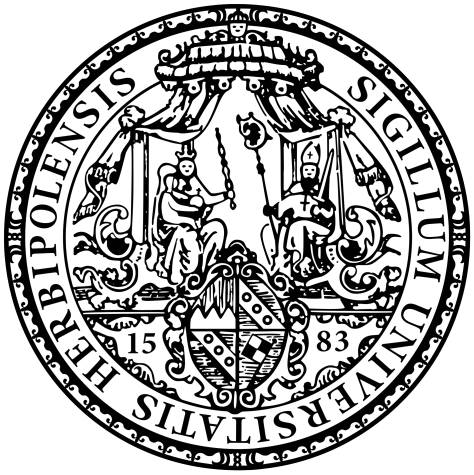

Global mapping of RNA-RNA interactions in *Salmonella* via RIL-seq

Globale Analyse der RNA-RNA-Interaktionen in *Salmonella* mittels RIL-seq

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Infection & Immunity

submitted by

Gianluca Matera

from Matera

Würzburg 2021

Submitted on:

Members of the Thesis Committee

Chairperson: Prof. Dr. Christoph Sotriffer

Primary Supervisor: Prof. Dr. Jörg Vogel

Supervisor (Second): Dr. Franziska Faber

Supervisor (Third): Prof. Dr. Kai Papenfort

Date of Public Defence:

Date of Receipt of Certificates:

Summary

RNA represents one of the most abundant macromolecules in both eukaryotic and prokaryotic cells. Since the discovery that RNA could play important gene regulatory functions in the physiology of a cell, small regulatory RNAs (sRNAs) have been at the center of molecular biology studies. Functional sRNAs can be independently transcribed or derived from processing of mRNAs and other non-coding regions and they often associate with RNA-binding proteins (RBPs). Ever since the two major bacterial RBPs, Hfq and ProQ, were identified, the way we approach the identification and characterization of sRNAs has drastically changed. Initially, a single sRNA was annotated and its function studied with the use of low-throughput biochemical techniques. However, the development of RNA-seq techniques over the last decades allowed for a broader identification of sRNAs and their functions. The process of studying a sRNA mainly focuses on the characterization of its interacting RNA partner(s) and the consequences of this binding. By using RNA interaction by ligation and sequencing (RIL-seq), the present thesis aimed at a high-throughput mapping of the Hfq-mediated RNA-RNA network in the major human pathogen *Salmonella enterica*.

RIL-seq was at first performed in early stationary phase growing bacteria, which enabled the identification of ~1,800 unique interactions. In-depth analysis of such complex network was performed with the aid of a newly implemented RIL-seq browser. The interactome revealed known and new interactions involving sRNAs and genes part of the envelope regulon. A deeper investigation led to the identification of a new RNA sponge of the MicF sRNA, namely OppX, involved in establishing a cross-talk between the permeability at the outer membrane and the transport capacity at the periplasm and the inner membrane. Additionally, RIL-seq was applied to *Salmonella enterica* grown in SPI-2 medium, a condition that mimicks the intracellular lifestyle of this pathogen, and finally extended to *in vivo* conditions during macrophage infection. Collectively, the results obtained in the present thesis helped unveiling the complexity of such RNA networks.

This work set the basis for the discovery of new mechanisms of RNA-based regulation, for the identification of a new physiological role of RNA sponges and finally provided the first resource of RNA interactions during infection conditions in a major human pathogen.

Zusammenfassung

RNA ist eines der am häufigsten vorkommenden Makromoleküle sowohl in eukaryontischen als auch in prokaryontischen Zellen. Seit der Entdeckung, dass RNA wichtige genregulatorische Funktionen in der Physiologie einer Zelle spielen könnte, stehen kleine regulatorische RNAs (sRNAs) im Mittelpunkt molekularbiologischer Studien. Funktionelle sRNAs können alleinstehend von nicht-codierenden oder codierenden Bereichen des Genoms transkribiert werden, aber sie können auch durch die Prozessierung einer mRNA entstehen. Des Weiteren sind sRNAs häufig mit RNA-bindenden Proteinen (RBPs) assoziiert. Seitdem die beiden wichtigsten bakteriellen RBPs, Hfq und ProQ, identifiziert wurden, hat sich die Art und Weise, wie wir an die Identifizierung und Charakterisierung von sRNAs herangehen, drastisch verändert. Ursprünglich wurden sRNAs annotiert und anschließend für einzelne sRNAs die Funktion mit biochemischen Techniken untersucht. Die Entwicklung von RNA-seq-Techniken in den letzten Jahrzehnten ermöglichte nun jedoch eine globale Identifizierung von sRNAs und ihren Funktionen. Der Prozess der Untersuchung einer sRNA konzentriert sich hauptsächlich auf die Charakterisierung ihrer interagierenden RNA-Partner und die Folgen dieser Bindung. Mit Hilfe der RNA-Interaktion durch Ligation und Sequenzierung (RIL-seq) wurde in der vorliegenden Arbeit eine Hochdurchsatzkartierung des Hfq-vermittelten RNA-RNA-Netzwerks in dem wichtigen humanen Krankheitserreger Salmonella enterica durchgeführt.

