizes the present knowledge of their biological properties.

MicA

The sRNA initially named SraD was discovered as a 67 nucleotide long transcript which is conserved in many enterobacteria, and was shown to be induced in the stationary growth phase (Argaman *et al.*, 2001; Zhang *et al.*, 2003). SraD was later shown to inhibit synthesis of outer membrane protein A (OmpA) ((Rasmussen *et al.*, 2005) and Paper I), and was therefore renamed MicA, in line with two other sRNAs that control expression of OMPs (MicF/*ompF* and MicC/*ompC*). Transcription of MicA is strictly dependent on sigma factor σ^E (Johansen *et al.*, 2006; Papenfort *et al.*, 2006; Udekwu and Wagner, 2007), which is activated upon envelope stress due to accumulation of misfolded outer membrane proteins (reviewed in (Ruiz and Silhavy, 2005)). OmpA was the first MicA target to be found (Paper I), but since then many additional targets have been identified (Bossi and Figueroa-Bossi, 2007; Coornaert *et al.*, 2010; Udekwu, 2010; Gogol *et al.*, 2011). Recently it was demonstrated that MicA, together with the sRNA RybB, endows a post-transcriptional repressor arm in the σ^E regulon (Gogol *et al.*, 2011).

OmrA and OmrB

In screens for novel sRNAs in *E. coli*, the sRNAs RygA and RygB were initially found as 88 and 82 nucleotide long transcripts, respectively (Argaman *et al.*, 2001; Wassarman *et al.*, 2001; Vogel *et al.*, 2003). Transcription of both sRNAs is dependent on the OmpR-EnvZ two-component system, and they were therefore renamed to OmrA and OmrB (for OmpR-regulated sRNA A and B (Guillier and Gottesman, 2006)). The *omrA* and *omrB* genes reside next to each other between the *aas* and *galR* genes on the *E. coli* chromosome and have probably arisen from a gene duplication event (Argaman *et al.*, 2001; Skippington and Ragan, 2012). Comparison of the OmrA and OmrB sequences reveals identical 5'- and 3'-ends, while the middle segments differ remarkably between the two RNAs (Paper II). The identical 5'-segments contain the sequences that are used for binding to their shared targets ((Guillier and Gottesman, 2008) and Paper II and IV, Figure 3). Both RNAs bind the Hfq protein with dissociation constants in the low nanomolar range, and are Hfq-associated *in vivo* ((Zhang *et al.*, 2003) and Paper II). OmrA and OmrB inhibit synthesis of the CirA, OmpT and OmpR proteins; OmpR is a positive regulator of *omrA* and *omrB* expression

(Guillier and Gottesman, 2006; Guillier and Gottesman, 2008). Recently, OmrA and OmrB were also shown to inhibit synthesis of CsgD and FlgM (Paper II and IV). Expression of these sRNAs is induced during the transition between the exponential and stationary growth phases (Argaman *et al.*, 2001; Wassarman *et al.*, 2001; Vogel *et al.*, 2003).

MicF

Almost three decades ago, the first chromosomally encoded sRNA was discovered. Shortly thereafter, this sRNA was shown to inhibit synthesis of the outer membrane protein F (OmpF), by binding to a complementary region in the *ompF* mRNA. It was therefore named MicF (for "mRNA-interfering-complementary RNA for *ompF*" (Mizuno *et al.*, 1984)). MicF is a 93 nucleotide long Hfq-associated transcript that originates from the region upstream of the *ompC* gene. Its expression is controlled by several TFs, in particular H-NS, Lrp, HU (repression) and OmpR, SoxS, MarA, Rob (activation) (Delihas and Forst, 2001; Zhang *et al.*, 2003). Steady-state levels of MicF are regulated in response to environmental factors such as temperature, oxidative stress, nutritional status and osmolarity (Delihas and Forst, 2001). Although MicF has been thoroughly characterized, no target mRNAs other than *ompF* were until recently described. It is now clear, however, that MicF has multiple targets (Paper III and (Corcoran *et al.*, 2012)).

Properties of sRNA regulation

One intriguing question concerns why cells at all use sRNAs to regulate gene expression, instead of only proteins as was previously assumed. Suggestions to answer this question deal with the energy cost of producing regulators and the modes of regulation. As regulators, sRNAs may be favored over proteins since they are energetically cheaper to produce (Altuvia and Wagner, 2000). To produce a regulatory protein, the cell needs to transcribe an mRNA of roughly ten times the size of an sRNA, and subsequently translate the mRNA into the final protein product. However, since a TF shuts down transcription of target genes, whereas an sRNA can act only after the target mRNA is already present, the energy saved by transcribing an sRNA is in part lost in the synthesis of the target mRNA. Concerning modes of regulation: by modeling regulation mediated by sRNAs versus TFs, it has been suggested that regulation by sRNAs is faster (Shimoni *et al.*, 2007). However, the fastest way of regulation is when the regulator is already present, for example through protein-protein interactions. In this case, the regulator is present in an inactive form, ready to support a very rapid response when activated (Shimoni *et al.*, 2007). In contrast to protein-based regulation, where expression of the target gene only depends on promoter activity,

sRNA regulation shows a threshold-linear behavior in which the abundance of target mRNAs affects the level of repression (Levine *et al.*, 2007).
Thus, depending on expression of the target, sRNA-mediated regulation can have a strong effect or no effect at all (Levine *et al.*, 2007). One advantage of using sRNAs for regulation is that it can counteract noise. If a transcriptionally controlled gene is accidently fired, a few mRNAs will be produced, which will be translated into even more protein molecules (Levine *et al.*, 2007). By contrast, failure of repression at the post-transcriptional level will not be amplified in a second step. Additionally, sRNA regulation was recently shown to reduce leakiness and to alter the temporal dynamics of target gene expression (Beisel and Storz, 2011). Since gene expression is often controlled not only by one but several regulators, and on several levels, one interesting aspect concerns the properties of combinatorial regulation.

sRNAs in regulatory networks

When studying gene regulation, the regulator and the regulated gene have traditionally been treated as isolated entities. In a complex regulatory system such as a cell this is of course an oversimplified view. It is now obvious that genes and their products are connected in mesh-like networks with several layers of hierarchy. When small subsets of gene regulatory networks are studied in detail, regulatory motifs (often referred to as network motifs) are revealed (Alon, 2007). Some motifs are found much more frequently in gene regulatory networks than in randomly generated networks, suggesting that these motifs have been successfully established throughout evolution (Shen-Orr *et al.*, 2002). Figure 4 shows some of the motifs that are often found in regulatory networks, including feed-forward and feed-back variants. The vast majority of research on network motifs has been carried out on systems in which the regulators are proteins, whereas network motifs that include both sRNAs and proteins (so called mixed motifs) have only recently begun to be understood (Shimoni *et al.*, 2007; Beisel and Storz, 2011). However, there are a few examples of properties of network motifs that include sRNAs. The classical sRNA MicF forms a feed-forward loop with its target *ompF* and the transcription factor OmpR, which activates expression of *micF* and simultaneously represses expression of *ompF* (Shimoni *et al.*, 2007). Expression of *ompF* is high during low osmolarity and low during high osmolarity (Lugtenberg *et al.*, 1976).

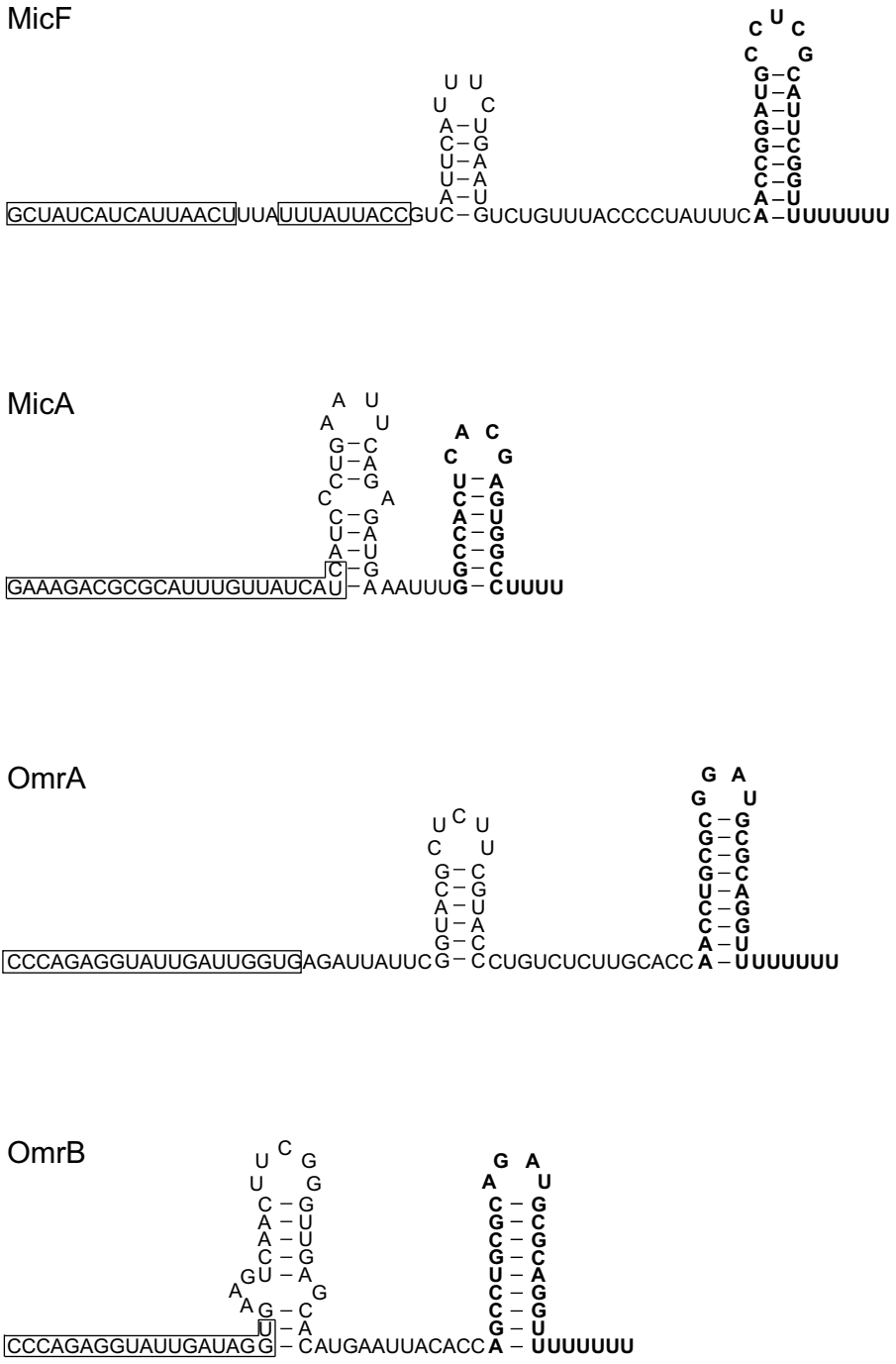


Figure 3. Secondary structure representations of the four sRNAs studied in this thesis. Nucleotides that are implicated in binding to target mRNAs are boxed. Nucleotides encompassing Rho-independent terminators are in bold letters.

During shifts from low to high osmolarity, OmpF synthesis is rapidly repressed (Aiba *et al.*, 1987). Interestingly, the time to reach a new steady-state level of *ompF* expression is dependent on MicF, since a longer response time is observed in a *micF* deletion strain (Aiba *et al.*, 1987). The sRNA MicC also forms a feed-forward loop with its target OmpC, and its regulator OmpR. In this case OmpR represses expression of *micC* and activates expression of *ompC* (Shimoni *et al.*, 2007). MicF also participates in a mixed double negative feedback loop together with its target and regulator Lrp (Paper III). In this case, MicF inhibits Lrp synthesis post-transcriptionally, while Lrp represses *micF* expression on the transcriptional level ((Ferrario *et al.*, 1995) and Paper III). Another example of a mixed double negative feedback loop is that of the sRNA RyhB and its repressor, the iron-responsive TF Fur (Vecerek *et al.*, 2007).

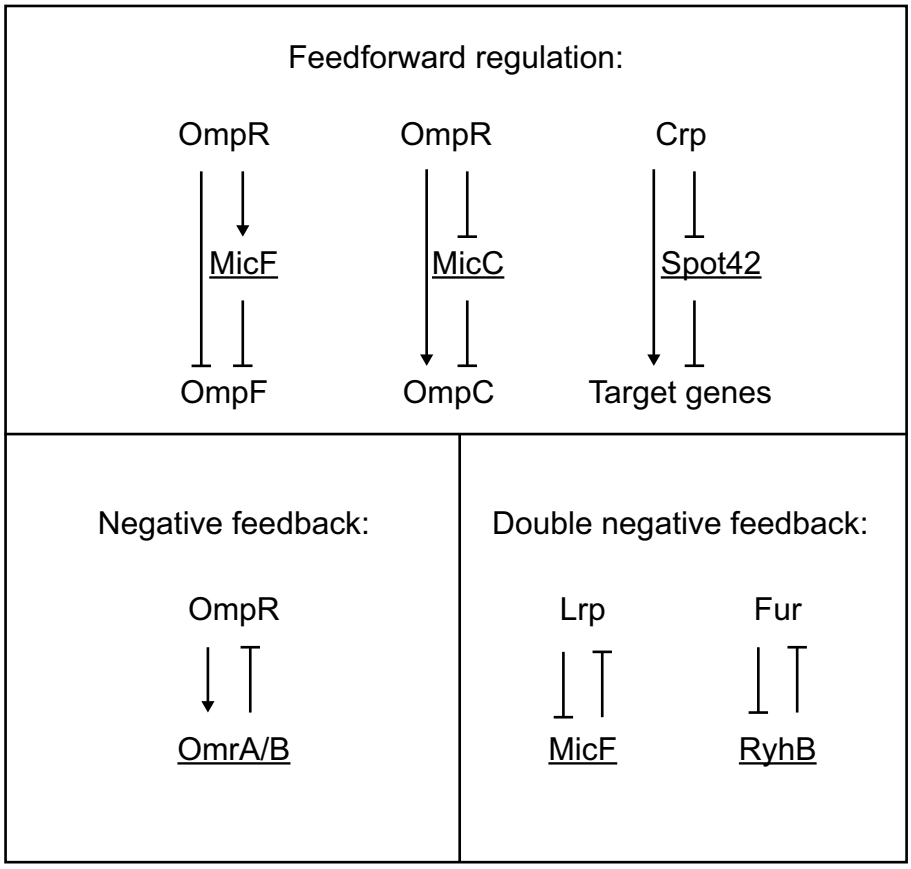


Figure 4. Examples of gene network motifs that contain sRNAs (underlined).

Present study

The studies that this thesis is based on deals with three major aspects: identification of mRNAs that are targeted by sRNAs, mechanisms of sRNA action, and physiological impact of sRNAs. With these three aspects as major themes, the results presented in Papers I-IV are summarized below, and discussed in relation to the present literature.

Identification of target mRNAs

Since most of the so far characterized sRNAs regulate protein synthesis by binding to target mRNAs, it is reasonable to assume that this may also apply to so far not characterized small RNAs. For some sRNAs, only one or a few targets have yet been found, which raises the question whether they have other, so far unknown targets. To identify novel targets, one can use several different experimental and computational approaches.

