UPTEC X 05 043
AUG 2005

ISSN 1401-2138

ERIK HOLMQVIST

Functional analysis of chromosomally encoded small regulatory RNAs in *E. coli*

Master's degree project

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 05 043	**Date of issue 2005-08**

Author

Erik Holmqvist

Title (English)

Functional analysis of chromosomally encoded small regulatory RNAs in *E. coli*

Title (Swedish)

Abstract

The *E. coli* small regulatory RNA MicA has here been shown to act as an antisense RNA targeting *ompA*-mRNA *in vivo*. Reporter gene fusions further showed that over-expression of the sRNAs MicA and MicF leads to decreased translation of *lrp*-mRNA, encoding the leucin responsive regulator protein.

Keywords

sRNA, post-transcriptional regulation, reporter gene system, structural probing

Supervisors

Gerhart Wagner
Department of Cell and Molecular Biology, Uppsala University

Scientific reviewer

Fredrik Söderbom
Departement of Molecular Biology, Swedish University of Agricultural Sciences

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **23**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

Functional analysis of chromosomally encoded small regulatory RNAs in *E. coli*

Erik Holmqvist

Sammanfattning

Små regulatoriska RNA-molekyler (sRNA) är en nyligen upptäckt grupp icke-kodande RNA-molekyler i bakterier. Av de sRNA som studerats noggrant verkar de flesta fungera som hämmare av proteinsyntes. Detta sker genom att ett sRNA basparar med ett specifikt mRNA och därmed försämrar ribosomens bindningsmöjligheter. Ett RNA som binder till en motsvarande region på ett annat RNA kallas för antisens-RNA. I detta projekt har de tre sRNA-molekylerna MicA, MicC och MicF studerats. Dessa har tidigare visat sig verka hämmande på syntes av de respektive proteinerna OmpA, OmpC och OmpF. För att studera MicA:s effekt på OmpA-syntes *in vivo* länkades OmpA till en så kallad reportergen, vars genprodukt är möjlig att läsa av. Försöken visade tydligt att hög produktion av MicA leder till kraftigt hämmad syntes av OmpA. För att studera om denna hämning sker pga. RNA-RNA interaktioner infördes mutationer i den förmodade bindningsregionen. Resultatet var entydigt, hämningen var inte längre möjlig när bindningsregionen förstörts genom mutationer. Dessutom kunde hämningen återskapas med kompenserande mutationer. Dessa resultat visar att MicA verkar hämmande på OmpA-syntes genom en antisensbindning till *ompA*-mRNA. Reportergenen användes även till att testa ett möjligt mål-mRNA för MicA och MicF. Försöken visade att hög produktion av i synnerhet MicF, men även MicA, hämmar syntes av proteinet Lrp. För att i detalj studera sRNA-mRNA interaktioner utfördes bindningsstudier *in vitro*. Dels studerades bindningsstyrkan mellan de tre sRNA-molekylerna och deras mål-mRNA-molekyler, dels gjordes detaljerade studier av själva bindningsregionerna. Resultaten är viktiga för vidare förståelse av genreglering med hjälp av antisens-RNA.

Examensarbete 20 p i Molekylär bioteknikprogrammet

Uppsala Universitet augusti 2005

1. TABLE OF CONTENTS

2 INTRODUCTION

2.1 Introduction to sRNAs

In the last five years the knowledge of RNA molecules and their functions has broadened considerably. Generally the focus on RNA function has been in genetic information transfer and protein synthesis. The messenger RNA (mRNA) is the template that brings the genetic information, stored in the DNA, to the translational machinery for protein synthesis. The transfer RNA (tRNA) and ribosomal RNA (rRNA) have functional roles in protein synthesis. RNAs, such as tRNA and rRNA that are not translated into a protein belong to the so-called non-coding RNAs (ncRNAs). The functions of ncRNAs can be structural, regulatory or catalytic.

In the recent years many novel ncRNAs have been discovered in bacteria. Since they are relatively small, the term small RNAs (sRNAs) have been used. Among the sRNAs found so far their sizes range from 50 to 400 nucleotides [1]. The searches for sRNAs have mainly been focused on the bacterium *Escherichia coli* in which more than 50 novel chromosomally encoded sRNAs have been identified [1]. Only a small fraction of these are characterized in terms of function. Similar types of RNAs have been found in eukaryotic cells, such as microRNAs (miRNAs) and small interfering RNAs (siRNAs).

2.2 Roles of sRNAs in post-transcriptional gene regulation

The production of a protein can be regulated at several levels. At the transcriptional level, proteins referred to as transcription factors bind to promoter sequences to induce or repress transcription. At the post-transcriptional level, the translation of the mRNA can be induced or repressed. The function of a protein can also depend on modifications of the protein itself.

Many of the studied sRNAs seem to be involved in post-transcriptional gene regulation [2]. In many cases these RNAs interact with mRNA targets through a so-called antisense binding. Although antisense RNAs encoded by plasmids and bacteriophages have been studied for long, the chromosomally encoded antisense binding sRNAs are a relatively recent discovery. An antisense RNA can be either *cis*- or *trans*-encoded [2] (Fig. 1). A gene encoding a *cis*-encoded sRNA is located in the same locus as the gene encoding its target-mRNA, but is transcribed from the opposite strand and therefore in the opposite direction. This implies that the sRNA is fully complementary to its target mRNA. A *trans*-encoded sRNA is transcribed from a separate locus and therefore not necessarily fully complementary to its target mRNA [2]. The sRNAs studied in this project are *trans*-encoded. The binding of an sRNA to its target mRNA can lead to either inhibition or activation of the mRNA function. For example, the sRNA DsrA binds to the rpoS-mRNA and opens up a stretch of nucleotides that in absence of DsrA sequesters the translation initiation region (TIR). The TIR thereby becomes accessible for the translation machinery. MicC, MicF, IstR-1 (and many others) are examples of sRNAs that functions in the opposite way. By binding to the TIR these sRNAs inhibit the initiation of translation resulting in inhibition of protein synthesis.

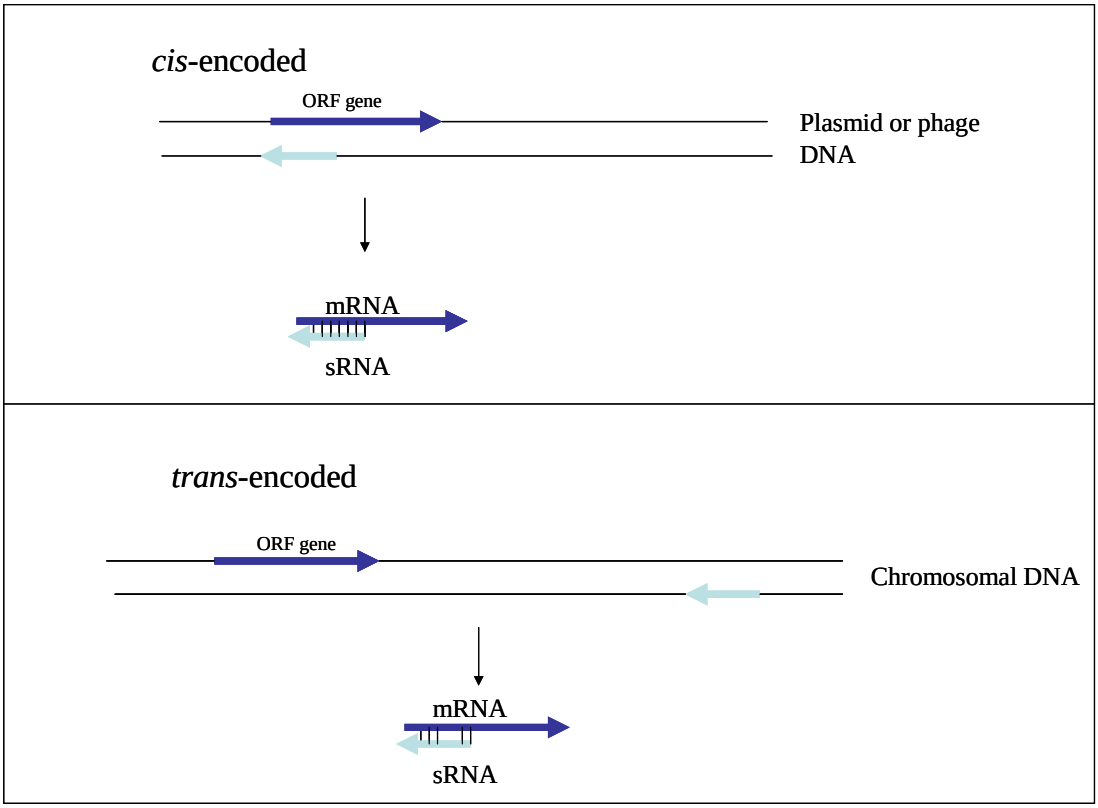

Figure 1. Schematic view of *cis*- versus *trans*-encoded antisense RNAs.