Experimental approaches

Experimental approaches such as global gene expression analyses using microarrays or deep RNA-sequencing have been fruitful in identifying mRNAs that are targeted by sRNAs (e.g. (Guillier and Gottesman, 2006; Papenfort *et al.*, 2006; Sharma *et al.*, 2011)). The success of this approach relies on the observation that target mRNA levels usually decrease (repression) or increase (activation) upon sRNA induction, irrespective of the molecular mechanism of sRNA action. Proteomic approaches such as two-dimensional protein gel electrophoresis have also revealed novel sRNA targets ((Sharma *et al.*, 2007) and Paper I). Although such methods potentially may detect many target mRNAs simultaneously, they have limitations. Since regulation can be dependent on the growth condition used in the particular experiment, target mRNAs that are only expressed under specific environmental conditions may not be detected. For example, the *csgD* mRNA would not be detected in such an experiment unless samples were taken from cultures in late stationary phase grown at 28°C (Paper II). Likewise, the *phoE* mRNA is only expressed at low phosphate concentrations, which does not apply to standard growth media such as LB and M9, and would therefore easily be missed (Paper III). sRNA-dependent regulation often relies on in-

hibition of translation followed by degradation of target mRNAs. However, the degree of sRNA-dependent mRNA degradation varies from case to case. Target mRNAs that are translationally inhibited, but not destabilized, will not be identified by methods that monitor mRNA expression levels. Proteome analyses using two-dimensional protein gel electrophoresis have the disadvantage that only proteins present at sufficiently high levels can be detected. However, by using stable isotope amino acid labeling and mass spectrometry, it is now possible to improve proteome coverage significantly (Ong *et al.*, 2002). Combining RNA sequencing and mass spectrometry read-outs will certainly be fruitful in future analyses of sRNA-dependent gene expression. In addition to global expression analyses, genetics have in some cases successfully identified new sRNA targets (Figueroa-Bossi *et al.*, 2009).

In silico target prediction approaches

An alternative to experimental approaches is to use *in silico* methods. There are several publicly available computational algorithms that aim at predicting interactions between bacterial sRNAs and mRNAs (Busch *et al.*, 2008; Tjaden, 2008; Cao *et al.*, 2009; Eggenhofer *et al.*, 2011; Ying *et al.*, 2011). We have used an in-house developed algorithm called *AntisenseRNA* to predict putative sRNA-mRNA interactions (Reimegård, 2010). This algorithm is based on two general rules:

- A short initial intermolecular structure is formed between the sRNA and the target. The nucleotides that interact should be single-stranded.
- Interaction is directional; it propagates in one direction from the initial intermolecular structure and allows for bulges and loops.

Among the structures that fulfill the rules, the most stable sRNA-mRNA interaction is selected. The stability is calculated by both considering the formed intermolecular structure and the changes of the intramolecular structures. *AntisenseRNA* can also include information from closely related species, where the sRNA and the target exist, by reducing the stability of intermolecular structures that are not conserved.

To increase the probability of finding true targets, we chose to use *AntisenseRNA* to predict targets for sRNAs that had at least one previously verified target and which the algorithm ranked among the 20 best candidates.

Mechanisms employed by *E. coli* sRNAs

The majority of sRNAs in *E. coli* control gene expression by antisense, i.e. by binding to mRNAs. However, the outcome of sRNA-mRNA pairing varies dependent on where the interaction site is located within the mRNA, and which proteins are associated with the sRNA (Figure 5).

sRNAs that bind to the ribosome binding site

The so far most commonly described mechanism of sRNA-mediated regulation relies on binding to the ribosome binding site (RBS) of target mRNAs. In bacteria, a ternary complex consisting of an mRNA, the 30S ribosomal subunit, and fMet-tRNA^{fMet} is formed to initiate translation of mRNAs into proteins. Ternary complex formation is achieved through interactions between the Shine and Dalgarno (SD) sequence of the mRNA and the 16S ribosomal RNA, and between the start codon of the mRNA and the anticodon of fMet-tRNA^{fMet}. The RBS is the segment of the mRNA which is recognized by the 30S ribosomal subunit and fMet-tRNA^{fMet} during translation initiation, and includes the SD sequence, the start codon, and additional flanking sequences via backbone contacts (Hüttenhofer and Noller, 1994). However, in contrast to canonical mRNAs, leaderless mRNAs are transcripts that, despite that they lack a SD sequence and start with an AUG start codon, are translationally active *in vivo* (reviewed in (Moll *et al.*, 2002)).

Binding of sRNAs within the RBS region sterically hinders association of 30S-tRNA^{fMet} and leads to inhibition of translation. In a recent study, it was suggested that sRNAs may inhibit translation by binding to the coding region of the mRNA. However, for efficient inhibition of translation, the sRNA binding site must be within the first five codons of the target mRNA, referred to as the "five codon window" (Bouvier *et al.*, 2008). Sterical hindrance of 30S-tRNA^{fMet} binding may also occur if the sRNA binding site is located upstream, but in close vicinity of the SD sequence (Chen *et al.*, 2004). This fits well with studies of 30S-tRNA^{fMet} occupancy on mRNAs when bound to the start codon (Hüttenhofer and Noller, 1994). Regulation by binding to the RBS ("RBS-targeting") applies to MicA when inhibiting OmpA synthesis (Paper I), MicF when inhibiting Lrp synthesis (Paper III), and OmrA and OmrB when inhibiting FlgM synthesis (Paper IV).

MicA inhibits OmpA synthesis through base-pairing (Paper I)

In an attempt to find proteins whose expression was dependent on MicA levels, the entire protein content from cells that overexpressed MicA was analyzed by 2D-gel electrophoresis, and compared to that of control cells. In this experiment, the levels of the OmpA protein were 10-fold lower in a strain that overexpressed MicA, indicating direct or indirect regulation of OmpA expression by MicA. Overexpression of MicA also resulted in an

Hfq-dependent decrease in *ompA* mRNA levels, and a strain deleted for the *micA* gene showed higher *ompA* mRNA levels in stationary growth phase, i.e. at a time at which MicA is induced. The *AntisenseRNA* algorithm was used to search for putative base-pairing between MicA and all *E. coli* mRNAs. Strikingly, *ompA* turned up as the best hit, and a region encompassing nucleotides 8-24 of MicA was predicted to interact with a complementary sequence in the *ompA* 5'UTR (including the SD sequence). Analyses of *ompA-lacZ* translational fusions showed that base pair changes in the predicted binding site of either MicA or *ompA* abolished regulation, while a combination of mutations allowing for full sequence complementarity restored regulation. The interaction site was confirmed by *in vitro* foot-printing using partial chemical or enzymatic digestion, and *in vitro* mobility shift assays demonstrated duplex formation. Finally, MicA was shown to sterically interfere with 30S-tRNA^{fMet} binding to the *ompA* mRNA, suggesting that inhibition of translation is the mechanisms by which MicA controls expression of OmpA protein levels.

Inhibition of Lrp synthesis by MicF (Paper III)

According to *AntisenseRNA* predictions, the *lrp* mRNA was one of the top candidates to interact with the sRNA MicF. The predicted interaction site in *lrp* consists of thirteen nucleotides and overlaps the AUG start codon. Hence the interaction is within the "five-codon window" and could potentially interfere with initiation of translation. Data that supported such an inhibition came from *in vivo* experiments that showed that deletion of the chromosomal *micF* gene resulted in elevated Lrp levels, whereas overexpression of MicF from a plasmid strongly reduced Lrp levels. As observed for many other sRNAs, deletion of *hfq* resulted in total loss of regulation. Several pieces of evidence support the hypothesis that the MicF-dependent effect on *lrp* expression is due to direct basepairing between the RNAs, as predicted.

First, mutations of nucleotides in the predicted binding site of MicF, or deletion of the entire binding site, resulted in abolished repression of *in vivo* Lrp levels. Even a single C to G mutation at position 2 of MicF resulted in loss of repression. However, when the complementary nucleotide in *lrp*, G-278, was changed to a C, regulation was restored. This result is a strong indication that basepairing between MicF and *lrp* is required for regulation.

Second, *in vitro* translation from an *lrp-3xflag* mRNA was inhibited upon pre-incubation with MicF, and the inhibition was Hfq-dependent. Additionally, binding of 30S-tRNA^{fMet} to *in vitro* transcribed *lrp* mRNA was inhibited after pre-incubation with MicF. Thus, binding of MicF to the *lrp* mRNA seems to sterically prevent formation of a translation initiation complex.

Third, with the use of partial chemical or enzymatic digestion of radiolabeled *lrp* or MicF, the interaction between these molecules was demonstrated and analyzed at nucleotide resolution. In conclusion, the analyses described above strongly indicate that MicF and *lrp* mRNA physically inter-

act in the cell, and that the outcome of the interaction is inhibition of Lrp synthesis.

OmrA and OmrB inhibit synthesis of FlgM (Paper IV)

Biocomputational predictions identified a putative interaction between the 5'-end of OmrA/OmrB and the *flgM* mRNA, which encodes an anti-sigma factor of the flagella-specific sigma factor σ^{28} (also known as FliA) (Ohnishi *et al.*, 1992). To test the possible effect of OmrA/B on FlgM synthesis, fluorescence from a translational *flgM::gfp* fusion was monitored in the presence or absence of OmrA or OmrB. Overexpression of either of the sRNAs resulted in significantly lower fluorescence levels, indicating OmrA/B-dependent control of FlgM synthesis. To address whether the observed regulation was due to a direct interaction between the *flgM* mRNA and OmrA or OmrB (as predicted), point mutations that would weaken the predicted interaction were introduced in both *flgM* and the sRNA sequences. Analysis of wild type and mutant constructs clearly indicated that OmrA/B-dependent inhibition of FlgM synthesis only occurs when target and sRNA sequences provide full complementarity.

Two different biocomputational algorithms (*AntisenseRNA* (Reimegård, 2010) and *IntaRNA* (Busch *et al.*, 2008)) suggested that OmrA/OmrB can bind to nucleotides 21 to 27 (codons 7 to 9) of the *flgM* mRNA. In addition *IntaRNA* predicted additional shorter stretches of complementarity within the *flgM* RBS. Binding within the RBS often results in inhibition of translation initiation (see above), while binding to coding sequences outside of the RBS has little effect on translation initiation (see below). *In vitro* translation experiments indicated that OmrA and OmrB independently can inhibit FlgM synthesis, tentatively suggesting that OmrA/B bind within the *flgM* RBS, as predicted *IntaRNA*. However, a direct demonstration of OmrA/OmrB-dependent inhibition of ternary complex formation on *flgM*, i.e. in "toe-print" experiments, will still need to be conducted to further strengthen this interpretation.

Binding to regions upstream of the RBS

There are a few reported cases in which sRNAs inhibit translation by binding to mRNA sequences far upstream of the RBS, rather than within the sequence that is normally bound by 30S-$tRNA^{fMet}$ during translation initiation. Binding of the sRNA IstR-1 to the *tisB* mRNA blocks a ribosome standby site at which the ribosome must bind prior to relocation to the RBS (Darfeuille *et al.*, 2007; Unoson and Wagner, 2007). The *tisB* SD sequence is sequestered in an intra-molecular structure which limits access to freely diffusing ribosomes for translation initiation. For efficient translation, ribosomes must therefore first associate with a single-stranded region upstream of the RBS (the standby site).

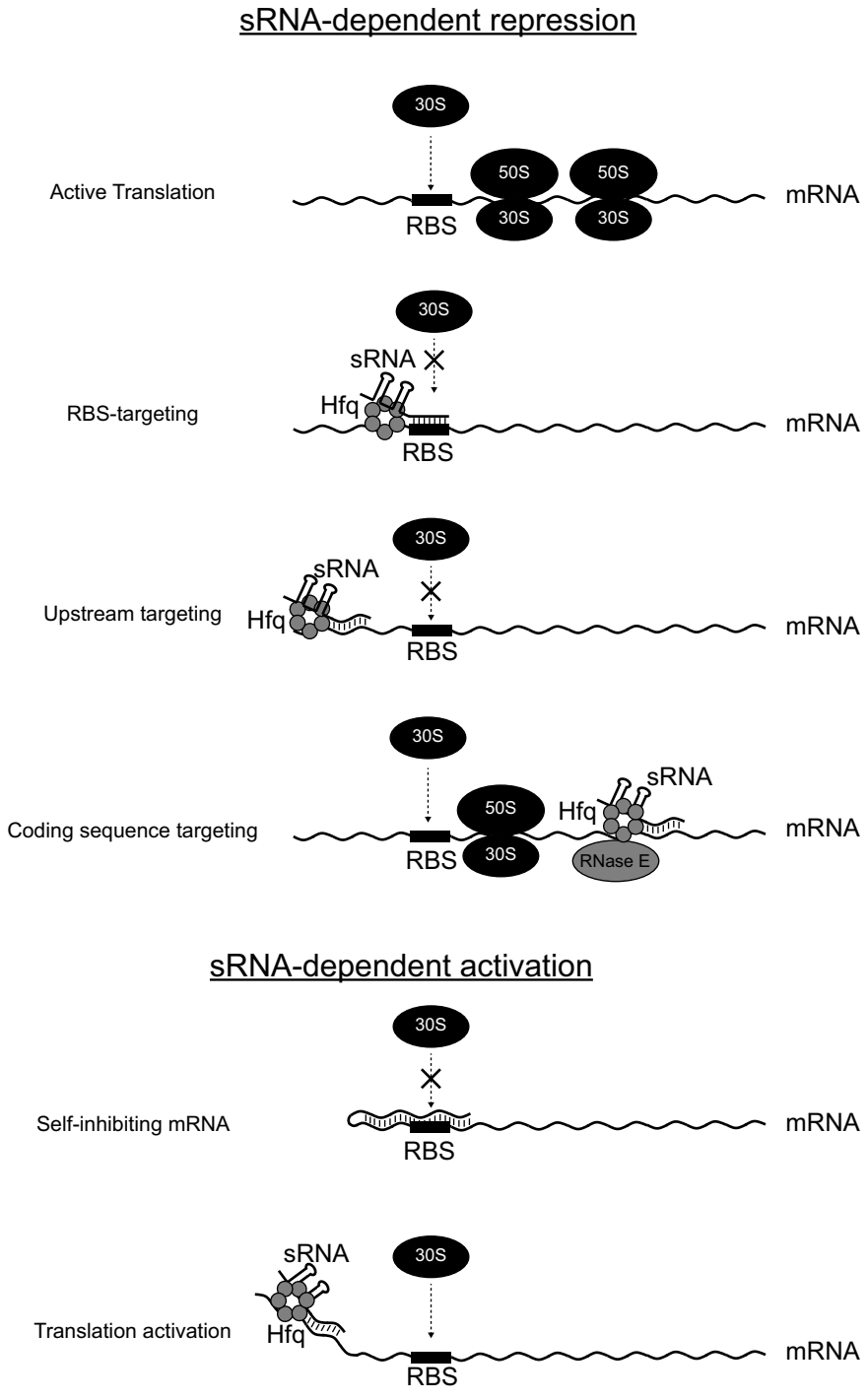


Figure 5. Regulatory mechanisms employed by bacterial sRNAs.