2.3 Examples of sRNAs

Many of the studied sRNAs in *E. coli* are not expressed in exponentially growing cells, but are rather up-regulated when cells enter stationary phase. The transition to stationary phase requires that cells adapt to the new conditions, and several stress responses are indeed activated during the transition. The up-regulation of sRNAs under these conditions suggests that they may be part of the stress response regulation.

OxyS is an sRNA that is transcriptionally controlled by the OxyR protein and up-regulated during oxidative stress [3]. OxyS has been shown to target two mRNAs; *fhlA* and *rpoS*. Since *fhlA* encodes a transcriptional activator, the negative regulation by OxyS may affect genes regulated by FhlA. It has been shown that OxyS acts as an antisense RNA blocking the ribosome binding region of the *fhlA*-mRNA [3]. The negative regulation of RpoS by OxyS is not well understood [1].

Iron levels in the cell have to be carefully controlled since excess iron can cause damage [1]. However, since iron is an essential nutrient, cells need specific iron assimilation and iron storage systems. When iron is plentiful in the cell, the Fur protein represses transcription of genes encoding iron assimilation proteins. Fur also represses transcription of RyhB, an sRNA. RyhB targets two mRNAs (*sodB* and *sdhD*), both encoding iron storage proteins. When iron is lacking, Fur becomes inactive and iron assimilation proteins are translated. Simultaneously, RyhB accumulates and represses synthesis of iron storage proteins [4, 5]. Interestingly, RyhB targets

the TIR in the second cistron of the *sdhCDAB* operon. Previously, no other mechanism has been proposed by which only one gene in a bacterial operon is down-regulated by such a mechanism.

The sRNAs mentioned above act by an antisense mechanism. The sRNA CsrB acts by a different mechanism, called protein sequestration. CsrB is build up of 18 similar sequence motifs which interact with the protein CsrA. Since CsrA is a regulatory protein, the sequestration by CsrB thus upregulates the genes that are repressed by CsrA [6].

2.4 Regulation of outer membrane proteins by sRNAs

A typical example of proteins whose expression is regulated both transcriptionally and post-transcriptionally are the outer membrane proteins OmpC and OmpF. The *ompC* and *ompF* genes encode outer membrane porins that allow nonspecific passage of soluble low-molecular-weight molecules through the outer cell membrane of *E. coli* [7]. Of the two porins, OmpF has the larger pore diameter [8]. The two porins are thought to be dominant in opposite environmental conditions. In environments with high nutrient and toxin concentrations, high temperatures and high osmolarities, the smaller pore diameter porin OmpC is thought to be most important. In opposite environmental conditions, the bigger pore diameter porin OmpF is thought to be most important [7].

The OmpR response regulator regulates the expression of OmpC and OmpF at the transcriptional level. As a response to an external signal, the EnvZ sensor protein phosphorylates or dephosphorylates OmpR. When the osmolarity of the medium is high, OmpR-P is present at high concentration, thereby activating *ompC* transcription. At these conditions, the transcription of *ompF* is repressed. When OmpR-P is present at low concentrations, transcription of *ompF* is derepressed [8].

The sRNAs MicC and MicF have been reported to regulate *ompC* and *ompF* expression at the posttranscriptional level [7, 9]. Both sRNAs inhibit translation by base pairing to the ribosome binding regions of their respective target mRNAs; MicC to *ompC*-mRNA and MicF to *ompF*-mRNA [7, 10]. The transcription profiles of MicC and MicF are reciprocal to translation of *ompC*- and *ompF*-mRNA. At high osmolarities MicF is active and thereby represses and downregulates translation of *ompF*-mRNA. At low osmolarities MicC is active and thereby represses translation of *ompC*-mRNA.

2.5 Approaches to identify novel sRNAs and their targets.

Biocomputational searches have shown to be a fruitful approach in the hunt for novel sRNAs in *E. coli*. To identify sequences that may encode sRNAs, the properties of known sRNA-encoding genes have been considered. They often (but not always) reside in intergenic regions. Therefore focus has been put on intergenic regions with specific features such as promoters recognized by the major RNA polymerase sigma factor, and Rho-independent terminators. The length of the putative sRNAs as well as their conservation among close relatives has also been considered in these searches [11]. Detection of sRNA transcripts by microarrays has also turned out successful.

Various methods can be used to identify potential targets of sRNAs. Microarrays can be used to monitor changes in mRNA levels due to sRNA overexpression. 2D-PAGE can be used to assess protein levels. Significant changes in protein levels due to over-expression of an sRNA give a clue as to which proteins may be regulated by the sRNA. In-house investigations have shown that biocomputational screening of an sRNA against all annotated mRNAs may be a powerful tool for sRNA target prediction [12].

2.6 Aims of the project

One aim of this project was to use various techniques to find target mRNAs for antisense sRNAs in *E. coli*. There are some key properties that are typical for an sRNA-mRNA system. The interaction between RNA species consists of basepairing, either contiguous or non-contiguous. The part of the sRNA that interacts with the mRNA in most cases is weakly structured, so that the interacting nucleotides are accessible. The interacting region of the mRNA either covers the translation initiation region (TIR), or lies in the immediate vicinity of the TIR, so that basepairing results in inhibition of translation initiation. As described above, bioinformatic searches can be used to predict putative sRNA targets. In such searches, properties of known sRNAs, as those described above, can be used to refine the searches. In this project I have used results from bioinformatic searches to select sRNA targets for experimental testing.

To study post-transcriptional sRNA regulation *in vivo*, I have used reporter gene technology where the target mRNA gene is fused to a reporter gene and provided on a plasmid. By measuring the activity of the reporter gene product, effects due to sRNA over-expression can be detected.

To study the specificity of sRNA-mRNA binding, I have used a technique called gel shift. This is an *in vitro* technique where sRNA-mRNA complexes can be detected by gel-electrophoresis. A third technique called structural probing has been used to map the sRNA-mRNA interaction region *in vitro*.

3 MATERIALS AND METHODS

3.1 Plasmid constructions

All oligodeoxyribonucleotides used were purchased from Sigma-Genosys (Table 1). Plasmids constructed and used in this study are listed in Table 2. A DNA fragment of the *ompA* leader starting at position -88 and ending at the 12th codon and a fragment of the *lrp* leader starting at position –72 and ending at the 23rd codon was generated by a PCR reaction with primer pairs ompA-pMC874-5'/ompA-pMC874-3' and Lrp_fwd/Lrp_rev respectively (Table 1). The reaction was performed with 50 ng of chromosomal *E. coli* K12 DNA as template and 100 pmol of each primer in a 30 μl final volume with PuReTaq Ready-To-Go™ PCR Beads (GE Healthcare). The reaction conditions were as follows: 10 min at 95°C; 35 cycles of 1 min at 95°C, 40 s at 52°C, 1

min at 72°C; and a final 10 min extension at 72°C. PCR product was analyzed on a 2% agarose gel in 0.5 x TEB buffer. 2 μg of each PCR fragment were cleaved at 37°C for 1 hour with 20 U *Bam*HI and 10 x *Bam*HI Reaction Buffer (Fermentas) in a 50 μl final volume. Proteins were removed by phenol-chloroform extraction and DNA was precipitated by adding 2.5 V 99.5% ice-cold EtOH and 0.1 vol. 3 M NaAc and put in 20°C for 1 hour. Precipitated DNA was resuspended in 10 μl ddH_20. The cleaved DNA was ligated to cleaved and dephosphorylated plasmid pMC874 using Ready-To-Go™ T4 DNA Ligase (GE Healthcare) according to the manufacturer (for details on pMC874, see [13]). The ligation mix was added to 50 μl One Shot® TOP 10 Chemically Competent *E. coli* (Invitrogen). The transformation mix was incubated 30 min on ice, 1 min at 42°C, and 1 hour at 37°C after addition of 200 μl SOC media. 200 μl culture was spread on LA plates with 50 μg/ml kanamycin and 40 μg/ml X-gal. Plasmids from blue colonies were isolated and sequenced. Sequencing confirmed that the *ompA* and *lrp* fragments were ligated in frame with the 8th codon of the *lacZ* gene of pMC874.

Table 1. Oligonucleotides used in PCR amplification.

Name	Sequence (5' to 3')	Amplified fragment
T7-ompC	GAAATTAATACGACTCACTATAGCCGACTGATTAATGAGGGT	*ompC*
3'ompC	TCAGAGAAATAGTGCAGGCC	
T7-MicC	GAAATTAATACGACTCACTATAGTTATATGCCTTTATTGTCACA	*micC*
3'MicC	AAAAAGCCCGGACGACTGTT	
T7-OmpA	GAAATTAATACGACTCACTATAGGCCAGGGGTGCTCGGCATAA	*ompA*
3'OmpA	GCCAGTGCCACTGCAATCGCGATA	
T7-MicA	GAAATTAATACGACTCACTATAGGAAAGACGCGCATTTGTTAT	*micA*
3'MicA	GAAAAAGGCCACTCGTGAGT	
T7-OmpF	GAAATTAATACGACTCACTAATAGAGACACATAAAGACACCAAACTC	*ompF*
3'OmpF	GATCTACTTTGTTGCCATCTTG	
T7-MicF	GAAATTAATACGACTCACTATAGGCTATCATCATTAACTTTAT	*micF*
3'MicF	GAAAAAAAACCGAATGCGAGGCATCCG	
Lrp_fw	CGGGATCCTATCTGGCATGTTGTACT	*Lrp*
Lrp-rev	CGGGATCCAACTCATTAAGAATGTTACG	
OmpAM6+	GGCGTATTTTGGTAGTATTCGAGGCGCAA	*ompA-M6*
OmpAM6-	TTGCGCCTCGAATAGTACCAAAATACGCC	
MicAM4+	GACGCGCATTTGAATACATCATCCCTGAATTCAG	*micA-M4*
MicAM4-	CTGAATTCAGGGATGATGTATTCAAATGCGCGTC	
MicAM6+	GACGCGCATTTGAATACTACATCCCTGAATTCAG	*micA-M6*
MicAM6-	CTGAATTCAGGGATGTAGTATTCAAATGCGCGTC	
ompA-pMC874-5'	CGGGATCCGATTAAACATACCTTATACAAGAC	*ompA*
ompA-pMC874-3'	CGGGATCCAGTGCCACTGAATCGCGATAGCT	

3.2 Transformations

100 ng of plasmids pControl, pMicA, pMicAM4 and pMicAM6 were added to 50 μl of competent *E. coli* MC4100 cells. The transformation mix was incubated for 30 min on ice, 1 min at 42°C and 1 hour at 37°C after addition of 200 μl SOC media. 200μl culture was spread on LA plates with 50 μg/ml ampicillin. 100 ng of plasmids pEH2 and pOmpAM6 were added to 50 μl of competent *E. coli* MC4100 containing one of the four plasmids pControl, pMicA, pMicAM4 and pMicAM6. 100 ng of plasmid pEH5 was added to 50 μl of competent *E. coli* MC4100 containing one of the three plasmids pControl, pMicA and pMicF. Transformations were performed as described above. Cultures were spread on LA plates with 50 μg/ml kanamycin. Colonies were re-streaked on LA plates with 50 μg/ml ampicillin, 50 μg/ml kanamycin and 40 μg/ml X-gal.

Table 2. Plasmids and their construction

Plasmids contructed in this study			
Name	Fragment	Parental plasmid	Marker
pEH2	*ompA-lacZ*	pMC874	Km
pEH5	*lrp-lacZ*	pMC874	Km
pOmpAM6	*ompAM6-lacZ*	pMC874	Km
pMicAM4	*micAM4*	ColE1	Amp
pMicAM6	*micAM6*	ColE1	Amp
Previously constructed plasmids used in this study			
pControl (pJV968-1)	Promoterless *lacZ*	ColE1	Amp
pMicA	*mica*	ColE1	Amp
pMicF	*micF*	ColE1	Amp

3.3 Site-directed mutagenesis

Site-directed mutagenesis on plasmid pEH2 carrying the *ompA-lacZ* fusion was performed with the QuikChange® XL Site-Directed Mutagenesis Kit (Stratagene) to create pOmpAM6. Mutagenesis of pMicA was performed with QuikChange® Site-Directed Mutagenesis Kit (Stratagene) to create pMicAM4 and pMicAM6. Mutagenesis was performed according to the manufacturer. Successful mutagenesis was confirmed by sequencing.

3.4 β-galactosidase activity assays

100 μl of an overnight culture were inoculated in 20 ml LB pre-warmed to 37°C. Ampicillin and kanamycin concentrations were 50 μg/ml. At OD_{540} = 0.5, 1 ml cells were spun down at 4°C for 5 minutes at 15000 rpm in a tabletop centrifuge. Pellets were resuspended in Z-buffer to 1.0 OD/ml. To lyse the cells, 75 μl chloroform and 50 μl 0.1% SDS was added followed by 15 seconds vortexing. 200 μl were then transferred to a microtiterplate. 40 μl ONPG were added to each well followed by a direct spectrophotometric measurement on a Labsystems Multiskan MS at 414 and 540 nm. Measurements were taken every 10 minutes during one hour. Between the

measurements samples were incubated in a 28°C shaker. β-galactosidase activity was evaluated with the DeltaSOFT 3 software (BioMetallics) and transformed to Miller units (1000 x (activity/cell OD)).

3.5 Generation of templates for *in vitro* transcription

PCR fragments were generated for *ompA, micA, ompC, micC, ompF* and *mic F* with primers containing a T7 RNA-polymerase promoter sequence as shown in table 1. The reaction was performed with 50 ng of chromosomal *E. coli* K12 DNA and 100 pmol of each primer in a 30 µl final volume with PuReTaq Ready-To-Go™ PCR Beads (GE Healthcare). The reaction conditions were as follows: 10 min at 95°C; 35 cycles of 1 min at 95°C, 40 s at 52°C, 1 min at 72°C; and a final 10 min extension at 72°C. PCR products were analyzed on a 2% agarose gel in 0.5 x TEB buffer. The bands corresponding to the appropriate sizes were then purified with QIAquick™Gel Extraction Kit (Qiagen).

3.6 *In vitro* transcription

To generate RNAs, an *in vitro* transcription reaction was performed with 500 ng DNA template, 7 mM rNTPs (GE Healthcare), 1.25 µl RNAguard™ Ribonuclease Inhibitor (GE Healthcare), 4 x reaction buffer and T7 RNA polymerase in a 250 µl final reaction volume. The reaction was performed for 4 hours at 37°C. To degrade DNA, 5 U of RQ1 DNase were added and incubation continued for 30 min at 37°C. The samples were then put on ice. Proteins were removed with phenol, chloroform and isoamylalcohol extraction. RNA were precipitated by adding 2.5 V 99.5% ice-cold EtOH and 0.1 V 3M NaAc and put in -20°C for 1 hour. Samples were then centrifuged for 30 min at 4°C to pellet precipitated RNA. The dried pellets were resuspended in 50 µl dH_2O. RNA was further purified by gel filtration through a Sephadex^{TM} G-50 (GE Healthcare) column and subsequently stored in -20°C. The purified RNA was analyzed by gel electrophoresis on a 10% polyacrylamide gel.

3.7 RNA 5'-end-labeling

20 pmol of each RNA species were dephosphorylated with 1 U shrimp alkaline phosphatase (USB) and 10 x buffer in a 20 µl final volume. The reaction mix was incubated for 30 min at 37°C. Proteins were removed by phenol-chloroform extraction followed by EtOH-precipitation as described above. The dissolved RNAs were mixed with 10 x Reaction Buffer, 30 U T4 polynucleotide kinase (USB) and 20 µCi γ-^{32}P-ATP in a 20 µl final volume. The reaction was carried out for 1 hour at 37°C. Samples were purified on an 8% polyacrylamide gel. Bands of appropriate sizes were cut out and the RNA was eluted. Finally a phenol-chloroform extraction followed by an EtOH precipitation was performed. Labeled RNA was resuspended in ddH_2O.

3.8 Gel shift experiments

The gel shift experiments shown in Figure 6 were performed as follows. Labeled sRNAs and unlabeled mRNA species were denatured for 1 minute at 95° and then put on ice to refold. TMN-buffer was added to a 1 x TMN final concentration. sRNA and mRNA were then pooled so that the final concentration of mRNA was 250 nM. The mix was incubated for 30 minutes at 37°C. After addition of 2 μl loading buffer (48% glycerol, Bromophenol Blue) the samples were loaded on a native polyacrylamide gel (5% polyacrylamide, 0.5 x TBE, 1% APS, 0.1% Temed). The PAGE was run in 0.5 x TBE buffer at 5 W in cold-room. The gel was then dried and analyzed by a Phosphor Imager (Molecular Dynamics).

3.9 Structural probing

Labeled RNA was denatured for 1 minute at 95°C and put on ice followed by addition of TMN-buffer (1 x TMN final concentration) and 1 μg yeast tRNA. Unlabeled RNA was then added to a final concentration of 1 μM and the mix was incubated 15 minutes at 37°C. The RNA was then cleaved with RNase T1 (0.01 U; Ambion), RNase T2 (0.02 U; Invitrogen) or lead(II) acetate (5 mM; Sigma-Aldrich). The cleavage reactions were performed at 37°C and continued for 5 minutes for the RNases and 1 minute for lead(II) acetate. The reactions were stopped by adding 5 μl 0.1 M EDTA and were then directly put on ice. The RNA was precipitated by adding 75 μl ice-cold 95% EtOH and then put at -20°C for 1 hour. The precipitate was spun down at 15000 rpm for 30 minutes at 4°C. Pellets were dried and resuspended in 10 μl H_20 and 10 μl denaturing loading buffer II (Ambion). An alkaline hydrolysis ladder was prepared by mixing the labeled RNA with an alkaline buffer (Ambion) and incubating for 5 minutes at 95°C followed by addition of loading buffer II (Ambion). An RNaseT1 ladder was prepared by first denaturing the labeled RNA in a sequencing buffer for 1 minute at 95°C followed by addition of 0.1 U RNase T1. The cleavage reaction continued for 5 minutes at 37°C and was stopped by addition of loading buffer II (Ambion). Samples were denatured for 1 minute at 95°C and then loaded on a denaturing polyacrylamide gel (8% polyacrylamide, 7 M Urea, 1% APS, 0.1% Temed in 1 x TBE buffer). The PAGE was run at 38 W.