This increases the local concentration of ribosomes and allows for relocation to the RBS upon "breathing" of the inhibitory structure. By binding to the standby site, IstR thus inhibits translation initiation. A second example of upstream targeting is the inhibition of Fur synthesis by the sRNA RyhB. Similar to *tisB*, the SD sequence upstream of the *fur* ORF is sequestered in an intra-molecular structure (Vecerek *et al.*, 2007). However, in contrast to the standby mechanism, efficient Fur translation is achieved through translational coupling to an upstream ORF (*uof*). RyhB blocks translation of the *uof*, leading to inhibition of downstream Fur synthesis (Vecerek *et al.*, 2007). A similar phenomenon has been described for the *Pseudomonas* sRNA PhrS, but in this case translation of the *uof* upstream of the *pqsR* RBS is activated through structural changes upon PhrS binding (Sonnleitner *et al.*, 2011). A third example is the upstream targeting exerted by the sRNA GcvB on several target mRNAs. Here, GcvB base-pairs to CA-rich translation enhancer sequences in mRNA leaders, thereby decreasing translation rates of the downstream ORFs (Sharma *et al.*, 2007).

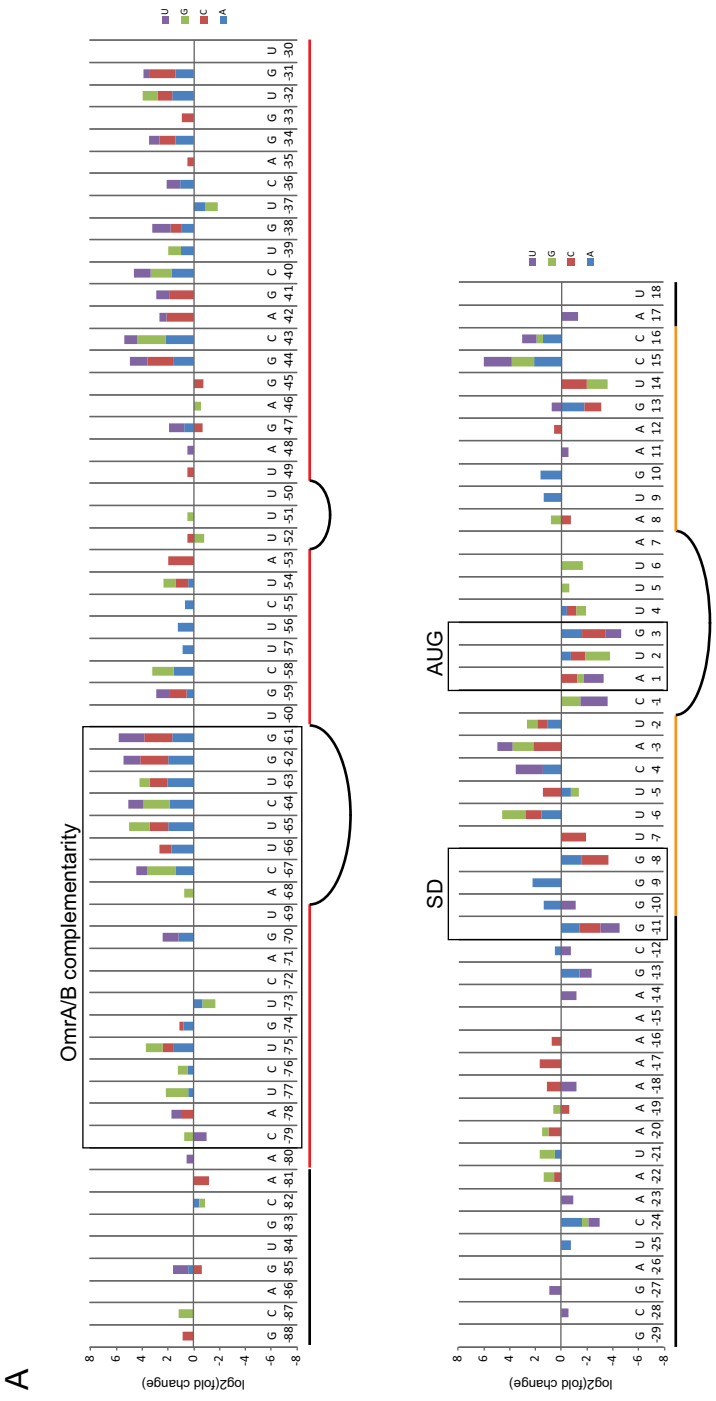
Mechanistic aspects of OmrA/OmrB-dependent regulation of CsgD

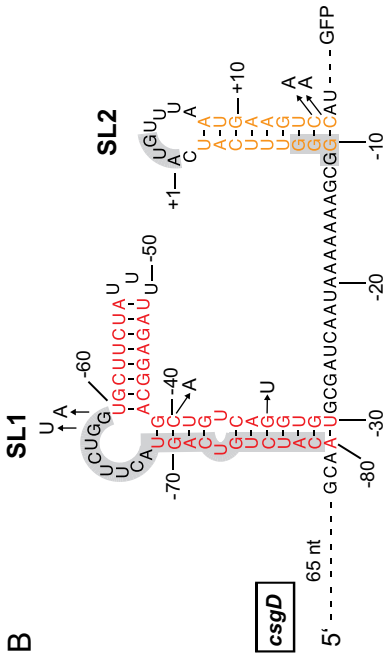
Another case of upstream binding is presented in Paper II. Results from computational predictions, *in vivo* mutational analyses, and *in vitro* structure probing strongly suggest that the sRNAs OmrA and OmrB bind to a region far upstream of the RBS of the *csgD* mRNA (Paper II). Despite the fact that the sRNA-binding site on *csgD* is located about 60 nucleotides upstream of the AUG start codon, binding primarily causes inhibition of translation initiation. This is based on several experimental results presented in Paper II. First, *in vitro* translation from an *in vitro* transcribed *csgD-3xflag* mRNA was inhibited when pre-incubated with either OmrA or OmrB, but not with mutant sRNAs carrying base-changes in the binding region. Inhibition could be restored by changing nucleotides in *csgD* to restore complementarity to the mutant sRNAs. These experiments were done in a translation system that was reconstituted from purified components (Puresystem), in which added RNA is stable throughout an experiment (unpublished observations and (Sharma *et al.*, 2007)) Second, formation of stable 30S-tRNA^{fMet} complexes on *csgD* mRNA was less effective when the mRNA had been preincubated with OmrA or OmrB. In this assay, the only added components were mRNAs, sRNAs, purified 30S subunits, and purified fMet-tRNA^{fMet}. These results indicate that translational inactivation of the *csgD* mRNA requires binding of OmrA or OmrB, but no other factors. Additionally, the results indicate that inhibition of translation as opposed to induced mRNA degradation is the primary outcome of OmrA/B binding. How then does binding of OmrA/B transform the *csgD* mRNA from a translationally active to a translationally inactive form? Secondary structure predictions and *in vitro* structure probing suggests that the *csgD* 5'UTR folds into two stem-loop structures, SL1 and SL2 (Paper II and Figure 6). The OmrA/B interaction site

starts in a large bulge in SL1 and continues through the lower part of the stem. SL2 sequesters the major part of the SD sequence, and the AUG start codon is located in the loop. The sequestration of the SD in the stem of SL2 appears to limit translation rates, since mutations that destabilize the stem lead to increased CsgD steady state levels (Paper II and Figure 6). The 148 nucleotide long *csgD* 5'UTR can be truncated from the 5'-end to nucleotide position -20 (with respect to the A in the start codon being +1) without losing translation activity, suggesting that there are no elements in the upstream sequence that are required for efficient translation and that could potentially could be blocked or altered upon OmrA/B binding. When OmrA or OmrB is bound to the *csgD* mRNA, the descending arm of SL1, which is complementary to the OmrA/B binding site, becomes single-stranded. However, the single-stranded character of this sequence does not seem to be responsible for inhibition of translation, since a 5'-truncation up to position -60, which leaves this sequence unpaired, is translated at the same rate as the wild type mRNA (Paper II).

A saturation mutagenesis approach to analyze *csgD* translation and regulation

Traditionally, mutational studies have successfully been used to map positions in RNA sequences that are important for e.g. secondary structure or gene expression. A disadvantage of these methods is that they are time-consuming, since every mutant has to be investigated individually. To minimize the time and cost of experimentation, and simultaneously increase the number of mutants that can be analyzed, we developed a strategy in which all possible mutations can be analyzed in parallel. Specifically, we used this strategy in order to map all positions in the *csgD* 5'UTR that are important for translational efficiency, or OmrA-dependent regulation of translation. On a *csgD*::*gfp* translational fusion plasmid, the DNA segment encoding the 5'UTR with all elements necessary for translation and regulation by OmrA/B were mutagenized by error-prone PCR, and mutant plasmids were transformed into strains that either harbored an OmrA overexpression plasmid or a control vector. Pools of cells with mutant plasmids were fractionated dependent on GFP fluorescence using fluorescence-activated cell sorting (FACS), and the representation of individual mutations in each fraction was identified by high-throughput DNA sequencing. The graph in Figure 6A shows fold change differences of single mutations in the *csgD* 5'UTR. Mutations in cells fractionated for high fluorescence in the presence of OmrA were compared to all mutations present in the original culture (without fractionation). The mutations that were enriched (positive value) should be those that either confer higher CsgD levels overall, or those that confer loss-of-regulation by OmrA.


A
OmrA/B complementarity
log2(fold change)
SD
AUG



C

csgD variant	sRNA plasmid	Relative expression*	Relative regulation+
wt	vector	1.00 ± 0.04	1.00 ± 0.04
	OmrA	0.37 ± 0.02	0.37 ± 0.02
	OmrB	0.30 ± 0.03	0.30 ± 0.03
(+15)C>A	vector	3.63 ± 0.24	1.00 ± 0.07
	OmrA	0.57 ± 0.02	0.16 ± 0.01
	OmrB	0.93 ± 0.01	0.26 ± 0.00
(+15)C>A (+16)C>A	vector	6.55 ± 0.28	1.00 ± 0.04
	OmrA	0.50 ± 0.02	0.08 ± 0.00
	OmrB	1.03 ± 0.04	0.16 ± 0.01
(-62)G>U	vector	0.49 ± 0.01	1.00 ± 0.03
	OmrA	0.58 ± 0.01	1.19 ± 0.03
	OmrB	0.52 ± 0.01	1.06 ± 0.03
(-61)G>A	vector	0.58 ± 0.02	1.00 ± 0.03
	OmrA	0.66 ± 0.01	1.12 ± 0.02
	OmrB	0.54 ± 0.01	0.92 ± 0.02
(-40)C>A (-34)G>U	vector	0.62 ± 0.02	1.00 ± 0.03
	OmrA	0.82 ± 0.01	1.13 ± 0.02
	OmrB	0.66 ± 0.02	1.05 ± 0.04
Δ(-60 to -42)	vector	1.91 ± 0.13	1.00 ± 0.07
	OmrA	1.00 ± 0.02	0.52 ± 0.01
	OmrB	0.79 ± 0.02	0.41 ± 0.01

Figure 6. Enrichment of randomly generated mutations in the *csgD*::*gfp* 5'UTR from cells expressing OmrA. (A) Fold change of number of mutations from cells with increased fluorescence versus number of all mutations present originally. Red and orange lines below the graph denote the SL1 and SL2 secondary structures, respectively (see B). (B) Secondary structure representation of the *csgD* 5'UTR. Base-paired nucleotides in SL1 and SL2 are in red and orange, respectively. The OmrA binding region, the SD sequence, and the AUG start codon are in grey. Arrows indicate mutations analyzed in C. (C) Relative values of fluorescence/OD600 from cells with *csgD*::*gfp* (and variants thereof), in combination with a vector control or overexpression of OmrA or OmrB. *All values were normalized to those from the wt strain with the control vector. + For each *csgD* variant, values were normalized to those from the control vector strain.

Mutations in SL1

Strikingly, all positions in the SL1 bulge which comprise part of the interaction site for OmrA show an enrichment of mutations that weaken the *csgD*-OmrA interaction (Figure 6, nts -68 to -61). In almost all of these positions, all three possible base-changes are present. Enrichment of mutations are also found in the remaining part of the interaction site, spanning the long stem in SL1 (nts -69 to -79), although not as pronounced as in the bulge. This may indicate that the bulge, within which the 5'-most nt's in the OmrA/B tail have matches, is the site at which formation of the *csgD*-OmrA duplex is initiated. Hence, mutations in this region are most detrimental for OmrA binding. Possibly, binding to only the bulge would be sufficient for regulation. To

confirm the sequencing results, some of the mutations were re-introduced in the wild type *csgD*::*gfp* fusion and assayed for fluorescence levels in the absence or presence of OmrA or OmrB. These experiments confirmed that mutations in the OmrA binding site (nts -62 and -61) confer loss of regulation (Figure 6). Enrichment of mutations in SL1 downstream of the OmrA interactions site correlates well with decreased stability of secondary structure. Mutations that confer increased stability are, on the other hand, underrepresented, e.g. mutations in U_{-73} and U_{-37}. Both of these positions show an underrepresentation of mutations to A or G, both of which would create a base-pair. This suggests that SL1 must be properly structured for OmrA to effectively bind to the SL1 bulge. The surrounding structure elements probably constrain the presentation of the bulge so it becomes "visible" and easily recognizable by OmrA. In line with this argument, deletion of the stem-loop encompassing nucleotides -60 to -42 does not impair expression or regulation by OmrA or OmrB (Figure 6C). In this mutant construct, the SL1 bulge becomes an apical loop in which the OmrA/OmrB interaction site is preserved. However, mutations in SL1 that break intramolecular base-pairs result in loss of regulation (Figure 6C, mutant (-40)C>A, (-34)G>U), perhaps because these mutations change the SL1 structure so that the OmrA interaction site becomes less available (as suggested by secondary structure predictions (not shown)).

Mutations between SL1 and SL2

The single-stranded region between SL1 and SL2 encompasses an A-rich stretch (nts -20 to -14). Since Hfq is known to bind to A- or AU-rich single-stranded sequences, this region may be an Hfq binding site. Interestingly, enriched mutations in this region are often base-changes from A to C, which would create a weaker Hfq binding site. Hfq is required for regulation of CsgD synthesis by both OmrA and OmrB (Paper II). The enriched mutations in the A-rich region may therefore tentatively be explained by an abolished Hfq binding site, which could result in less efficient regulation.

Mutations in SL2

The SL2 stem sequesters the SD sequence and is therefore a negative determinant for CsgD synthesis (Paper II). Interestingly, in cells with elevated fluorescence, several positions in the SL2 stem show enrichment for mutations that weaken the stem (Figure 6, nts -10, -9, -6 to -2, 7 to 9, 15, 16). Mutations in positions -10 and -9 may, in addition to opening up the SL2 stem, increase translation rates by creating better SD sequences (GGGG to GAGG or GGAG). Re-introduction of some of the mutations that destabilize SL2 confer increased CsgD-GFP levels without loss of regulation by OmrA (Figure 6C, mutants +15C>A, +16C>A).