4 RESULTS

4.1 The expression of an *ompA-lacZ* translational fusion is inhibited by MicA.

OmpA is an outer membrane protein with somewhat different properties from OmpC and OmpF (described above). It is not clear whether OmpA works as a porin, and there have been reports with contradictory results [14]. Transcriptional regulation of *ompA* has been shown to be conducted by the cyclic AMP receptor protein-cyclic AMP complex [15].

The *micA* gene (earlier *sraD*) lies in the intergenic region between *luxS* and *gshA* on the *E. coli* chromosome. The *micA* gene is transcribed but does not have an open reading frame, thus expressing a non-coding RNA. Recently, results from 2D-PAGE have shown that OmpA levels are reduced in cells where MicA RNA is constitutively expressed [16]. When entering stationary phase, *ompA* is down-regulated about 4-fold [17]. Analyses by Northern blot have shown that this down-regulation is abolished in a Δ*micA* strain [16].

A reporter gene system was set up for measuring sRNA regulation *in vivo* (Figure 2). A fragment of the target mRNA-encoding gene was first generated by PCR. This fragment contained the promoter, the translation initiation region and some of the first codons. On plasmid pMC874, this fragment was placed in frame with the reporter gene *lacZ*. Transcription and translation of *lacZ* was thereby depending on the regulation of the target mRNA gene. *lacZ* encodes β-galactosidase which hydrolyzes lactose to glucose and galactose. To measure β-galactosidase activity the substrate ONPG was used. When ONPG is hydrolyzed by β-galactosidase, ortho-nitrophenol is produced. Ortho-nitrophenol gives a yellow color that can be measured spectrophotometrically. Post-transcriptional sRNA regulation was studied by measuring the β-galactosidase activity in the presence or absence of the sRNA.

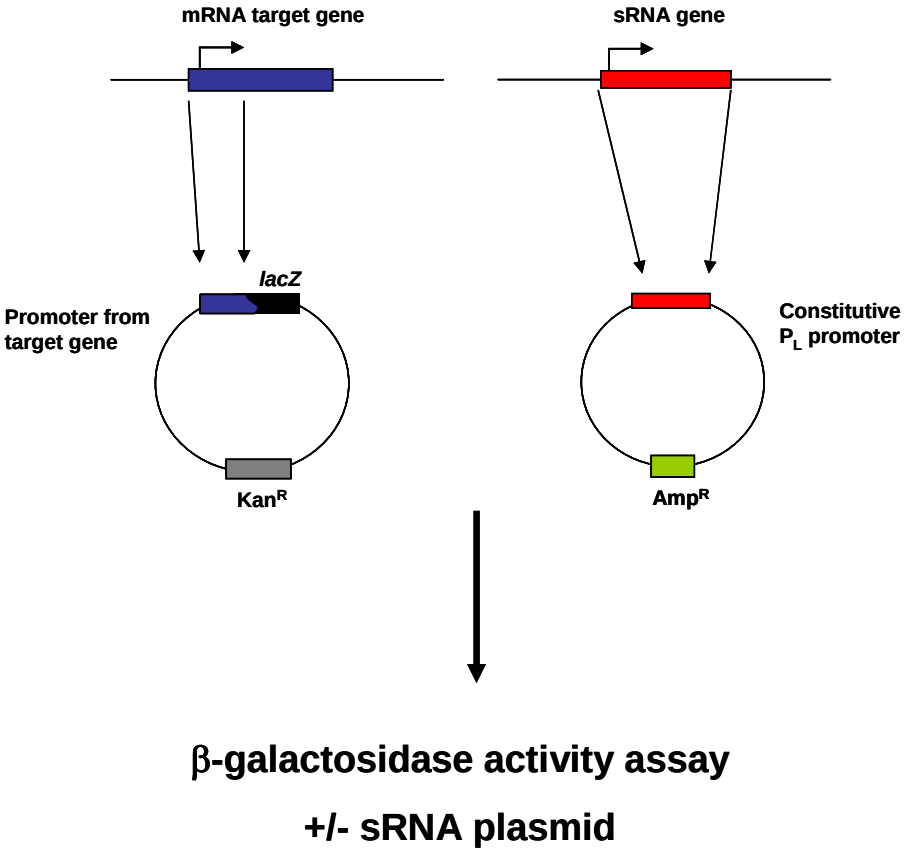

Figure 2. Schematic picture of the reporter gene system. The reporter gene *lacZ* is under target mRNA gene regulation, both transcriptional and translational. The sRNA is constitutively expressed from a high-copy plasmid.

To test this system, the proposed inhibition of *ompA*-mRNA translation by MicA was measured. *E. coli* MC4100 competent cells were transformed with the reporter plasmid pEH2 carrying an *ompA-lacZ* translational fusion. In addition to pEH2, cells were also transformed with one of the three high copy plasmids pMicA, pAnti or pControl. pMicA constitutively expresses MicA RNA, pAnti constitutively expresses the antisense of MicA, and pControl carries a promoterless *lacZ* fragment. Cells were grown in LB media and harvested at an OD_{540} of 0.5. After lysing the cells, ONPG was added to the extracts and production of ortho-nitrophenol was measured at several time points. As expected, over-expression of MicA did affect the translation of the *ompA-lacZ* fusion negatively. As shown in Figure 3, the β-galactosidase activity was 4 times lower in cells carrying pMicA compared to cells carrying pControl. These results verified that the reporter gene system worked satisfactorily and could be used for subsequent experiments. Cells carrying pAnti showed an even higher β-galactosidase activity than the control. This may be due to out-titration of chromosomally encoded MicA since pAnti encodes an RNA that is antisense to MicA.

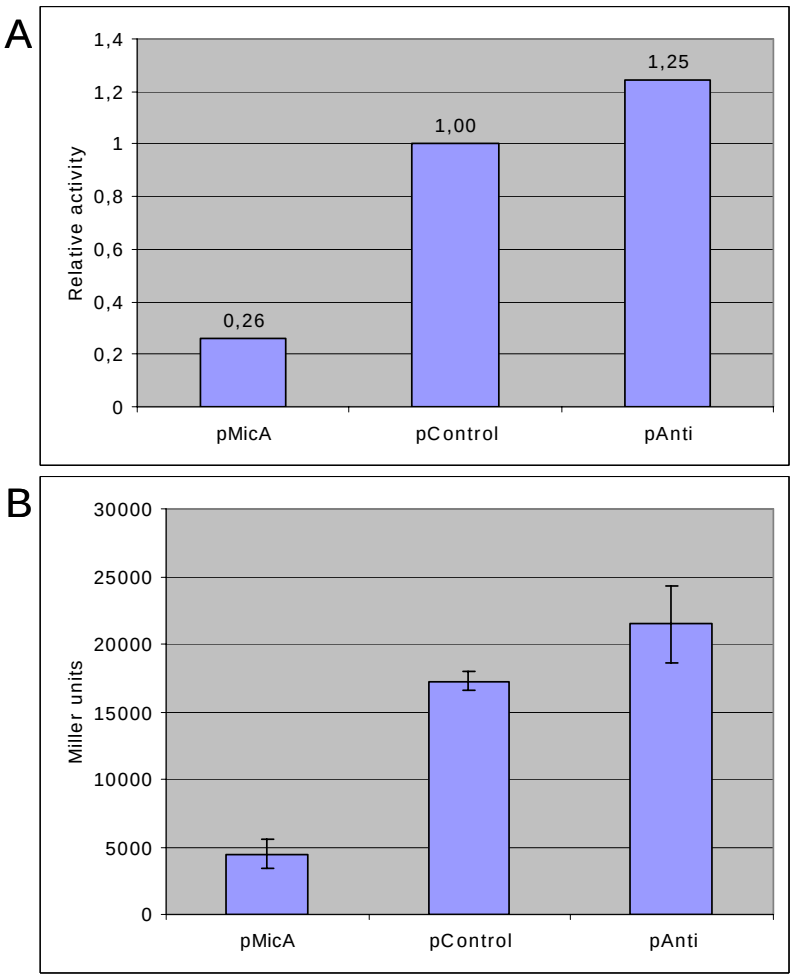

Figure 3. The activity of β-galactosidase encoded by an *ompA-lacZ* translational fusion decreased 4-fold in *E. coli* MC4100 cells over-expressing MicA. pMicA is a MicA over-expressing plasmid, pAnti encodes an antisense RNA complementary to MicA and pControl carries a promoterless *lacZ* fragment. Cells were harvested at OD_{540}=0.5. A) Relative activity compared to cells with pControl. B) Specific activity in Miller units.

4.2 Mutations in the binding region abolish the inhibition of *ompA-lacZ* expression.

As described recently, MicA seems to inhibit *ompA*-mRNA translation by acting as an *antisense* RNA [16]. To rigorously test the proposed base-pairing interaction and its effect on regulation, mutations were introduced in the proposed binding region. pMicAM4 and pMicAM6 carry 4 and 6 nucleotide changes, respectively, in the binding region of MicA. pOmpAM6 carries six nucleotide changes which are complementary to the changes in pMicAM6 and should therefore restore the putative antisense binding (Fig. 4). As seen in Figure 5, the β-galactosidase activity increases in strains carrying pEH2 and pMicAM4, or pMicAM6, as compared to strains providing wild-type MicA. Loss of down-regulation is also obtained in the strain carrying pMicA and pOmpAM6. The strain containing pMicAM6 and pOmpAM6 shows a 4- to 5-fold lower β-galactosidase activity compared to the strain with pOmpAM6 and pControl. This indicates that inhibition was restored by the compensatory mutations. These results strongly suggest that MicA is a true antisense RNA targeting *ompA*-mRNA. It also shows that the proposed binding region indeed is where the interaction occurs *in vivo*.

Figure 4. Mutations introduced in the proposed MicA-*ompA*-mRNA binding region. Changed nucleotides are indicated in red and the *ompA*-mRNA start codon is indicated in blue.

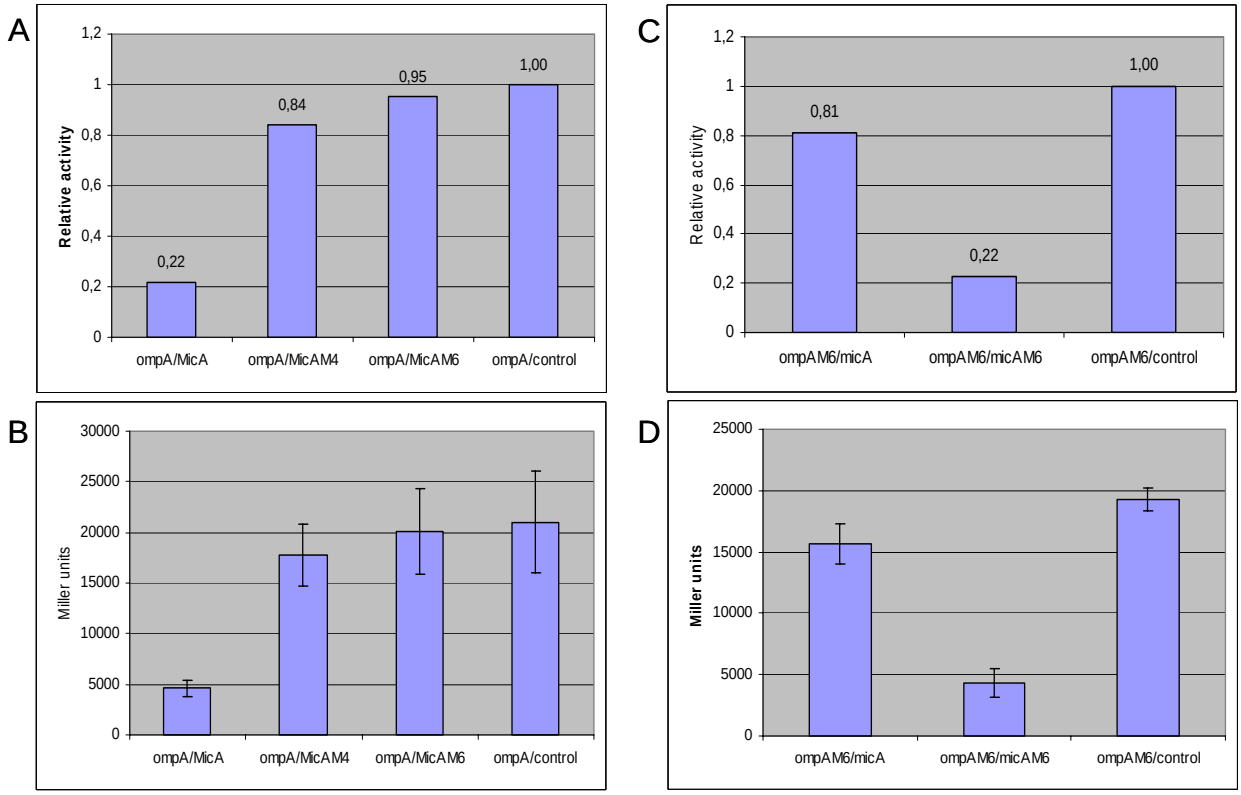

Figure 5. β-galactosidase activity from cells with mutations in the MicA-*ompA*-mRNA binding region of plasmids pEH2 and pMicA. See Figure 4 for details on introduced mutations. When mutations are introduced in the MicA (ompA/MicAM4, ompA/MicAM6) or ompA-mRNA (OmpAM6/MicA) binding regions, the β-galactosidase activity increases compared to cells where both *micA* and *ompA* are wild-type (ompA/micA). The decrease in β-galactosidase activity seen with wild-type *micA* and *ompA* (ompA/micA)is restored when compensatory mutations are introduced (ompAM6/micAM6). Relative activity compared to cells with plasmid pControl are shown in a) and c). Specific activity in Miller units are shown in b) and d).

4.3 Antisense RNA binding specificity *in vitro.*

The specificity of binding between an antisense RNA and its cognate target mRNA is essential for the regulation. Therefore, to qualitatively monitor the specificity of the *in vitro* binding of MicA, MicC and MicF to their respective cognate target mRNAs *ompA*, *ompC* and *ompF*, a gel shift experiment was carried out. Each sRNA was 5'-end-[^{32}P]-labeled and incubated with one of the three unlabeled mRNAs separately. Formed complexes were monitored as slower migrating bands on a native polyacrylamide gel. As seen in Figure 6, MicA, MicC and MicF form complexes with their cognate target mRNAs; *ompA*, *ompC* and *ompF*, respectively. The non-cognate RNA pairs do not form complexes except for one case: MicC forms a complex with *ompA*-mRNA. However, all the labeled MicC forms complex with *ompC*-mRNA whereas only about half of the labeled MicC forms a complex with *ompA*-mRNA. Thus, the three sRNAs indeed have a specific affinity for their cognate mRNA targets *in vitro* as concluded by this qualitative comparison.

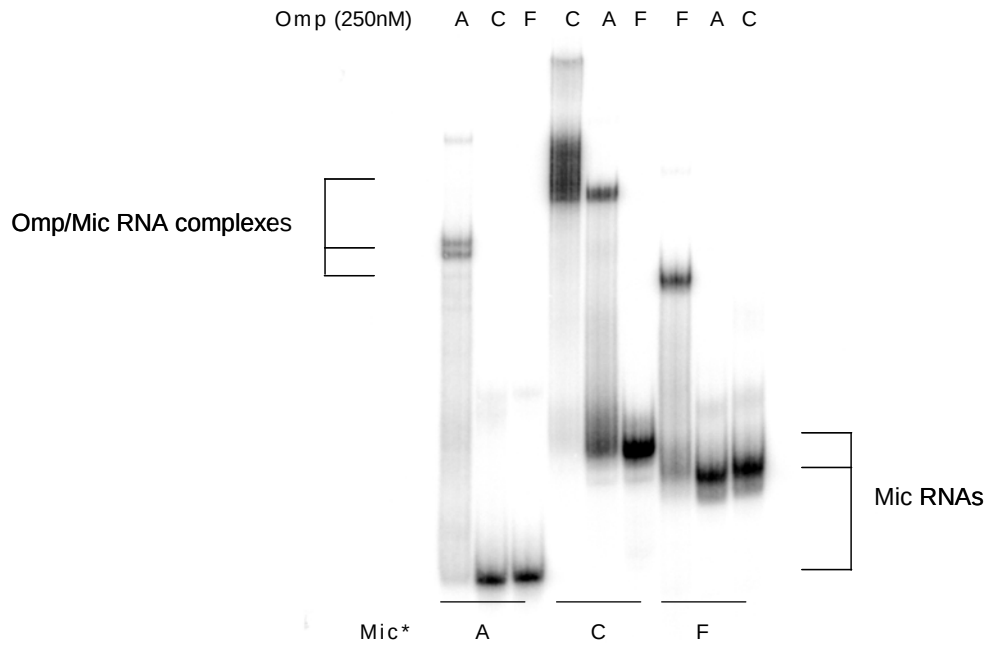

Figure 6. Gel shift experiment where 5'-end-[^{32}P]-labeled MicA, MicC and MicF was incubated with unlabeled *ompA*, *ompC* and *ompF* mRNAs in excess followed by electrophoresis on a native polyacrylamide gel. Slower migration bands show sRNA/mRNA complexes. Every sRNA show higher binding affinity for the cognate mRNA than for the non-cognate mRNAs.