In conclusion, this saturation mutagenesis strategy allows for the identification of all possible mutations important for post-transcriptional regulation,

as well as those that are phenotypically silent, and can be used for any RNA of interest. In addition to identification of single mutations as shown in Figure 6, many sequence reads contained two or more mutations. Thus, subsequent analysis will be conducted to identify compensatory mutations that revert the effect observed for each mutation alone, and hopefully result in a deeper understanding of the elusive regulatory mechanism that underlies OmrA/OmrB-dependent inhibition of CsgD synthesis. Additionally, it will be interesting to see whether inhibition of CsgD synthesis by other sRNAs, e.g. McaS and RprA (Jørgensen *et al.*, 2012; Mika *et al.*, 2012; Thomason *et al.*, 2012), rely on the same mechanism as that used by OmrA/OmrB.

sRNA-induced mRNA degradation

In addition to inhibition of translation, sRNA binding often increases the rate of mRNA degradation. Actively translated mRNAs are covered by elongating ribosomes protecting them from attack by RNases. When translation initiation is blocked by sRNAs, mRNAs become more sensitive to endonucleolytic RNase activity due to exposure of RNA segments in the absence of ribosome occupancy. In this model, increased mRNA degradation is a consequence of decreased translation rate upon sRNA binding.

Recent data however suggests that sRNAs may sometimes specifically induce mRNA cleavage. The sRNA RyhB was shown to, in addition to inhibiting translation by binding to the RBS of target mRNAs, induce RNase E cleavages at specific sites in the mRNAs, even in the absence of translation (Afonyushkin *et al.*, 2005; Prevost *et al.*, 2011). Several studies indicate that degradation of sRNAs and mRNAs is coupled, either by recruitment of RNase E or, if the sRNA-mRNA duplex is sufficiently long, through concomitant cleavage of the duplex by the double-strand-specific endoribonuclease RNase III (Masse *et al.*, 2003; Huntzinger *et al.*, 2005; Jørgensen *et al.*, 2012). One recent example of the latter is the indication that MicA is degraded by RNase III in a target-dependent manner (Viegas *et al.*, 2011). When the sRNA RyhB targets the *sodB* mRNA to inhibit translation and induce RNase E-dependent mRNA degradation by, RyhB is concomitantly degraded by RNase III (Afonyushkin *et al.*, 2005). A consequence of such a phenomenon is that as soon as sRNA expression is stopped, the intracellular sRNA pool will rapidly be depleted through targeting-dependent degradation. This would provide a mechanism for the cell to rapidly increase target gene expression.

There are also a few cases where sRNAs primarily or exclusively induce cleavage of target mRNAs, without effects on initiation of translation. This mechanism of regulation seems to be preferred when sRNAs bind within the coding sequence of the target mRNA (called “CDS targeting”) downstream of the “five codon window” (Bouvier *et al.*, 2008). The *Salmonella* sRNA MicC binds to the coding sequence of the *ompD* mRNA and recruits RNase

E to cleave the mRNA (Pfeiffer *et al.*, 2009). The recruitment of RNase E by the sRNA seems to be mediated through the Hfq protein. Further evidence that supports this model shows that sRNAs can form a complex with Hfq and RNase E *in vivo*, and that formation of this complex requires the C-terminal scaffold domain of RNase E (Morita *et al.*, 2005). The sRNA MicF post-transcriptionally regulates synthesis of the PhoE and CpxR proteins by binding to the coding sequence of the corresponding mRNA. Since these binding sites are downstream of the "five codon window", their regulation may be dependent on Hfq/RNase E (Paper III). However, this hypothesis needs to be tested, for example by analyzing MicF's effect on the stability of these mRNAs in RNase E mutant strains (Morita *et al.*, 2005).

sRNA-dependent activation of translation

Several sRNAs have been shown to activate translation of target mRNAs. The by far best studied case is the *rpoS* mRNA that codes for the stationary phase sigma factor σ^S, on which translation is activated by no less than three sRNAs (Majdalani *et al.*, 1998; Majdalani *et al.*, 2001; Mandin and Gottesman, 2010). The 5'-UTR of the *rpoS* mRNA can fold into a stable secondary structure in which the SD sequence becomes sequestered by intra-molecular base-pairing, leading to lower translation rates. Binding of either of the sRNAs RprA, DsrA, or ArcZ, changes the structure of the *rpoS* leader which renders the SD sequence accessible for translation, a mechanism referred to as anti-antisense (Lease *et al.*, 1998; Majdalani *et al.*, 1998).

Physiological roles of sRNAs

From the early discoveries of chromosomal sRNAs and antisense RNAs in plasmids, it was plausible to think that each sRNA had only one specific target mRNA. However, it is now clear that the majority of sRNAs in *E. coli* are regulators with multiple targets. It is therefore adequate to talk about sRNA regulons, in line with regulons of TFs. Additionally, the increasing number of examples of TFs that are regulated by sRNAs, suggests that the majority of the genes in *E. coli* may be controlled, directly or indirectly, by sRNAs.

Many sRNAs are induced by specific environmental signals to regulate the expression of genes that are physiologically connected. The sRNA RyhB is repressed by the TF Fur that responds to the intracellular iron concentration (Masse and Gottesman, 2002). RyhB is induced upon iron starvation and represses target mRNAs that encode iron-containing proteins, while simultaneously repressing the synthesis of its own repressor Fur (Masse *et al.*, 2007; Vecerek *et al.*, 2007). Expression of the sRNAs MicA and RybB is activated by the sigma factor σ^E in response to membrane stress. When in-

duced, RybB and MicA repress many mRNAs, several of which encode outer membrane proteins ((Johansen *et al.*, 2006; Papenfort *et al.*, 2006; Udekwu and Wagner, 2007; Gogol *et al.*, 2011) and Paper I). The sRNA FnrS is induced during anaerobic conditions to repress synthesis of enzymes needed in the presence of oxygen (Boysen *et al.*, 2010; Durand and Storz, 2010). In *S. aureus*, the quorum sensing activated sRNA RNAIII is unique in its ability to act as a post-transcriptional inhibitor, activator, and additionally as an mRNA encoding a hemolysin (Felden *et al.*, 2011).

Regulation of a sessile lifestyle

Motility and sessility are two opposite lifestyles for bacterial cells. When cells are motile, they produce flagella that enable them to swim and move directionally, which is the preferred state when nutrients are available. However, when nutrients in the surrounding environment become scarce, or when cells suffer from other types of stresses, they tend to aggregate in sessile communities known as biofilms. In order to aggregate, *E. coli* cells produce extracellular appendices such as pili, fimbriae, cellulose and curli, which make up a meshwork in which the cells are trapped. The switch between motility and sessility is governed by complex interconnected genetic reprogramming. As an activator of biofilm component expression, and a repressor of flagella synthesis, the CsgD protein is at the center of gene regulation when cells switch between the motile and sessile states. CsgD is expressed from the *csgDEFG* operon which is located next to the divergently transcribed *csgBAC* operon on the *E. coli* chromosome. Among the genes activated by CsgD are the sessility-associated operons *csgBAC*, *adrA*, and *csgDEFG* (Hammar *et al.*, 1995; Brombacher *et al.*, 2006; Ogasawara *et al.*, 2011), while the flagella operons *fliE* and *fliFGHIJK* are repressed by CsgD (Ogasawara *et al.*, 2011). The gene products of the two *csg* operons (csg stands for "curli-specific gene") are all required for formation of extracellular curli fibers (Hammar *et al.*, 1995). These proteinaceous fibers mediate contact between bacterial cells in biofilms and to abiotic surfaces, but also mediate contact with host cells through interaction with host surface proteins such as major histocompatibility complex 1 molecules (Hammar *et al.*, 1995; Olsén *et al.*, 1998; Wang *et al.*, 2006). Curli may have a role in pathogenesis, since curli-proficient strains are internalized into host cells to a higher extent than are curli-deficient strains (Gophna *et al.*, 2001). However, curli also seem to be important outside the host by aiding transmission (White *et al.*, 2008). The fiber proteins CsgA and CsgB are secreted through the CsgG pore, where CsgA proteins are nucleated into a fiber by CsgB (Robinson *et al.*, 2006; Nenninger *et al.*, 2011). CsgE and CsgF mediate specificity and localization of CsgA and CsgB, while CsgC is required for proper fiber formation (Gibson *et al.*, 2007; Nenninger *et al.*, 2009; Nenninger *et al.*, 2011). The AdrA protein is a diguanylate cyclase, an enzyme that converts GTP

into cyclic di-3',5'-guanylate (c-di-GMP), which in turn acts as a second messenger to activate cellulose production (Römling *et al.*, 2000; Zogaj *et al.*, 2001; Garcia *et al.*, 2004; Simm *et al.*, 2004). The CsgD-repressed operons *fliE* and *fliFGHIJK* both express proteins for flagellar assembly.

The *csgD* promoter is one of the most complex in *E. coli,* regulated by many TFs such as OmpR (Vidal *et al.*, 1998; Gerstel *et al.*, 2003; Gerstel *et al.*, 2006), IHF(Gerstel *et al.*, 2003; Gerstel *et al.*, 2006), Crp (Zheng *et al.*, 2004), MlrA (Brown *et al.*, 2001; Ogasawara *et al.*, 2010), RstA (Ogasawara *et al.*, 2007) and CsgD itself (Ogasawara *et al.*, 2011) (activation), and CpxR (Jubelin *et al.*, 2005) and H-NS (Gerstel *et al.*, 2003) (repression). On the post-transcriptional level, the two sRNAs OmrA and OmrB independently inhibit CsgD synthesis by binding to a complementary sequence in the *csgD* 5'UTR (Paper II). In addition, at least two other sRNAs (McaS and RprA) target the *csgD* mRNA (Jørgensen *et al.*, 2012; Mika *et al.*, 2012; Thomason *et al.*, 2012). In addition to inhibition of CsgD synthesis, McaS also activates synthesis of the FlhD protein of the $FhlD_2C_2$ flagellar master regulator, and is thus involved in the cross-regulation of motility and sessility (Thomason *et al.*, 2012). RprA is also an activator of σ^S synthesis, a sigma-factor required for expression of *csgD* (Hammar *et al.*, 1995; Majdalani *et al.*, 2001). Thus, synthesis of CsgD is controlled on several levels and through multiple signals. An explanation for the need of such complex control may lie in the irreversible nature of curli expression.

OmrA and OmrB and bistability of *csgD* expression (Paper IV)

As soon as CsgD is synthesized, cells start to produce curli fibers and become sessile. For the individual cell, as well as for the population, the decision to start curli production must be well-founded. In line with this is the observation that synthesis of CsgD, as well as curli production, is bistable. That is, individual cells in a population either express CsgD fully (ON state), or not at all (OFF state) (Grantcharova *et al.*, 2010). Bistable expression has also been demonstrated in production of type 1 pili, which – like curli – are extracellular appendix structures. Since pili aid adherence and infection of host cells, but at the same time stimulate an inflammatory response, there is a trade-off between increased virulence and immune system attraction. Balancing the trade-off by bistable expression of pili to create a heterogeneous population may be advantageous for successful infection (Wolf and Arkin, 2002). Interestingly, the fraction of cells that are ON in terms of curli production is strongly influenced by the presence of OmrA or OmrB, as seen by monitoring single-cell CsgB-YFP levels in wild type or isogenic sRNA deletion strains (Paper IV). Upon deletion of one or both sRNAs, the fraction of ON cells becomes much larger, suggesting a role for OmrA/B in controlling the stochastic nature of curli production. A possible reason for inhibition of CsgD synthesis by OmrA/B could therefore be to balance the fraction of the population that switches to the sessile state. Another possibility is that

OmrA/B function as dampers of stochastic transcriptional bursts when *csgD* expression should be OFF, preventing cells from inadvertently switching to the non-preferred sessile state. Related to the latter is the fact that OmrA/B inhibits synthesis of the TF and response regulator OmpR, which acts as a positive regulator of *csgD* transcription (Guillier and Gottesman, 2008). Simultaneous repression of both OmpR and CsgD may therefore impact on bistability of *csgD* expression. To define the contribution of OmpR versus OmrA/B to the bistable nature of *csgD* and curli expression, strains in which OmpR or CsgD synthesis is unresponsive to OmrA/B should be analyzed.

OmrA/OmrB promote motility by inhibiting FlgM synthesis (Paper IV)

The interconnected transcriptional regulation of the opposing motile and sessile states suggests that also post-transcriptional regulators such as sRNAs may control genes in both pathways. In fact, a recent study suggested that the sRNA McaS simultaneously acts as an inhibitor of CsgD and an activator of FhlD (Thomason *et al.*, 2012). Based on similar reasoning, we could show that OmrA/OmrB also regulate genes in the flagella pathway by inhibiting synthesis of the FlgM protein (see above). Decreased levels of FlgM would indirectly lead to increased levels of active σ^{28}, and eventually affect motility. To test this hypothesis, the motility of wild type versus *omrA/omrB* deletion strains was analyzed on soft agar plates. In line with the mutational data discussed above, the *omrA/omrB* deletion strain was less motile. This observation suggests that the proposed regulation of FlgM synthesis by OmrA/OmrB is physiologically relevant, and not only an overexpression effect.

Synthesis of flagella involves at least 17 operons. The promoters driving these operons are temporally regulated and designated class 1, 2, and 3 (reviewed in (Chilcott and Hughes, 2000)). The *flgM* mRNA is expressed from two promoters, an $FhlD_2C_2$-activated σ^{70}-dependent class 2 promoter that generates the *flgAMN* transcript, and a σ^{28}-dependent class 3 promoter which generates the *flgMN* transcript. Upon completion of the hook-basal body (a sub-step in flagellar assembly), the FlgM protein is actively secreted, which in turn activates σ^{28} for transcription of the class 3 operons, including *flgMN*. Producing FlgM from a class 3 promoter only to be secreted is believed to regulate the length of the final flagellum. The intriguing question of whether OmrA/B can regulate synthesis from both the *flgAMN* and *flgMN* transcripts, and how this influences the temporal expression of flagellar synthesis, is an important aspect that will be addressed in future studies.

Several sRNAs converge on the same targets

Many sRNAs are known to regulate multiple targets. However, there are also mRNAs which are regulated by several sRNAs. The classical example is the regulation of RpoS synthesis. The 5'UTR of the *rpoS* mRNA folds into a

secondary structure element in which the SD sequence and AUG start codon are sequestered by intramolecular base-pairing (Majdalani *et al.*, 1998).

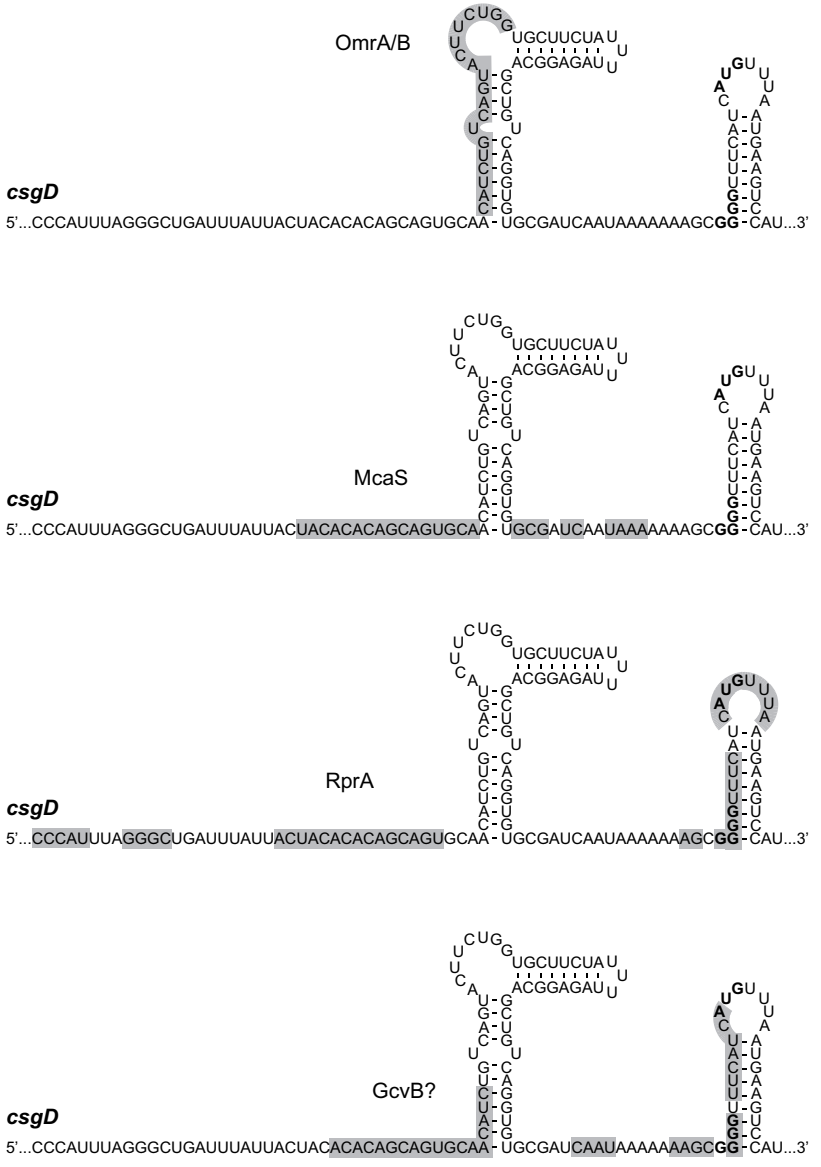


Figure 7. Schematic figure of the *csgD* 5'UTR with proposed binding sites of the five sRNAs OmrA, OmrB, McaS, RprA, and GcvB highlighted in grey. The SD sequence and the AUG start codon are in bold. The question mark after GcvB is inserted to emphasize that this interaction is not yet supported by experimental evidence.

Three different sRNAs, RprA, DsrA, and ArcZ, use a so called anti-antisense mechanism to disrupt the intramolecular structure in the *rpoS* mRNA and thereby activate translation (Majdalani *et al.*, 1998; Majdalani *et al.*, 2001; Mandin and Gottesman, 2010). In addition, a fourth sRNA, OxyS, has been shown to inhibit translation by binding to *rpoS* (Zhang *et al.*, 1998), although the regulatory mechanism is not yet fully understood. It was recently proposed that three sRNA, in addition to OmrA/B, inhibit translation of the *csgD* mRNA by binding to complementary sequences in the *csgD* 5'UTR (Figure 7, (Jørgensen *et al.*, 2012; Mika *et al.*, 2012; Thomason *et al.*, 2012)). RprA and McaS both target the same region, which is located immediately upstream of the OmrA/B binding site, and RprA may in addition base-pair to a downstream sequence that includes the SD (Figure 7, (Jørgensen *et al.*, 2012; Mika *et al.*, 2012)). Whether the third sRNA, GcvB, is a direct regulator of CsgD remains to be established (Jørgensen *et al.*, 2012). Why do several sRNAs converge on the same target with – as it seems – the same outcome, translational activation (on *rpoS*), or inhibition (on *csgD*)? Both σ^S and CsgD are central regulatory proteins that have a great impact on cell physiology or behavior. Both gene products are also under extensive transcriptional regulation. It therefore seems plausible that the control of these genes need to be responsive to many different environmental signals, so that tuning of their expression is optimized. Future work will show whether the levels of other central regulatory proteins, e.g. the flagellar master regulator $FlhD_2C_2$, are also controlled by multiple sRNAs.

MicF regulates Lrp in response to nutritional status

Recent data indicate that the expression of up to ten percent of the ORFs on the *E. coli* genome are controlled by the TF Lrp, strengthening the view of Lrp as being one of the most important global TFs in this bacterium (Cho *et al.*, 2008). Many Lrp-controlled genes are involved in transport of small molecules or amino acid biosynthesis (Calvo and Matthews, 1994; Newman and Lin, 1995). Among the Lrp-regulated genes are *micF* and *lrp* (autoregulation). The data presented in Paper III shows that MicF inhibits synthesis of Lrp by binding to a complementary sequence in the *lrp* mRNA. Hence, Lrp and MicF make up a double negative feedback loop through the transcriptional regulation by Lrp on *micF* expression, and the post-transcriptional regulation of MicF on Lrp protein synthesis. Previous studies suggested that Lrp levels are regulated in response to nutrient availability, leading to higher Lrp levels in nutrient-poor than in nutrient-rich environments (Lin *et al.*, 1992). However, the repression of Lrp levels in nutrient-rich environments has been elusive. A comparison of Lrp levels during growth in high and low nutrient media, monitored as fluorescence of a plasmid-encoded Lrp-GFP fusion protein, is presented in Paper III. Importantly, expression of the fusion (which is controlled by the native *lrp* promoter) from a plasmid does not

perturb the chromosomally encoded regulatory system. Hence, the regulatory impact on Lrp levels from chromosomally encoded MicF and Lrp could be studied. Since Lrp-GFP levels were upregulated in rich growth medium in a strain deleted for *micF*, the previously observed lower Lrp levels in nutrient-rich environments may be explained by inhibition of Lrp synthesis by MicF (Paper III). Lrp autoregulation has a minor effect in nutrient-rich environments but is important in nutrient-poor environments ((Lin *et al.*, 1992; Wang *et al.*, 1994) and Paper III). This fits well with the observed steady state levels of MicF, which are high in nutrient-rich media and low in nutrient-poor media ((Coyer *et al.*, 1990) and Paper III). In line with the above mentioned results, the model described in Paper III suggests that Lrp levels are regulated in a nutrient-dependent fashion, either through post-transcriptional regulation by MicF (nutrient-rich environments) or by transcriptional autoregulation by Lrp (nutrient-poor environments).

Concluding remarks

The results presented in this thesis deal with mechanisms and physiological impact of sRNAs-dependent regulation of gene expression. By binding to complementary sequences in the 5'-part of target mRNAs, sRNAs affect translation and mRNA stability. The sRNA MicA inhibits synthesis of the outer membrane protein OmpA by base-pairing with the SD sequence of the *ompA* mRNA (Paper I). The same mechanism governs inhibition of Lrp synthesis by MicF (Paper III), and inhibition of FlgM synthesis by OmrA and OmrB (Paper IV). A somewhat different, and presently elusive, mechanism is responsible for the translational inhibition that OmrA/OmrB exert on the *csgD* mRNA (Paper III). In this case, regulation does not rely on sRNA-binding at the RBS, but to an upstream sequence.

sRNA-dependent regulation often results in physiological changes in cells. MicF regulates the steady-state levels of Lrp as a consequence of nutrient availability in the environment surrounding the bacteria, and reciprocal expression of these two regulators is mediated through mixed double negative feedback (Paper III). OmrA and OmrB attenuate production of the biofilm components curli and cellulose by inhibition of CsgD synthesis (Paper II), and activate motility by inhibiting synthesis of FlgM (Paper IV), suggesting that motility and sessility are inversely regulated also at the post-transcriptional level. Finally, phenotypic variation in clonal populations due to bistability in *csgD* expression is strongly dependent on OmrA and OmrB (PaperIV), suggesting that sRNA-dependent post-transcriptional regulation may be a common way for controlling bistability in gene expression.

Future perspectives

Although some questions regarding sRNA-dependent regulation have been answered in the course of this thesis work, even more new questions have emerged. I will in the following section discuss some of the aspects that I am curious to learn more about in the future.

Mechanism of inhibition of CsgD synthesis

Saturation mutagenesis of the *csgD* 5'UTR revealed single mutations that affect CsgD steady state levels in an OmrA-dependent or independent fashion. However, at this stage, the link between OmrA/OmrB-binding and translation inhibition remains unclear. Several hypotheses have been tested and ruled out:

- *Transcription attenuation*. Single-round transcription in the presence or absence of OmrA/OmrB did not reveal induced attenuation - a mechanism we initially favored. *In vitro* translation from an mRNA template is inhibited in the presence of OmrA/OmrB (and in the absence of transcription).
- *Induced mRNA degradation*. Regulation is unaffected in strains deleted for RNase III or RNase E. Regulation of *in vitro* translation in the absence of mRNA degradation.
- *Inhibition by Hfq*. In this model, sRNAs would recruit Hfq to the *csgD* mRNA and Hfq would act as an inhibitor of translation. Regulation in *in vitro* translation assays is effective in the absence of Hfq.
- *Masking of ribosome standby site*. Truncating the *csgD* 5'UTR up to position -20 does not affect the translation rate, suggesting that no sequence or structure elements upstream of position -20 are needed for efficient translation.
- *Inhibition of leader peptide synthesis*. Two short ORFs are present in the *csgD* 5'UTR. However, translational fusions to either of these ORFs show that they are inactive, irrespective of the presence or absence of OmrA or OmrB.

Among the remaining plausible mechanisms by which CsgD synthesis becomes inhibited upon OmrA/OmrB-binding is induction of secondary struc-

ture changes. One could imagine that the 3'-half of SL1 that is rendered single-stranded upon OmrA/B binding forms an alternative structure that promotes inhibition of translation (Figure 6 in Paper II). Although this has so far not been detected in *in vitro* structure probing experiments, these techniques may not in all cases reveal subtle structure changes. Another possibility is that the sRNAs interact with additional sequences in *csgD*, e.g. closer to the RBS, and thereby directly interfere with translation initiation. However, this model is not supported by *in vitro* structure probing and biocomputational predictions. Further and deeper analysis of the data set generated by saturation mutagenesis might result in new hypotheses that can be tested, or point to one of the models mentioned above. Identification of pairs of mutations that compensate for the altered phenotype observed for each single mutation alone may be the most promising route to obtain a deeper mechanistic understanding of the OmrA/OmrB-dependent regulation of CsgD synthesis.

Stochastic regulation of motility versus sessility

The results presented in Paper IV indicate that OmrA and OmrB participate in regulation of *both* motility and sessility. These behaviors are most likely mutually exclusive for a single cell, but not necessarily for the population. In a comparative study of different *E. coli* strains, the most motile strains were also best at forming biofilms (Wood *et al.*, 2006). Since production of curli fibers is bistable ((Grantcharova *et al.*, 2010) and Paper IV), it is tempting to speculate whether motility is also regulated in a bistable fashion. One way of testing this hypothesis would be to engineer *E. coli* so that single cells express two different (and wave length-wise distinguishable) fluorescent proteins under the control of promoters specific for flagella or curli production (e.g. *fliC* and *csgB*). This strategy would reveal whether single cells show a mutually exclusive expression of the two pathways. In addition, the influence of any sRNAs known to affect these pathways can be investigated using such an experimental set-up.

Bistability is known to be caused by stochastic fluctuations in transcription. When the expression level of a positive regulator reaches a certain threshold, downstream genes are activated. Positive autoregulation, if applicable, of the regulator further enhances downstream activation. This is exemplified by activation of the competent state in *Bacillus subtilis* by stochastic fluctuations in ComK concentration (reviewed in (Dubnau and Losick, 2006; Veening *et al.*, 2008)). Bistability in the curli pathway has been shown by analysis of CsgD and CsgB protein levels ((Grantcharova *et al.*, 2010) and Paper IV), and OmrA and OmrB were shown to impact on the bistability in CsgB levels (Paper IV). Hence, the cause of bistability in this pathway lies either on the translational level of *csgD* expression, or at a level that controls *csgD* translation. It will thus be interesting to analyze whether transcription

of *csgD* is also bistable, and if it is, whether one or several regulators of *csgD* transcription also show bistable expression. One possibility is that OmrA and OmrB (and possibly other sRNAs such as McaS and RprA) are the main cause of bistability in *csgD* expression. It would in that case be the first example of bistability generated by sRNA-dependent regulation.

Intracellular localization of sRNAs

A more general question deals with how RNA localize within the cell. The traditional view of the bacterial cell as a "bag of enzymes" with freely diffusing macromolecules has recently been challenged. Improved methodology has allowed for spatial detection of intracellular RNA, suggesting both restricted mRNA movement from the site where transcription occurs, as well as migration of mRNAs for sub-cellular localization (Montero Llopis *et al.*, 2010; Nevo-Dinur *et al.*, 2011). Regarding sRNA-dependent regulation, restricted movement of mRNAs implies that sRNAs either are produced in the vicinity of their targets, or move to the site of transcription of target mRNAs. The latter view is supported by the fact that plasmid-encoded sRNAs can regulate mRNAs transcribed from both plasmids and chromosomal loci. One study suggested that sRNA-dependent repression of membrane protein synthesis requires target mRNA localization to the inner membrane (Kawamoto *et al.*, 2005), and membrane localization has also been observed for Hfq (Diestra *et al.*, 2009). Further studies on intracellular localization and movement of sRNAs, mRNAs, and Hfq, may reveal the spatial and temporal dynamics of sRNA-dependent regulation.

Svensk sammanfattning

RNA-baserad genreglering

Celler svarar på förändringar i den omgivande miljön genom att modifiera genuttryck, dvs geners aktivitet. Genreglering är ett samlingsnamn för de processer som styr geners aktivitet. Traditionellt har man trott att proteiner, så kallade transkriptionsfaktorer, är de enda molekyler som är ansvariga för genreglering. Transkriptionsfaktorer binder till DNA-sekvenser i närheten av gener för att stimulera eller motverka att genen transkriberas till RNA. Denna form av kontroll kallas för transkriptionell reglering. RNA-molekyler kan delas upp i två huvudgrupper. De som kodar för proteiner, dvs kan translateras, kallas för mRNA. RNA-molekyler som saknar denna egenskap kallas för icke-kodande RNA. Med posttranskriptionell reglering menas både de processer som påverkar i vilken grad mRNA translateras till protein, och de processer som påverkar hur mRNA bryts ner. Translation av mRNA utförs av ribosomen, ett stort komplex bestående av både protein och RNA. Nedbrytning av RNA utförs av enzymer som kallas för RNaser. Intensiv forskning under de senaste tio åren har visat att de flesta organismer använder korta icke-kodande RNA-molekyler för posttranskriptionell genreglering. Detta fenomen upptäcktes först i bakterier, och man trodde då att det rörde sig om enstaka fall. När man sedan systematiskt sökte efter okända gener i tarmbakterien *Escherichia coli* kunde man påvisa många som uttrycker korta RNA-molekyler. Genom min och andras forskning vet vi nu att dessa RNA-molekyler (ofta kallade sRNA) fungerar som kontrollelement för posttranskriptionell genreglering. I likhet med transkriptionsfaktorer är sRNA integrerade delar i regulatoriska nätverk och kan påverka uttrycket från flera gener simultant. Genreglering utfört av sRNA har visat sig påverkar många bakteriella egenskaper, t ex motilitet, näringsrespons, biofilmsbildning, kommunikation och virulens.