4.4 MicC 5'end binds immediately upstream of the ribosome binding site of *ompC*-mRNA.

As described above, the interacting region of an sRNA is often weakly structured and often binds close to, or overlaps, the ribosome binding site of the mRNA. To delineate the binding regions of MicC and *ompC*-mRNA, structural probing was performed by using RNase T1, RNase T2 or lead(II) acetate which all specifically cleave single stranded RNA. RNase T1 specifically cleaves 3' of G residues, RNase T2 and lead(II) cleaves after all four residues. The protection from cleavage of 5'-end-labeled MicC due to the double stranded region obtained through the MicC-*ompC* binding is shown as a "footprint" in Figure 7. The binding region maps from the MicC 5'-end to the uracil residue at position 16. In predictions of the binding region by BLASTN, an additional binding region was proposed, starting from the uracil at position 25 and ending at the uracil at position 30 [7]. This region was not protected in the experiment shown here. No protection is seen when MicC is incubated with *ompA*-mRNA. Furthermore, as seen in the lanes without addition of mRNAs, the binding region is easily cleaved, indicating that this region is weakly structured and thereby accessible for RNA-RNA interaction. Figure 8 shows the reverse experiment, unlabeled MicC incubated with 5'-end-labeled *ompC*-mRNA. The "footprint"

obtained by MicC binding, lies immediately upstream of the ribosome binding site (RBS) of *ompC*-mRNA. The MicC binding thereby may inhibit initiation of *ompC*-mRNA translation.

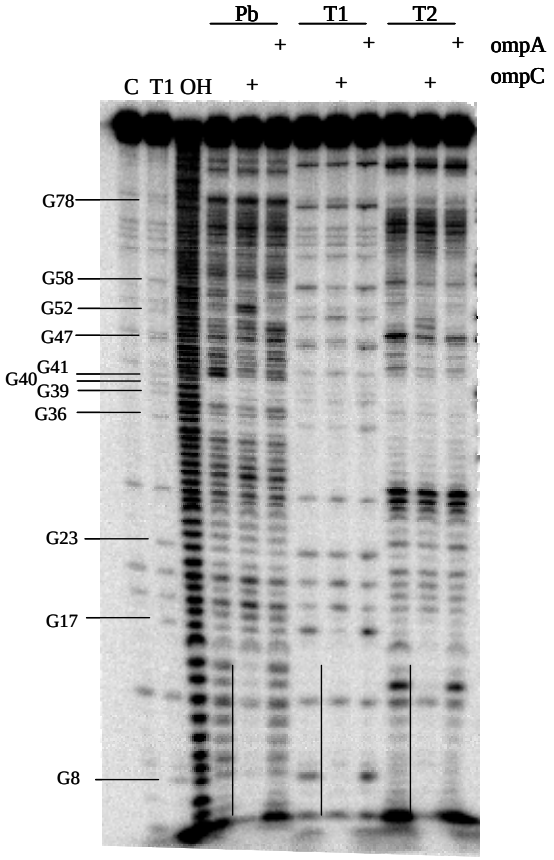

Figure 7. Structural probing of MicC RNA with lead(II)acetate, RNase T1 and RNase T2, conducted on 5'-end-labeled MicC RNA. + indicates the presence of of unlabeled *ompA*-mRNA or *ompC*-mRNA. C shows the mock-treated control, T1 shows RNase T1 cleavage under denaturing conditions and OH shows an alkaline ladder. The lines show the protection from cleavage due to *ompC*-mRNA binding to MicC.

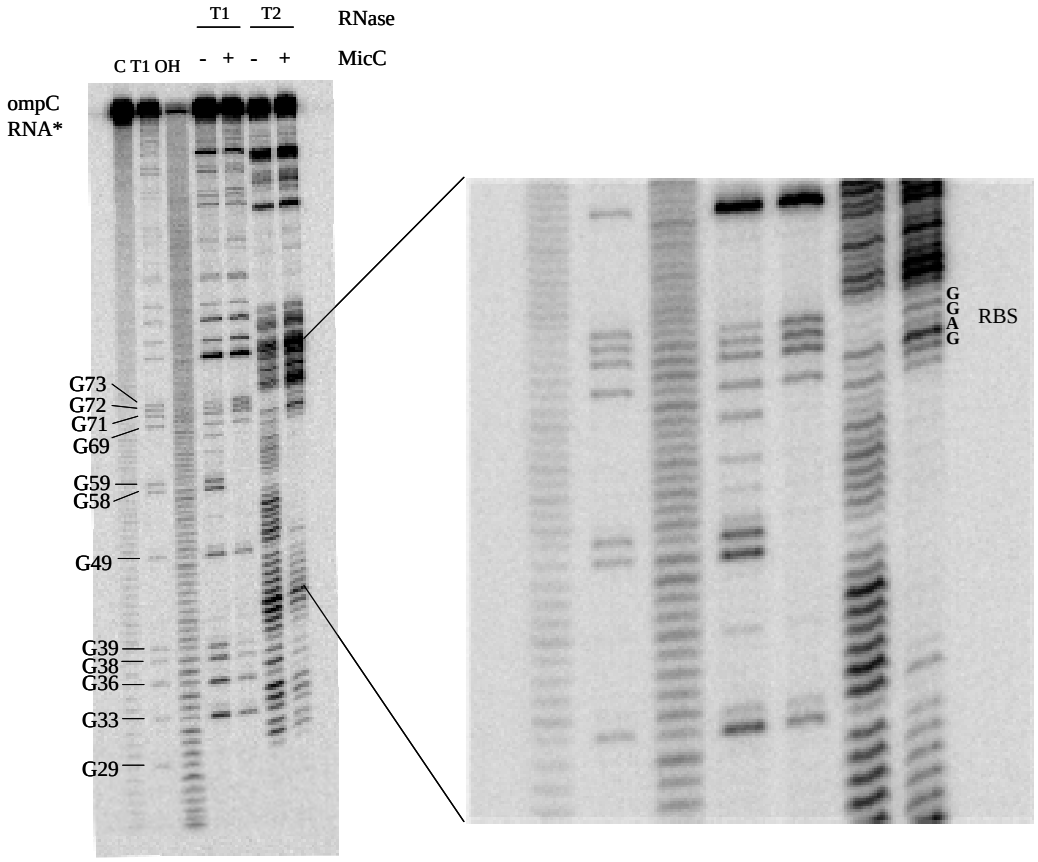

Figure 8. Structural probing with RNase T1 and RNase T2, conducted on 5'-end-labeled *ompC*-mRNA. + indicates the presence of unlabeled MicC RNA. C shows the mock-treated control, T1 shows RNase T1 cleavage under denaturing conditions and OH shows an alkaline ladder. The magnification shows the protection from cleavage due to binding of MicC. The *ompC*-mRNA ribosome binding site (RBS) is indicated.

4.5 Accessibility and conservation of binding regions make *lrp*-mRNA a good putative target for MicA and MicF.

As mentioned above, bioinformatics can be used to predict putative target mRNAs for sRNAs. In an in-house bioinformatic search for targets of the sRNA MicF, the already verified *ompF*-mRNA showed up at first place as the top score candidate [12]. At third place the *lrp*-mRNA showed up, which encodes the leucine responsive regulator protein (Lrp), known to be one of the global transcription regulators in *E. coli*. In a search for targets for MicA, *lrp*-mRNA also showed up as a top ten candidate. How do the properties of MicA-*lrp*-mRNA and MicF-*lrp*-mRNA agree with known sRNA-mRNA couples? As outlined in Figure 9, the predicted binding regions of MicA and MicF are weakly structured. This indicates that the binding regions may be accessible. The binding regions of the sRNAs are also well conserved among related species (Figure 10), indicating that these regions may be under a selective pressure. The binding regions of *lrp*-mRNA do not cover the TIR, but cover the start codon (Figure 10). A binding to the start codon may inhibit initiation of translation. The binding regions of *lrp*-mRNA are also conserved among related species (Figure 10). Since these indications are in line with properties of known RNAs, an experimental test of *lrp*-mRNA as a target for MicA and MicF was carried out.