I min forskning har jag försökt att förstå de mekanismer som ligger till grund för den reglering som utförs av sRNA, och hur detta påverkar bakteriecellens fysiologi. Jag har också analyserat hur sRNA verkar i samarbete med andra regulatoriska molekyler, t ex transkriptionsfaktorer, för att förstå funktionen av så kallade regulatoriska nätverksmotiv. Den bakterie som jag har studerat, *Escherichia coli*, utgör vanligtvis en viktig del i friska människors tarmflora, men kan i vissa fall även orsaka sjukdom. I *E. coli* verkar sRNA oftast genom att baspara (binda) till specifika mål-mRNA. Detta för-

utsätter att sRNA och mål-mRNA har komplementära nukleotidsekvenser mellan vilka basparning kan ske. I många fall basparar sRNA till ribosombindningsstället, den del av mål-mRNAt där translation påbörjas. Detta leder till att ribosomers åtkomst till mRNAt blockeras, vilket i sin tur resulterar i att proteinsyntesen förhindras. Aktivt translaterade mRNA är täckta av ribosomer och därmed skyddade från att attackeras av RNaser. Detta är dock inte fallet för mRNA-molekyler till vilka sRNA har basparat, eftersom basparningen leder till att mRNA blir fritt från ribosomer. Basparning mellan sRNA och mRNA leder därför ofta till ökad nedbrytning av mRNA.

Identifiering av mål-mRNA

Genom att använda datorbaserade beräkningsalgoritmer har jag och mina forskarkollegor gjort förutsägelser om vilka mRNA-molekyler som teoretiskt kan baspara till specifika sRNA-molekyler. Experimentellt har jag sedan analyserat vilka av dessa mRNA-molekylers aktivitet som verkligen är styrd av sRNA i bakteriecellen. Genom att använda ett reportergensystem i vilket proteinproduktion från mål-mRNA enkelt kan mätas, har jag identifierat mRNA vars uttryck är kontrollerat av sRNA. Genom att introducera mutationer i basparningsställen i både sRNA och mål-mRNA, har jag kunnat visa att den observerade regleringen är helt beroende av basparning.

Studie av sRNA-molekylers regleringsmekanismer

För att på molekylär nivå studera basparning mellan sRNA och mRNA har jag använt en teknik där RNA klyvs med antingen enzymer eller kemikalier. Eftersom klyvningseffektiviteten beror på om RNA-substratet är basparat eller ej, ger denna typ av analys information om vilka delar av en RNA-molekyl som deltar i intra- eller intermolekylära interaktioner. På detta sätt har jag kunnat påvisa basparning av sRNA till ribosombindningsställen på flera mål-mRNA, och även kunnat visa vilka delar av sRNA som deltar i basparning. För att direkt visa att basparning till ribosombindningsställen minskar ribosomers möjlighet att associera till mRNA, har jag använt en variant av en teknik som kallas ”primer extension”. Jag har även kunnat påvisa en tidigare okänd styrmekanism. I detta fall basparar inte sRNA inom ribosombindningsstället, men bindningen leder oavsett till inhiberad translationsinitiering (arbete II). Även om detaljerna i denna mekanism inte är helt klarlagda visar resultat från en mutationsstudie att flera områden av mål-mRNAt, utöver sRNA-basparningsstället, är viktiga för att reglering ska kunna ske.

sRNA kontrollerar bakteriell motilitet och biofilmsbildning

Två av de mål-mRNA som jag har identifierat kodar för proteiner som kontrollerar biofilmsbildning (proteinet CsgD) och motilitet (proteinet FlgM). När det finns mycket näring i bakteriens levnadsmiljö producerar de flageller, extracellulära roterande organeller med vilka de kan förflytta sig. FlgM är en negativ regulator av flagellproduktion. När bakterier å andra sidan befinner sig i näringsfattiga miljöer eller utsätts för olika typer av stress, till exempel i närvaro av antibiotika, hindras flagellproduktionen och bakterierna aggregerar i en biofilm. Biofilmer är multicellulära bakteriesamhällen uppbyggda av en extracellulär matrix bestående av polysackarider och proteinpolymerer. Matrixen gör att bakterier kan fästa vid varandra, värdceller, eller abiotiska ytor. En proteinpolymer i biofilmer kallas curli och byggs upp av proteinerna CsgA och CsgB. Produktionen av dessa aktiveras i sin tur av transkriptionsfaktorn CsgD. Jag har kunnat visa att två sRNA som kallas OmrA och OmrB *aktiverar* produktion av flageller genom att inhibera syntes av FlgM, och *förhindrar* produktion av curli genom att inhibera syntes av CsgD. På detta sätt styr OmrA och OmrB om bakterier ska vara motila eller bilda biofilm. Eftersom biofilm är ett stort problem både inom sjukvården och livsmedelsindustrin är förståelsen för hur biofilm kontrolleras av oerhört stor betydelse, såväl ekonomiskt som humanitärt.

sRNA reglerar fenotypisk variation i klonala bakteriepopulationer

Populationer av bakterier kan, trots att de är genetiskt identiska och lever i samma miljö, dela upp sig i fenotypiskt distinkta subpopulationer. Till exempel har man i studier av biofilmer observerat att vissa individuella bakterier kan bryta sig loss, producera flageller och därmed bli motila. Även om detta beteende kan tyckas vara ofördelaktigt för den individuella bakterien, kan det för hela populationen vara en fördel genom att det ökar möjligheterna att möta snabba miljöförändringar. Jag har kunnat visa att denna typ av fenotypisk variation kan vara styrd av sRNA, och att avsaknad av sRNA-gener leder till minskad fenotypisk variation (arbete IV). För att studera detta har jag använt fluorescens-aktiverad cellsortering, en teknik där genuttryck från enskilda celler kan mätas. I en population av vildtypsceller kunde jag se

att curliproteinet CsgB visade ett bistabilt uttryck; i enskilda celler var genuttrycket antingen fullt aktiverat (PÅ) eller inaktiverat (AV), medan få celler hade ett intermediärt genuttryck. I en population av celler som saknade OmrA och OmrB var däremot genuttrycket mer homogent, och de flesta celler hade ett fullt aktiverat genuttryck av CsgB. Till skillnad från tidigare förklaringar till fenotypisk variation, t ex stort brus vid lågt genuttryck, visar mina resultat att stokastiska skillnader i genuttryck skulle kunna härledas till posttranskriptionell reglering.

Återkoppling mellan sRNA och ett regulatoriskt protein involverat i näringsrespons

Det första sRNA som påvisades i *E. coli* kallas MicF. Trots att detta sRNA har varit känt väldigt länge kände man hittills bara till ett enda specifikt mRNA som kontrolleras av MicF. Arbete III beskriver för första gången att MicF reglerar flera andra mRNA, bland annat mRNAt Lrp. Proteinet Lrp som uttrycks från detta mRNA är en globalt verkande transkriptionsfaktor som påverkar aktiviteten av hela tio procent av alla gener i *E. coli*. Genom att analysera aktiviteten av Lrp-kontrollerade gener har jag kunnat visa att MicF indirekt får stor inverkan på cellens fysiologi genom att reglera Lrp. Då Lrp styr aktiviteten av genen *micF*, och MicF sRNA styr produktionen av proteinet Lrp, kontrolleras dessa båda styrmolekyler genom dubbel negativ återkoppling (arbete III). Denna återkoppling leder till att MicF och Lrp får motsatta uttrycksmönster. Uttryck av Lrp kontrolleras även genom negativ autoreglering, dvs proteinet Lrp utför negativ kontroll på uttryck av genen *lrp*. För att kunna studera både autoreglering och MicF-beroende reglering av Lrp, har jag låtit uttrycket av ett fluorescerande protein styras av de genetiska kontrollelement som normalt styr uttyck av Lrp. På detta sätt har jag kunnat visa att MicF har stor inverkan på uttrycksnivåer av Lrp när bakterier är i näringsrika miljöer, medan autoreglering av Lrp dominerar i näringsfattiga miljöer. Detta stämmer väl överens med mina resultat som visar att uttrycket av MicF är högre i näringsrika miljöer än i näringsfattiga miljöer.

Sammanfattning

De senaste femton åren har inneburit en revolution för förståelsen av kontroll av genuttryck. I de allra flesta organismer, från bakterier till djur, har RNA visat sig spela en viktig roll i detta. Jag har genom att studera *E. coli* försökt

förstå de mekanismer som ligger till grund för sRNA-baserad genreglering i bakterier. Jag har också försökt belysa den betydelse sRNA har för fysiologiska responser. Eftersom bakterier både är viktiga för människans normala fysiologi, och kan orsaka sjukdom, är förståelsen för den genetiska kontrollen av bakteriers beteende av stor betydelse för samhället. Jag hoppas att jag med denna avhandling har ökat förståelsen av bakteriell genreglering, och att denna kunskap kommer att vara till nytta för framtida forskning.

Acknowledgements

First and foremost I want to express my sincere gratitude to everyone that has contributed to this work. In particular I would like to thank the following people:

My supervisor **Gerhart**, when I started in your group I hardly knew anything about RNA, gene regulation or bacterial physiology. What I knew was that you were (and still are) a very enthusiastic and inspiring scientist and teacher. My years as a PhD student have been extremely educative and, much thanks to you, I have cultivated an everlasting passion for science in general, and bacterial research in particular. However, above all, the time in your group has been great fun! I want to thank you for always putting interest in, and encouraging all my ideas, for letting me work independently and putting trust in my capability, for giving me the opportunity to go to many meetings, and for hourly discussions about science, music, and life in general. I sincerely hope that I will have the possibility to do research with you also in the future.

Johan R (Movie Star), we have had (and still have) such a great collaboration. Your understanding of experimental work has been extremely important and I am looking forward to collaborating more with you in the future. I am also looking forward to more CL-nights at the pub!

Cia (Duracell-woman), so many good memories from all the meetings we went to together, and so many interesting and fruitful scientific discussions. Those six months of intense work on the MicF-story made us such a great team.

Klas, thanks for opening my eyes to the world of bacterial physiology, and for teaching me how to do *in vivo* experiments. Good luck in Stockholm!

Fabien ("it's all about loading"), my first teacher in the lab, thanks for sharing your excellent skills in doing *in vitro* experiments.

Maaike, thanks for good collaborations, I'm looking forward for more of that in the future.

Nina and **Ute**, thank you for introducing me to the mysteries of the curliated world.

Jörg, I had a great time in your lab in Berlin, I am looking forward to joining your team in Würzburg.

I also would like to thank ALL the post and present colleagues and friends at BMC. In particular I would like to thank:

The GW group: **Mirthe**, it has been great to share office with you the last year. Good luck with your PhD, I am sure you will do great. **Nadja**, we always had such good discussions, I am looking forward to more of that (colleagues soon?), **Aurelie**, **Amanda**, **Magnus Lindell** and **Helene**, thanks for all the fun in the lab and elsewhere.

The Dicty Crew: My second supervisor **Fredrik S**, thank you for interesting discussions, friendship, and for reading and correcting my thesis at the last minute, and welcome to Micro! **Lotta** (Dancing Queen), so many memorable moments discussing life (present and future), the Berlin trip was good fun. Good luck in Australia and I hope we can be colleagues again in the future! **Andrea**, thanks for helpful career advice and for always being encouraging.

The CRISPR Team: **Magnus Lundgren**, thanks for all scientific discussions and for advice about the scientific career. Good luck with your research and the new family member! **Lina** and **Devashish**, good luck with your projects.

The Giardians: **Staffan**, thanks for fun discussions, both scientific and other, I wish I had your broad knowledge and deep understanding of molecular biology. **Ankarklev** (World Master of Acknowledgements), I have really enjoyed being your colleague and friend, thanks for all discussions about science, music and life. You will do great in the future. **Jon** (the living encyclopedia), thanks for good times in the lab and during courses, and for teaching me so much, is there anything you don't know? The snake at EBC was hilarious. **Britta** (Queen of Swiss Chocolate) and **Elin** (horse-power), so much fun being around you two, you always make me smile. **Karin T** (The Manager), I miss you in the lab, and I look up to your ability to get things done. **Mattias**, **Emma** and all other post and present members of the Svärd group for making Micro such a social and fun working place.

Marie and **Disa**, we have had so much fun together during these years, good luck with your post-doc projects. **Shi-Ying**, we miss you in the lab, soon we are neighbors! **Fredrik P**, thanks for all help in the lab and fun times. **Ulrika**, thanks for all the help during my rookie years, and for all good moments. **Karin C**, thank you for very valuable comments on my Swedish summary.

Magnus J, I really enjoy our chats about science and life during lunches, family gatherings and while jogging along the river. Good luck in the land opportunities!

I also want to thank my friends and family for support and encouragement:

Magnus, Anna, **Jonas, Åsa**, **Erik, Maria**, **Niklas, Minna**, **Fredrik K**, **Matti**. I am glad to have such good friends.

David (Dr Bas), **Oscar** (Osman), **Petter** (Svetlana), **Johannes** (Vladimir): There is nothing like being on the stage with you, sweating, smiling, and receiving all the positive energy from the audience. Thank you for the music...

A special thanks to **Ulla** and **Anders** for being fantastic parents in law. You are such warm-hearted and generous persons. I will really miss being in "Stugan" this Easter.

My big family: **Mamma**, **Pappa**, **Rasmus**, **Kajsa**, **Stina**, **Pelle**, (Stora)**Axel**, **Mattias**, **Sara**, and all the **children**: thank you for all the support and love. I am looking forward to being together with all of you in the new winter house!

Axel and **Ellen**, life has forever changed since you came to us. You are the most wonderful children and it is a fantastic privilege to being part of your lives. I love you so much.

And finally, **Eva**, this book had never been finished without you supporting me. The last few months you have often sacrificed your own wishes so I could work on this book. For this I am sincerely grateful. I LOVE YOU.

References

Afonyushkin, T., Moll, I., Bläsi, U. and Kaberdin, V. R. (2003) Temperature-dependent stability and translation of Escherichia coli ompA mRNA. *Biochem Biophys Res Commun* **311**: 604-609.

Afonyushkin, T., Vecerek, B., Moll, I., Bläsi, U. and Kaberdin, V. R. (2005) Both RNase E and RNase III control the stability of sodB mRNA upon translational inhibition by the small regulatory RNA RyhB. *Nucleic Acids Res* **33**: 1678-1689.

Aiba, H., Matsuyama, S., Mizuno, T. and Mizushima, S. (1987) Function of micF as an antisense RNA in osmoregulatory expression of the ompF gene in Escherichia coli. *J Bacteriol* **169**: 3007-3012.

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J Bacteriol* **181**: 6361-6370.

Alon, U. (2007) Network motifs: theory and experimental approaches. *Nat Rev Genet* **8**: 450-461.

Altuvia, S., Kornitzer, D., Teff, D. and Oppenheim, A. B. (1989) Alternative mRNA structures of the cIII gene of bacteriophage lambda determine the rate of its translation initiation. *J Mol Biol* **210**: 265-280.