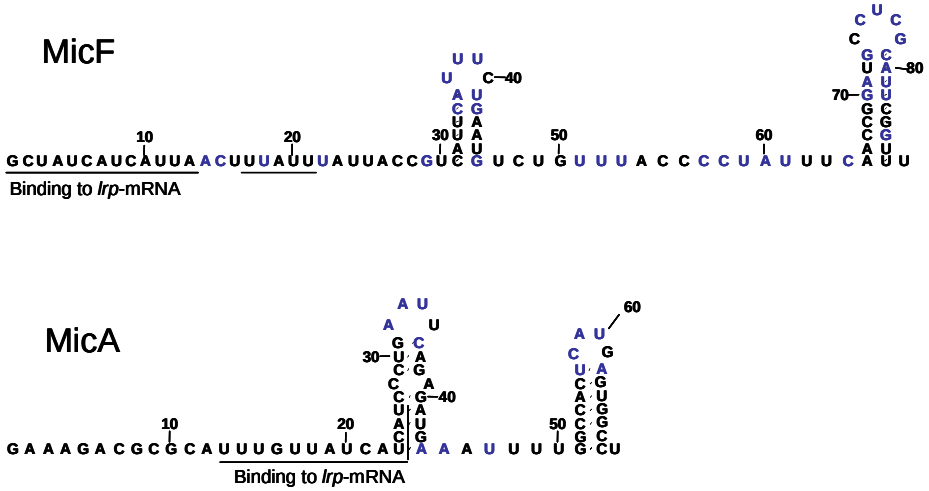

Figure 9. Secondary structures of *E. coli* MicA and MicF. Nucleotides in blue indicate mutational changes from the E. coli sequence shown here compared to *Shigella*, *Salmonella* and *Yersinia*. Note that few mutations are found in the regions predicted to bind to *lrp*-mRNA. The secondary structures are constructed from structural probing results and computational prediction (MicF; [18], MicA; [16])

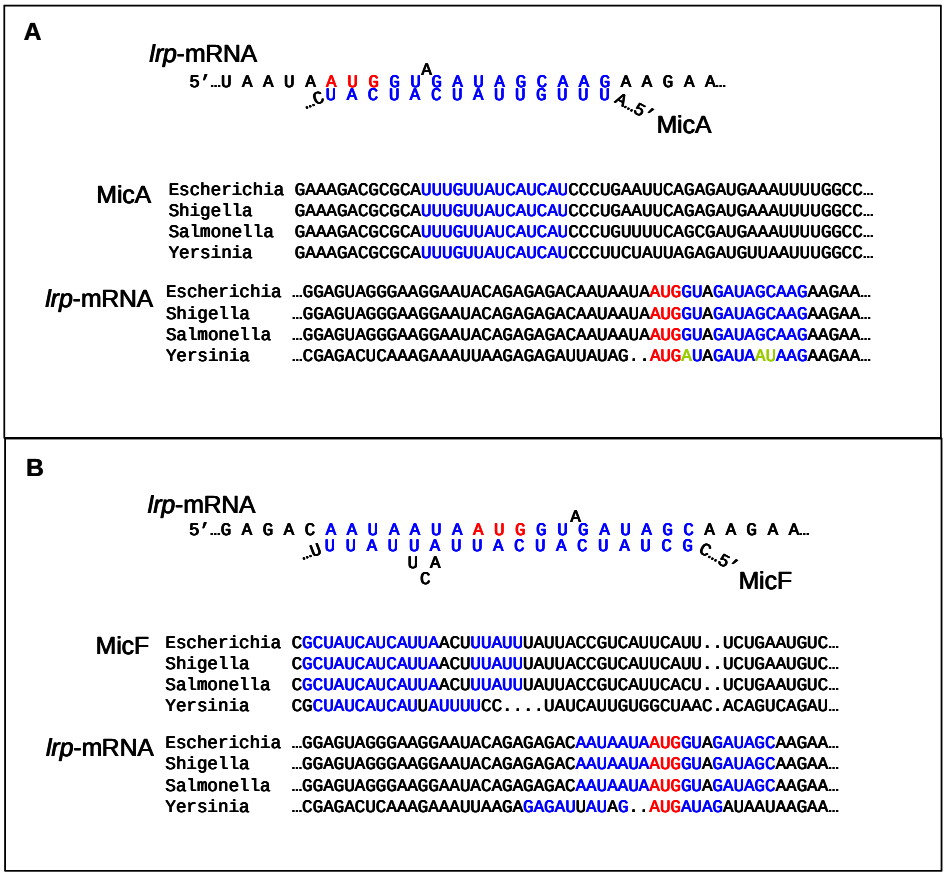

Figure 10. RNA alignments of *lrp*-mRNA and a) MicA, b) MicF. Predicted binding regions are shown in blue, the AUG start codon of *lrp*-mRNA is shown in red.

4.6 The expression of an *lrp-lacZ* translational fusion is inhibited by MicF.

To experimentally test Lrp as a putative target for MicA and MicF, a similar two plasmid system as for the MicA / *ompA*-mRNA experiments was used. The leader and the beginning of the coding region of *lrp* was fused in frame with *lacZ* on plasmid pMC874. Post-transcriptional regulation through interactions with the 5' untranslated region of *lrp*-mRNA 5' would thereby affect expression of β-galactosidase. The resulting fusion plasmid pEH5 and one of the three plasmids pMicA, pMicF and pControl was transformed to *E. coli* MC4100 competent cells. The β-galactosidase activity assay was then performed as described for OmpA. As seen in Figure 11, the β-galactosidase activity decreased 4-fold in cells over-expressing MicF, and 2-fold in cells over-expressing MicA. These results indicate that *lrp*-mRNA may be a target at least for MicF. It also shows that the bioinformatic search may serve as a powerful tool for prediction of sRNA targets.

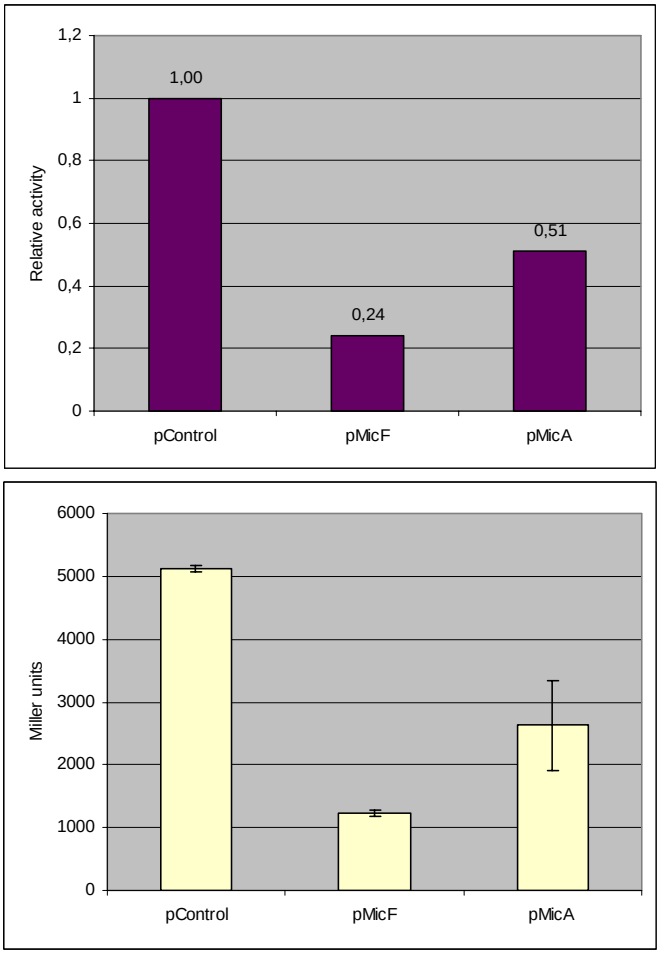

Figure 11. β-galactosidase activity decreases 5-fold due to MicF over-expression, and 2-fold due to MicA over-expression. All strains carry plasmid pEH5 encoding an *lrp-lacZ* translational fusion, plus one of the three plasmids pMicF, pMicA or pControl as indicated in the figure. Relative activities are normalized against the strain carrying pControl. Specific activity is given in Miller units.

5 DISCUSSION

In this project, post-transcriptional regulation through antisense-acting sRNAs in *E. coli* has been studied. To monitor this regulation *in vivo*, a reporter gene system was set up and tested. This system was then used to test bioinformatically predicted sRNA targets *in vivo*. To study the antisense binding specificity and localization, *in vitro* techniques were used.

As described recently, translation of *ompA*-mRNA decreases when MicA RNA expression increases [16]. In this project these observations have been confirmed and also further investigated. An *ompA*-*lacZ* translational fusion was used to study the MicA regulation of OmpA expression *in vivo*, by measuring the β-galactosidase activity. As expected, the β-galactosidase activity was lower in cells containing the fusion plasmid and a MicA over-expression plasmid, compared to cells with the fusion plasmid and control plasmid. We also wanted to see if the inhibition occurs through MicA acting as a true antisense RNA targeting *ompA*-mRNA. When mutations where introduced in the predicted binding region, two observations were made. First, the β-galactosidase activity was higher in cells with mutated MicA and wildtype *ompA*-mRNA compared to cells in which both RNAs were wildtype. Lower activity was also obtained with wildtype MicA and mutated *ompA*-mRNA than with wild type RNAs. This indicates that the nucleotide changes weaken the MicA-*ompA*-mRNA binding and thereby weaken the MicA inhibition of OmpA translation. Second, the decrease in β-galactosidase activity observed with wildtype RNAs was restored in cells where both RNAs were mutated so that the binding region was restored. These two observations strongly indicate that MicA is an antisense RNA targeting the *ompA*-mRNA *in vivo*.