Altuvia, S. and Wagner, E. G. H. (2000) Switching on and off with RNA. *Proc Natl Acad Sci U S A* **97**: 9824-9826.

Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L. and Storz, G. (1997) A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* **90**: 43-53.

Ambros, V. (2011) MicroRNAs and developmental timing. *Curr Opin Genet Dev* **21**: 511-517.

Argaman, L., Hershberg, R., Vogel, J., Bejerano, G., Wagner, E. G. H., Margalit, H. and Altuvia, S. (2001) Novel small RNA-encoding genes in the intergenic regions of Escherichia coli. *Curr Biol* **11**: 941-950.

Arnold, T. E., Yu, J. and Belasco, J. G. (1998) mRNA stabilization by the ompA 5' untranslated region: two protective elements hinder distinct pathways for mRNA degradation. *RNA* **4**: 319-330.

Avery, O. T., Macleod, C. M. and McCarty, M. (1944) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J Exp Med* **79**: 137-158.

Babitzke, P. and Romeo, T. (2007) CsrB sRNA family: sequestration of RNA-binding regulatory proteins. *Curr Opin Microbiol* **10**: 156-163.
Beisel, C. L. and 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.
Benito, Y., Kolb, F. A., Romby, P., Lina, G., Etienne, J. and Vandenesch, F. (2000) Probing the structure of RNAIII, the Staphylococcus aureus agr regulatory RNA, and identification of the RNA domain involved in repression of protein A expression. *RNA* **6**: 668-679.
Bossi, L. and Figueroa-Bossi, N. (2007) A small RNA downregulates LamB maltoporin in Salmonella. *Mol Microbiol* **65**: 799-810.
Bouvier, M., Sharma, C. M., Mika, F., Nierhaus, K. H. and Vogel, J. (2008) Small RNA binding to 5' mRNA coding region inhibits translational initiation. *Mol Cell* **32**: 827-837.
Boysen, A., Møller-Jensen, J., Kallipolitis, B., Valentin-Hansen, P. and Overgaard, M. (2010) Translational regulation of gene expression by an anaerobically induced small non-coding RNA in Escherichia coli. *J Biol Chem* **285**: 10690-10702.
Brombacher, E., Baratto, A., Dorel, C. and Landini, P. (2006) Gene expression regulation by the Curli activator CsgD protein: modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion. *J Bacteriol* **188**: 2027-2037.
Brown, P. K., Dozois, C. M., Nickerson, C. A., Zuppardo, A., Terlonge, J. and Curtiss, R., 3rd (2001) MlrA, a novel regulator of curli (AgF) and extracellular matrix synthesis by Escherichia coli and Salmonella enterica serovar Typhimurium. *Mol Microbiol* **41**: 349-363.
Busch, A., Richter, A. S. and Backofen, R. (2008) IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. *Bioinformatics* **24**: 2849-2856.
Calvo, J. M. and Matthews, R. G. (1994) The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. *Microbiol Rev* **58**: 466-490.
Cao, Y., Zhao, Y., Cha, L., Ying, X., Wang, L., Shao, N. and Li, W. (2009) sRNATarget: a web server for prediction of bacterial sRNA targets. *Bioinformation* **3**: 364-366.
Carmichael, G. G., Weber, K., Niveleau, A. and Wahba, A. J. (1975) The host factor required for RNA phage Qbeta RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate-cellulose chromatography. *J Biol Chem* **250**: 3607-3612.
Carpousis, A. J. (2007) The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E. *Annu Rev Microbiol* **61**: 71-87.
Celesnik, H., Deana, A. and Belasco, J. G. (2007) Initiation of RNA decay in Escherichia coli by 5' pyrophosphate removal. *Mol Cell* **27**: 79-90.

Chen, S., Zhang, A., Blyn, L. B. and Storz, G. (2004) MicC, a second small-RNA regulator of Omp protein expression in Escherichia coli. *J Bacteriol* **186**: 6689-6697.

Chilcott, G. S. and Hughes, K. T. (2000) Coupling of flagellar gene expression to flagellar assembly in Salmonella enterica serovar typhimurium and Escherichia coli. *Microbiol Mol Biol Rev* **64**: 694-708.

Cho, B. K., Barrett, C. L., Knight, E. M., Park, Y. S. and Palsson, B. O. (2008) Genome-scale reconstruction of the Lrp regulatory network in Escherichia coli. *Proc Natl Acad Sci U S A* **105**: 19462-19467.

Coornaert, A., Lu, A., Mandin, P., Springer, M., Gottesman, S. and Guillier, M. (2010) MicA sRNA links the PhoP regulon to cell envelope stress. *Mol Microbiol* **76**: 467-479.

Corcoran, C. P., Podkaminski, D., Papenfort, K., Urban, J. H., Hinton, J. C. and Vogel, J. (2012) Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. *Mol Microbiol*.

Coyer, J., Andersen, J., Forst, S. A., Inouye, M. and Delihas, N. (1990) micF RNA in ompB mutants of Escherichia coli: different pathways regulate micF RNA levels in response to osmolarity and temperature change. *J Bacteriol* **172**: 4143-4150.

Darfeuille, F., Unoson, C., Vogel, J. and Wagner, E. G. H. (2007) An antisense RNA inhibits translation by competing with standby ribosomes. *Mol Cell* **26**: 381-392.

Deana, A., Celesnik, H. and Belasco, J. G. (2008) The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**: 355-358.

Delihas, N. and Forst, S. (2001) MicF: an antisense RNA gene involved in response of Escherichia coli to global stress factors. *J Mol Biol* **313**: 1-12.

Diestra, E., Cayrol, B., Arluison, V. and Risco, C. (2009) Cellular electron microscopy imaging reveals the localization of the Hfq protein close to the bacterial membrane. *PLoS One* **4**: e8301.

Dubnau, D. and Losick, R. (2006) Bistability in bacteria. *Mol Microbiol* **61**: 564-572.

Dulebohn, D., Choy, J., Sundermeier, T., Okan, N. and Karzai, A. W. (2007) Trans-translation: the tmRNA-mediated surveillance mechanism for ribosome rescue, directed protein degradation, and nonstop mRNA decay. *Biochemistry* **46**: 4681-4693.

Durand, S. and Storz, G. (2010) Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol Microbiol* **75**: 1215-1231.

Eggenhofer, F., Tafer, H., Stadler, P. F. and Hofacker, I. L. (2011) RNApredator: fast accessibility-based prediction of sRNA targets. *Nucleic Acids Res* **39**: W149-154.

Felden, B., Vandenesch, F., Bouloc, P. and Romby, P. (2011) The Staphylococcus aureus RNome and its commitment to virulence. *PLoS Pathog* **7**: e1002006.

Fender, A., Elf, J., Hampel, K., Zimmermann, B. and Wagner, E. G. H. (2010) RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev* **24**: 2621-2626.

Ferrario, M., Ernsting, B. R., Borst, D. W., Wiese, D. E., 2nd, Blumenthal, R. M. and Matthews, R. G. (1995) The leucine-responsive regulatory protein of Escherichia coli negatively regulates transcription of ompC and micF and positively regulates translation of ompF. *J Bacteriol* **177**: 103-113.

Figueroa-Bossi, N., Lemire, S., Maloriol, D., Balbontin, R., Casadesus, J. and Bossi, L. (2006) Loss of Hfq activates the sigmaE-dependent envelope stress response in Salmonella enterica. *Mol Microbiol* **62**: 838-852.

Figueroa-Bossi, N., Valentini, M., Malleret, L., Fiorini, F. and Bossi, L. (2009) Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev* **23**: 2004-2015.

Franze de Fernandez, M. T., Eoyang, L. and August, J. T. (1968) Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. *Nature* **219**: 588-590.

Garcia, B., Latasa, C., Solano, C., Garcia-del Portillo, F., Gamazo, C. and Lasa, I. (2004) Role of the GGDEF protein family in Salmonella cellulose biosynthesis and biofilm formation. *Mol Microbiol* **54**: 264-277.

Gerstel, U., Kolb, A. and Römling, U. (2006) Regulatory components at the csgD promoter--additional roles for OmpR and integration host factor and role of the 5' untranslated region. *FEMS Microbiol Lett* **261**: 109-117.

Gerstel, U., Park, C. and Römling, U. (2003) Complex regulation of csgD promoter activity by global regulatory proteins. *Mol Microbiol* **49**: 639-654.

Gibson, D. L., White, A. P., Rajotte, C. M. and Kay, W. W. (2007) AgfC and AgfE facilitate extracellular thin aggregative fimbriae synthesis in Salmonella enteritidis. *Microbiology* **153**: 1131-1140.

Giuliodori, A. M., Di Pietro, F., Marzi, S., Masquida, B., Wagner, R., Romby, P., Gualerzi, C. O. and Pon, C. L. (2010) The cspA mRNA is a thermosensor that modulates translation of the cold-shock protein CspA. *Mol Cell* **37**: 21-33.

Gogol, E. B., Rhodius, V. A., Papenfort, K., Vogel, J. and Gross, C. A. (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.

Gophna, U., Barlev, M., Seijffers, R., Oelschlager, T. A., Hacker, J. and Ron, E. Z. (2001) Curli fibers mediate internalization of Escherichia coli by eukaryotic cells. *Infect Immun* **69**: 2659-2665.

Grantcharova, N., Peters, V., Monteiro, C., Zakikhany, K. and Römling, U. (2010) Bistable expression of CsgD in biofilm development of Salmonella enterica serovar typhimurium. *J Bacteriol* **192**: 456-466.

Guillier, M. and Gottesman, S. (2006) Remodelling of the Escherichia coli outer membrane by two small regulatory RNAs. *Mol Microbiol* **59**: 231-247.

Guillier, M. and Gottesman, S. (2008) The 5' end of two redundant sRNAs is involved in the regulation of multiple targets, including their own regulator. *Nucleic Acids Res* **36**: 6781-6794.

Hammar, M., Arnqvist, A., Bian, Z., Olsen, A. and Normark, S. (1995) Expression of two csg operons is required for production of fibronectin- and congo red-binding curli polymers in Escherichia coli K-12. *Mol Microbiol* **18**: 661-670.

Han, K., Kim, K. S., Bak, G., Park, H. and Lee, Y. (2010) Recognition and discrimination of target mRNAs by Sib RNAs, a cis-encoded sRNA family. *Nucleic Acids Res* **38**: 5851-5866.

Hershey, A. D. and Chase, M. (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol* **36**: 39-56.

Huntzinger, E., Boisset, S., Saveanu, C., Benito, Y., Geissmann, T., Namane, A., Lina, G., Etienne, J., Ehresmann, B., Ehresmann, C., Jacquier, A., Vandenesch, F. and Romby, P. (2005) Staphylococcus aureus RNAIII and the endoribonuclease III coordinately regulate spa gene expression. *EMBO J* **24**: 824-835.

Hüttenhofer, A. and Noller, H. F. (1994) Footprinting mRNA-ribosome complexes with chemical probes. *EMBO J* **13**: 3892-3901.

Jacob, F. and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**: 318-356.

Johansen, J., Rasmussen, A. A., Overgaard, M. and Valentin-Hansen, P. (2006) Conserved small non-coding RNAs that belong to the sigmaE regulon: role in down-regulation of outer membrane proteins. *J Mol Biol* **364**: 1-8.

Johansson, J., Mandin, P., Renzoni, A., Chiaruttini, C., Springer, M. and Cossart, P. (2002) An RNA thermosensor controls expression of virulence genes in Listeria monocytogenes. *Cell* **110**: 551-561.

Jørgensen, M. G., Nielsen, J. S., Boysen, A., Franch, T., Møller-Jensen, J. and Valentin-Hansen, P. (2012) Small regulatory RNAs control the multi-cellular adhesive lifestyle of Escherichia coli. *Mol Microbiol*.

Jubelin, G., Vianney, A., Beloin, C., Ghigo, J. M., Lazzaroni, J. C., Lejeune, P. and Dorel, C. (2005) CpxR/OmpR interplay regulates curli gene expression in response to osmolarity in Escherichia coli. *J Bacteriol* **187**: 2038-2049.

Kaberdin, V. R. and Bläsi, U. (2006) Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol Rev* **30**: 967-979.

Kajitani, M., Kato, A., Wada, A., Inokuchi, Y. and Ishihama, A. (1994) Regulation of the Escherichia coli hfq gene encoding the host factor for phage Q beta. *J Bacteriol* **176**: 531-534.

Kawamoto, H., Morita, T., Shimizu, A., Inada, T. and Aiba, H. (2005) Implication of membrane localization of target mRNA in the action

of a small RNA: mechanism of post-transcriptional regulation of glucose transporter in Escherichia coli. *Genes Dev* **19**: 328-338.

Kawano, M., Reynolds, A. A., Miranda-Rios, J. and Storz, G. (2005) Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli. *Nucleic Acids Res* **33**: 1040-1050.

Lease, R. A., Cusick, M. E. and Belfort, M. (1998) Riboregulation in Escherichia coli: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc Natl Acad Sci U S A* **95**: 12456-12461.

Lee, R. C., Feinbaum, R. L. and Ambros, V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**: 843-854.

Lee, T. and Feig, A. L. (2008) The RNA binding protein Hfq interacts specifically with tRNAs. *RNA* **14**: 514-523.

Levine, E., Zhang, Z., Kuhlman, T. and Hwa, T. (2007) Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* **5**: e229.

Lewis, B. P., Burge, C. B. and Bartel, D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**: 15-20.

Lin, R., Ernsting, B., Hirshfield, I. N., Matthews, R. G., Neidhardt, F. C., Clark, R. L. and Newman, E. B. (1992) The lrp gene product regulates expression of lysU in Escherichia coli K-12. *J Bacteriol* **174**: 2779-2784.

Link, T. M., Valentin-Hansen, P. and Brennan, R. G. (2009) Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci U S A* **106**: 19292-19297.

Lorenz, C., Gesell, T., Zimmermann, B., Schoeberl, U., Bilusic, I., Rajkowitsch, L., Waldsich, C., von Haeseler, A. and Schroeder, R. (2010) Genomic SELEX for Hfq-binding RNAs identifies genomic aptamers predominantly in antisense transcripts. *Nucleic Acids Res* **38**: 3794-3808.

Lugtenberg, B., Peters, R., Bernheimer, H. and Berendsen, W. (1976) Influence of cultural conditions and mutations on the composition of the outer membrane proteins of Escherichia coli. *Mol Gen Genet* **147**: 251-262.

Majdalani, N., Chen, S., Murrow, J., St John, K. and Gottesman, S. (2001) Regulation of RpoS by a novel small RNA: the characterization of RprA. *Mol Microbiol* **39**: 1382-1394.

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. and Gottesman, S. (1998) DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci U S A* **95**: 12462-12467.

Mandin, P. and Gottesman, S. (2010) Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. *EMBO J* **29**: 3094-3107.

Masse, E., Escorcia, F. E. and Gottesman, S. (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes Dev* **17**: 2374-2383.

Masse, E. and 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.

Masse, E., Salvail, H., Desnoyers, G. and Arguin, M. (2007) Small RNAs controlling iron metabolism. *Curr Opin Microbiol* **10**: 140-145.

Mika, F., Busse, S., Possling, A., Berkholz, J., Tschowri, N., Sommerfeldt, N., Pruteanu, M. and Hengge, R. (2012) Targeting of csgD by the small regulatory RNA RprA links stationary phase, biofilm formation and cell envelope stress in Escherichia coli. *Mol Microbiol*.

Mironov, A. S., Gusarov, I., Rafikov, R., Lopez, L. E., Shatalin, K., Kreneva, R. A., Perumov, D. A. and Nudler, E. (2002) Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111**: 747-756.

Mizuno, T., Chou, M. Y. and Inouye, M. (1984) A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proc Natl Acad Sci U S A* **81**: 1966-1970.

Moll, I., Afonyushkin, T., Vytvytska, O., Kaberdin, V. R. and Bläsi, U. (2003) Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. *RNA* **9**: 1308-1314.

Moll, I., Grill, S., Gualerzi, C. O. and Blasi, U. (2002) Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. *Mol Microbiol* **43**: 239-246.

Montero Llopis, P., Jackson, A. F., Sliusarenko, O., Surovtsev, I., Heinritz, J., Emonet, T. and Jacobs-Wagner, C. (2010) Spatial organization of the flow of genetic information in bacteria. *Nature* **466**: 77-81.

Morita, M., Kanemori, M., Yanagi, H. and Yura, T. (1999) Heat-induced synthesis of sigma32 in Escherichia coli: structural and functional dissection of rpoH mRNA secondary structure. *J Bacteriol* **181**: 401-410.

Morita, M. T., Tanaka, Y., Kodama, T. S., Kyogoku, Y., Yanagi, H. and Yura, T. (1999) Translational induction of heat shock transcription factor sigma32: evidence for a built-in RNA thermosensor. *Genes Dev* **13**: 655-665.

Morita, T., Maki, K. and Aiba, H. (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* **19**: 2176-2186.

Muffler, A., Fischer, D. and Hengge-Aronis, R. (1996) The RNA-binding protein HF-I, known as a host factor for phage Qbeta RNA replication, is essential for rpoS translation in Escherichia coli. *Genes Dev* **10**: 1143-1151.

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L. and Breaker, R. R. (2002) Genetic control by a metabolite binding mRNA. *Chem Biol* **9**: 1043.

Nenninger, A. A., Robinson, L. S., Hammer, N. D., Epstein, E. A., Badtke, M. P., Hultgren, S. J. and Chapman, M. R. (2011) CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation. *Mol Microbiol* **81**: 486-499.

Nenninger, A. A., Robinson, L. S. and Hultgren, S. J. (2009) Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. *Proc Natl Acad Sci U S A* **106**: 900-905.

Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S. and Amster-Choder, O. (2011) Translation-independent localization of mRNA in E. coli. *Science* **331**: 1081-1084.

Newman, E. B. and Lin, R. (1995) Leucine-responsive regulatory protein: a global regulator of gene expression in E. coli. *Annu Rev Microbiol* **49**: 747-775.

Nikulin, A., Stolboushkina, E., Perederina, A., Vassilieva, I., Bläsi, U., Moll, I., Kachalova, G., Yokoyama, S., Vassylyev, D., Garber, M. and Nikonov, S. (2005) Structure of Pseudomonas aeruginosa Hfq protein. *Acta Crystallogr D Biol Crystallogr* **61**: 141-146.

Ogasawara, H., Hasegawa, A., Kanda, E., Miki, T., Yamamoto, K. and Ishihama, A. (2007) Genomic SELEX search for target promoters under the control of the PhoQP-RstBA signal relay cascade. *J Bacteriol* **189**: 4791-4799.

Ogasawara, H., Yamamoto, K. and Ishihama, A. (2010) Regulatory role of MlrA in transcription activation of csgD, the master regulator of biofilm formation in Escherichia coli. *FEMS Microbiol Lett* **312**: 160-168.

Ogasawara, H., Yamamoto, K. and Ishihama, A. (2011) Role of the biofilm master regulator CsgD in cross-regulation between biofilm formation and flagellar synthesis. *J Bacteriol* **193**: 2587-2597.

Ohnishi, K., Kutsukake, K., Suzuki, H. and Lino, T. (1992) A novel transcriptional regulation mechanism in the flagellar regulon of Salmonella typhimurium: an antisigma factor inhibits the activity of the flagellum-specific sigma factor, sigma F. *Mol Microbiol* **6**: 3149-3157.

Olsén, A., Wick, M. J., Mörgelin, M. and Björck, L. (1998) Curli, fibrous surface proteins of Escherichia coli, interact with major histocompatibility complex class I molecules. *Infect Immun* **66**: 944-949.

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A. and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **1**: 376-386.

Panja, S. and Woodson, S. A. (2012) Hexamer to Monomer Equilibrium of E. coli Hfq in Solution and Its Impact on RNA Annealing. *J Mol Biol*.

Papenfort, K., Bouvier, M., Mika, F., Sharma, C. M. and Vogel, J. (2010) Evidence for an autonomous 5' target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci U S A* **107**: 20435-20440.

Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J. C. and Vogel, J. (2006) SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay. *Mol Microbiol* **62**: 1674-1688.

Papenfort, K., Said, N., Welsink, T., Lucchini, S., Hinton, J. C. and Vogel, J. (2009) Specific and pleiotropic patterns of mRNA regulation by ArcZ, a conserved, Hfq-dependent small RNA. *Mol Microbiol* **74**: 139-158.

Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J. C. and Vogel, J. (2009) Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* **16**: 840-846.

Pfeiffer, V., Sittka, A., Tomer, R., Tedin, K., Brinkmann, V. and Vogel, J. (2007) A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the Salmonella core genome. *Mol Microbiol* **66**: 1174-1191.

Prevost, K., Desnoyers, G., Jacques, J. F., Lavoie, F. and Masse, E. (2011) Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. *Genes Dev* **25**: 385-396.

Raghavan, R., Groisman, E. A. and Ochman, H. (2011) Genome-wide detection of novel regulatory RNAs in E. coli. *Genome Res* **21**: 1487-1497.

Rasmussen, A. A., Eriksen, M., Gilany, K., Udesen, C., Franch, T., Petersen, C. and Valentin-Hansen, P. (2005) Regulation of ompA mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Mol Microbiol* **58**: 1421-1429.

Reimegård, J. (2010) Making Sense of Antisense. *PhD Thesis, Acta Universitatis Upsaliensis Uppsala*.

Robinson, L. S., Ashman, E. M., Hultgren, S. J. and Chapman, M. R. (2006) Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. *Mol Microbiol* **59**: 870-881.

Römling, U., Rohde, M., Olsén, A., Normark, S. and Reinkoster, J. (2000) AgfD, the checkpoint of multicellular and aggregative behaviour in Salmonella typhimurium regulates at least two independent pathways. *Mol Microbiol* **36**: 10-23.

Ruiz, N. and Silhavy, T. J. (2005) Sensing external stress: watchdogs of the Escherichia coli cell envelope. *Curr Opin Microbiol* **8**: 122-126.

Sauter, C., Basquin, J. and Suck, D. (2003) Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. *Nucleic Acids Res* **31**: 4091-4098.

Schumacher, M. A., Pearson, R. F., Møller, T., Valentin-Hansen, P. and Brennan, R. G. (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. *EMBO J* **21**: 3546-3556.

Sharma, C. M., Darfeuille, F., Plantinga, T. H. and 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, C. M., Papenfort, K., Pernitzsch, S. R., Mollenkopf, H. J., Hinton, J. C. and Vogel, J. (2011) Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. *Mol Microbiol* **81**: 1144-1165.

Shen-Orr, S. S., Milo, R., Mangan, S. and Alon, U. (2002) Network motifs in the transcriptional regulation network of Escherichia coli. *Nat Genet* **31**: 64-68.

Shimoni, Y., Friedlander, G., Hetzroni, G., Niv, G., Altuvia, S., Biham, O. and Margalit, H. (2007) Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol* **3**: 138.

Shinhara, A., Matsui, M., Hiraoka, K., Nomura, W., Hirano, R., Nakahigashi, K., Tomita, M., Mori, H. and Kanai, A. (2011) Deep sequencing reveals as-yet-undiscovered small RNAs in Escherichia coli. *BMC Genomics* **12**: 428.

Simm, R., Morr, M., Kader, A., Nimtz, M. and Römling, U. (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol Microbiol* **53**: 1123-1134.

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C. M., Rolle, K., Binnewies, T. T., Hinton, J. C. and Vogel, J. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* **4**: e1000163.

Sittka, A., Pfeiffer, V., Tedin, K. and Vogel, J. (2007) The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol Microbiol* **63**: 193-217.

Skippington, E. and Ragan, M. A. (2012) Evolutionary dynamics of small RNAs in 27 Escherichia coli and Shigella genomes. *Genome Biol Evol*.

Sonnleitner, E., Gonzalez, N., Sorger-Domenigg, T., Heeb, S., Richter, A. S., Backofen, R., Williams, P., Hüttenhofer, A., Haas, D. and Bläsi, U. (2011) The small RNA PhrS stimulates synthesis of the Pseudomonas aeruginosa quinolone signal. *Mol Microbiol* **80**: 868-885.

Soper, T. J. and Woodson, S. A. (2008) The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* **14**: 1907-1917.

Storz, G. (1999) An RNA thermometer. *Genes Dev* **13**: 633-636.

Stougaard, P., Molin, S. and Nordström, K. (1981) RNAs involved in copy-number control and incompatibility of plasmid R1. *Proc Natl Acad Sci U S A* **78**: 6008-6012.

Thomason, M. K., Fontaine, F., De Lay, N. and Storz, G. (2012) A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in Escherichia coli. *Mol Microbiol*.

Tjaden, B. (2008) TargetRNA: a tool for predicting targets of small RNA action in bacteria. *Nucleic Acids Res* **36**: W109-113.

Tomizawa, J., Itoh, T., Selzer, G. and Som, T. (1981) Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA. *Proc Natl Acad Sci U S A* **78**: 1421-1425.

Tsui, H. C., Leung, H. C. and Winkler, M. E. (1994) Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. *Mol Microbiol* **13**: 35-49.

Tucker, B. J. and Breaker, R. R. (2005) Riboswitches as versatile gene control elements. *Curr Opin Struct Biol* **15**: 342-348.

Udekwu, K. I. (2010) Transcriptional and post-transcriptional regulation of the Escherichia coli luxS mRNA; involvement of the sRNA MicA. *PLoS One* **5**: e13449.

Udekwu, K. I. and Wagner, E. G. H. (2007) Sigma E controls biogenesis of the antisense RNA MicA. *Nucleic Acids Res* **35**: 1279-1288.

Unoson, C. and Wagner, E. G. (2007) Dealing with stable structures at ribosome binding sites: bacterial translation and ribosome standby. *RNA Biol* **4**: 113-117.

Vecerek, B., Moll, I., Afonyushkin, T., Kaberdin, V. and Bläsi, U. (2003) Interaction of the RNA chaperone Hfq with mRNAs: direct and indirect roles of Hfq in iron metabolism of Escherichia coli. *Mol Microbiol* **50**: 897-909.

Vecerek, B., Moll, I. and Bläsi, U. (2005) Translational autocontrol of the Escherichia coli hfq RNA chaperone gene. *RNA* **11**: 976-984.

Vecerek, B., Moll, I. and Bläsi, U. (2007) Control of Fur synthesis by the non-coding RNA RyhB and iron-responsive decoding. *EMBO J* **26**: 965-975.

Veening, J. W., Smits, W. K. and Kuipers, O. P. (2008) Bistability, epigenetics, and bet-hedging in bacteria. *Annu Rev Microbiol* **62**: 193-210.

Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M. and Lejeune, P. (1998) Isolation of an Escherichia coli K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. *J Bacteriol* **180**: 2442-2449.

Viegas, S. C., Silva, I. J., Saramago, M., Domingues, S. and Arraiano, C. M. (2011) Regulation of the small regulatory RNA MicA by ribonuclease III: a target-dependent pathway. *Nucleic Acids Res* **39**: 2918-2930.

Vogel, J., Bartels, V., Tang, T. H., Churakov, G., Slagter-Jäger, J. G., Hüttenhofer, A. and Wagner, E. G. H. (2003) RNomics in Escherichia coli detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res* **31**: 6435-6443.

Wadler, C. S. and Vanderpool, C. K. (2007) A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. *Proc Natl Acad Sci U S A* **104**: 20454-20459.

Wagner, E. G. H., Altuvia, S. and Romby, P. (2002) Antisense RNAs in bacteria and their genetic elements. *Adv Genet* **46**: 361-398.

Wang, Q., Wu, J., Friedberg, D., Plakto, J. and Calvo, J. M. (1994) Regulation of the Escherichia coli lrp gene. *J Bacteriol* **176**: 1831-1839.

Wang, X., Rochon, M., Lamprokostopoulou, A., Lunsdorf, H., Nimtz, M. and Römling, U. (2006) Impact of biofilm matrix components on interaction of commensal Escherichia coli with the gastrointestinal cell line HT-29. *Cell Mol Life Sci* **63**: 2352-2363.

Wassarman, K. M. (2007) 6S RNA: a small RNA regulator of transcription. *Curr Opin Microbiol* **10**: 164-168.

Wassarman, K. M., Repoila, F., Rosenow, C., Storz, G. and Gottesman, S. (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* **15**: 1637-1651.

White, A. P., Gibson, D. L., Grassl, G. A., Kay, W. W., Finlay, B. B., Vallance, B. A. and Surette, M. G. (2008) Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in Salmonella enterica serovar Typhimurium. *Infect Immun* **76**: 1048-1058.

Winkler, W., Nahvi, A. and Breaker, R. R. (2002) Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**: 952-956.

Winkler, W. C., Cohen-Chalamish, S. and Breaker, R. R. (2002) An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A* **99**: 15908-15913.

Wolf, D. M. and Arkin, A. P. (2002) Fifteen minutes of fim: control of type 1 pili expression in E. coli. *OMICS* **6**: 91-114.

Wood, T. K., Gonzalez Barrios, A. F., Herzberg, M. and Lee, J. (2006) Motility influences biofilm architecture in Escherichia coli. *Appl Microbiol Biotechnol* **72**: 361-367.

Ying, X., Cao, Y., Wu, J., Liu, Q., Cha, L. and Li, W. (2011) sTarPicker: a method for efficient prediction of bacterial sRNA targets based on a two-step model for hybridization. *PLoS One* **6**: e22705.

Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R. and Storz, G. (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *EMBO J* **17**: 6061-6068.

Zhang, A., Wassarman, K. M., Rosenow, C., Tjaden, B. C., Storz, G. and Gottesman, S. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* **50**: 1111-1124.

Zheng, D., Constantinidou, C., Hobman, J. L. and Minchin, S. D. (2004) Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Res* **32**: 5874-5893.

Zogaj, X., Nimtz, M., Rohde, M., Bokranz, W. and Römling, U. (2001) The multicellular morphotypes of Salmonella typhimurium and Escherichia coli produce cellulose as the second component of the extracellular matrix. *Mol Microbiol* **39**: 1452-1463.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 916

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology.

Distribution: publications.uu.se
urn:nbn:se:uu:diva-171642

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2012