To qualitatively study the specificity of binding between Mic RNAs and *omp*-mRNAs, gel shift experiments were carried out. As seen in Figure 6, Mic RNAs bind specifically to their cognate *omp*-mRNAs. One exception in the experiment is the binding of MicC to *ompA*-mRNA. However, compared to the cognate MicC-*ompC* complex, the MicC-*ompA* binding is less specific since only half of the added MicC does form a complex with *ompA*-mRNA while all MicC forms a complex with *ompC*-mRNA. Although this complex was formed in this *in vitro* experiment, earlier studies have shown that OmpA levels *in vivo* not are affected by over-expression of plasmid-borne *micC* [7]. Additionally, no binding of MicC to *ompA*-mRNA was seen in the structural probing experiment (Figure 8). The binding of MicC to *ompA*-mRNA may be too weak to promote protection the RNAs from cleavage by lead(II), RNase T1 or RNase T2.

The region by which MicC binds to *ompC*-mRNA has been predicted by a BLASTN search [7]. The predicted region consists of two stretches, nucleotides 1 to 16 and 25 to 30 counting from the MicC 5'-end. Here, enzymatic and chemical cleavage *in vitro* of radioactively 5'-end-labeled MicC incubated with *ompC*-mRNA has been used to map the interaction. As seen in Figure 7, the longer stretch from nucleotide 1 to 16 is well protected through binding of the two RNAs. On the other hand, the shorter stretch from nucleotide 25 to 30 is not protected. This indicates that this second stretch does not take part of the binding, or that binding is weak. The reverse experiment, probing of 5'-end-labeled *ompC*-mRNA incubated with unlabeled MicC reveals the binding region of the mRNA. As seen in Figure 8, the 16 nucleotide binding site lies immediately upstream the RBS. This region is indeed complementary to nucleotides 1 to 16 of the MicC 5'-

end. This confirms the earlier prediction. Through a primer extension inhibition assay it has been shown that MicC inhibits 30S ribosome from binding to *ompC*-mRNA [7].

To test the ability to predict sRNA targets by bioinformatic means, *lrp*-mRNA was picked out as a putative target for two sRNAs, MicA and MicF. Lrp was chosen not only because it showed up as a good candidate in the computationally created lists. As shown in Figure 9 and Figure 10, the MicA-*lrp*-mRNA and MicF-*lrp*-mRNA couples share several properties with known sRNA-mRNA couples. The binding regions are conserved, the sRNA binding regions are accessible, and the binding regions of *lrp*-mRNA cover the start codon. These findings make *lrp*-mRNA a good target for MicA and MicF, at least theoretically.

What biological connections are there between Lrp and MicA/MicF? Lrp has earlier been reported to negatively regulate transcription of *ompC* and *micF* [8]. This means that when Lrp is highly expressed, OmpC and MicF are down-regulated, which also implies that OmpF is expressed. As described above, OmpC and OmpF are expressed in opposite environmental conditions, so when OmpF is down-regulated by MicF, OmpC is expressed. As seen in Figure 11, a 4-fold repression of Lrp expression was observed when MicF was provided on an over-expression plasmid. Since Lrp is thought to be a transcriptional repressor of *ompC*, expression of OmpC indicates that Lrp levels are down. An inhibition of Lrp expression by MicF thus fits in the model, assuming that the up-regulation of OmpC is partly because of the absence of repression by Lrp. As seen in Figure 11, MicA over-expression inhibited Lrp expression 2-fold. In-house experiments have shown that OmpC levels increase when MicA is over-expressed. MicA may therefore (as well as MicF) affect OmpC expression positively through negative regulation of Lrp.

These speculations of course have to be thoroughly confirmed before they become more than just speculations. The down-regulation of Lrp shown in these experiments is due to over-expression of the sRNAs. This probably means that the inhibition is less at single-copy levels of the sRNA. How large the real inhibition is and what physiological effects this implies is an exciting question for future studies.

All sRNAs studied so far are non-essential under normal growth conditions. What biological importance do they then possess? Since a majority of the sRNAs is involved in stress responses they may increase the fitness of the bacteria under these conditions significantly. Why use an RNA as a regulator instead of a protein? Again the answer may relate to stress responses. RNA is both faster and energetically cheaper to produce. When the environment changes rapidly the sRNAs may be the fast response to the transition followed by the slower response in the form of protein regulators. Another interesting role may be the shut down of one gene in an operon as in the case of the sRNAs RyhB and Spot 42.

Since the study of sRNAs have been carried out mainly in *E. coli* and its close relatives, little is known about this phenomena in other bacteria. However, it seems unlikely that sRNAs are a special feature of enterobacteria. Future studies will probably reveal sRNAs in many bacterial species.

This project is an example of how the different properties of sRNAs can be studied *in vivo* and *in vitro*. Since the characterization of sRNAs is still in its infancy, there is an urgent need for

methods that are appropriate and effective. Bioinformatic searches combined with reporter gene technology seem to be a fruitful combination, although this has to be tested at a larger scale.

6 ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor Professor Gerhart Wagner for giving me the opportunity to carry out my degree project in his group. I am also very grateful to Doctor Fabien Darfeuille for giving me never ending supervision in the lab. His great knowledge about science and his practical skills have been invaluable for the project and for future work. For additional help with the project I would like to thank: Klas Udekwu, Ulrika Lustig, Johan Reimegård, Jaydip Ghosh, Cecilia Unoson and Ema Kikovska. Many thanks go to Fredrik Söderbom for reading and reviewing my report. Finally, thank you everyone at the Microbiology Program and ICM. I have had a great time!

7 REFERENCES

1. Gottesman S. (2004). The small RNA regulators of *Escherichia coli*: Roles and mechanisms. *Annu. Rev. Microbiol.* 58:303-328.
2. Wagner EGH, Vogel J. (2003). Noncoding RNAs Encoded by Bacterial Chromosomes. *Noncoding RNAs: molecular Biology and Molecular Medicine*. Edited by Erdman VA. Copyright 2003 Eurekah.com.
3. Altuvia S, Zhang A, Argaman L, Tiwari A, Storz G. (1998). The Escherichia coli OxyS regulatory RNA represses fhlA translation by blocking ribosome binding. *EMBO J.* 17(20):6069-6075.
4. Masse E, Gottesman S. (2002). A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *PNAS* 99(7):4620-4625.
5. Vecerek B, Moll I, Afonyushkin T, Kaberdin V, Blasi 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(3):897-909.
6. Liu MY, Gui G, Wei B, Preston JF 3rd, Oakford L, Yuksel U, Giedroc DP, Romeo T. (1997). The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli. *J Biol Chem*. 272(28):17502-17510.
7. Chen S, Zhang A, Blyn LB, Storz G. (2004). MicC, a second small-RNA regulator of Omp protein expression in *Escherichia coli*. *J. Bacteriol.* 186(20):6689-6697.
8. Ferrario M, Ernsting BR, Borst DW, Wiese II DE, Blumenthal RM, Matthwes RG. (1994). The leucine-responsive regulatory protein of *Escherichia coli* negatively regulates transcription of *ompC* and *micF* and positively regulates translation of *ompF*. *J. Bacteriol*. 177(1):103-113.
9. Aiba H, Matsuyama S, Mizuno T, Mizushima S. (1987). Function of micF as an antisense RNA in osmoregulatory expression of the ompF gene in Escherichia coli. *J Bacteriol*. 169(7):3007-3012.

10. Andersen J, Delihas N. (1990). *micF* RNA binds to the 5'end of *ompF* mRNA and to a protein from *Escherichia coli*. *Biochemistry*. 29(39);9249-9256.
11. Wagner EGH, Vogel J. (2004). Approaches to Identify Novel Non-messenger RNAs in Bacteria and to investigate their Biological Functions: Functional Analysis of Identified Non-mRNAs. *Handbook of RNA Biochemistry*. Edited by Hartmann RK, Bindereif A, Schön A, Westhof E. Copyright 2005 WISLEY_VCH GmbH & Co. KgaA, Weinheim ISBN: 3-527-30826-1
12. Reimegård J. *et al*. Unpublished results.
13. Casadaban MJ, Chou J, Cohen SN. (1980). In vitro gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. *J Bacteriol*. 143(2):971-980.
14. Bond PJ, Faraldo-Gómez JD, Sansom MSP. (2002). OmpA: A pore or not a pore? Simulation and modeling studies. *Biophysical Journal*. 83;763-775.
15. Gibert I, Barbe J. (1990). Cyclic AMP stimulates transcription of the structural gene of the outer-membrane protein OmpA of Escherichia coli. *FEMS Microbiol Lett.* 56(3):307-311.
16. Udekwu *et al.* Accepted in *Genes & Development* 2005.
17. Vytvytska O, Jakobsen JS, Balcunaite G, Andersen JS, Baccarini M, von Gabain A. (1998). Host factor I, Hfq, binds to Escherichia coli ompA mRNA in a growth rate-dependent fashion and regulates its stability.*PNAS* 95(24):14118-14123.
18. Schmidt M, Zheng P, Delihas N. (1995). Secondary structures of Escherichia coli antisense micF RNA, the 5'-end of the target ompF mRNA, and the RNA/RNA duplex. *Biochemistry.* 34(11):3621-3631.