RIL-seq wurde zunächst in Bakterien in der frühen stationären Wachstumsphase durchgeführt, was die Identifizierung von ~1.800 einzigartigen Interaktionen ermöglichte. Mit Hilfe eines neu implementierten RIL-seq-Browsers wurde daraufhin eine eingehende Analyse dieses komplexen Netzwerks durchgeführt. Das Interaktom enthüllte bekannte und neue Interaktionen zwischen sRNAs und mRNAs, die Teil des Zellwand-Regulons sind. Eine tiefergehende Untersuchung führte zur Identifizierung eines neuen RNA-Schwammes, OppX, welcher mit der sRNA MicF bindet und so die Herstellung eines Cross-Talks

zwischen der Permeabilität an der äußeren Membran und der Transportkapazität am Periplasma und der inneren Membran ermöglicht. Darüber hinaus wurde RIL-seq für Salmonella enterica angewandt, welche in SPI-2-Medium gewachsen waren, wobei diese Bedingung, die den intrazellulären Lebensstil dieses Erregers nachahmt. Durch die Infektion von Makrophagen mit dem Bakterium, wurde das RIL-seq Protokoll des Weiteren unter in vivo Bedingungen getestet. Insgesamt trugen die in dieser Arbeit erzielten Ergebnisse dazu bei, die Komplexität solcher RNA-Netzwerke zu enthüllen. Diese Arbeit bildete die Grundlage für die Entdeckung neuer Mechanismen der RNA-basierten Regulierung als auch für die Identifizierung einer neuen physiologischen Rolle von RNA-Schwämmen und lieferte letztendlich die erste Untersuchung für RNA-Interaktionen unter Infektionsbedingungen in einem wichtigen menschlichen Krankheitserreger.

Contents

List of Figures

List of Tables

Chapter 1

1 Introduction

1.1 Post-transcriptional control in bacteria

RNA molecules and molecular interactions involving them represent a vital feature for living organisms. RNA is a major component of ribosomes and builds up 80-90% of the bacterial transcriptome. Together with transfer RNAs (tRNAs), the respective messenger RNA (mRNA) and accessory proteins, ribosomes are able to actively support protein synthesis processes in a cell. Until a few decades ago, these RNA molecules were the only ones thought to exist in any type of cells, prokaryotic or eukaryotic. While elucidating the functioning of the *lac* operon in *E. coli* in the early 60s, Jacob and Monod proposed the existence of a "*lac* repressor" that would have the same chemical features of an RNA molecule and would be able to inhibit mRNA translation or transcription by interactions with the DNA locus or with the mRNA itself (Jacob and Monod, 1961). Despite being eventually wrong, as the *lac* repressor turned out to be a protein factor, these scientists inadvertently set the basis for what has nowadays become a whole research area. Bacterial small regulatory RNAs (sRNAs) have acquired over the decades a central role in orchestrating post-transcriptional control.

A common mechanism in RNA-based regulation is that a regulatory RNA acts by base-pairing with its target mRNAs, leading to alteration in translation and/or mRNA stability. Unlike eukaryotic microRNAs (miRNAs), sRNAs have a broader length, usually from 50 to 300 nucleotides (nt) (Gorski et al., 2017). With the exception of antisense RNAs (asRNAs) which are encoded on the opposite strand of their target RNAs (*cis*-encoded) and therefore base-pair with full complementarity, most sRNAs in Gram-negatives are *trans*-encoded. These regulatory RNAs base-pair with only short and limited complementarity and impact their target mRNAs by annealing with the help of RNA chaperones, e.g. Hfq or ProQ (Hör et al., 2020; Storz et al., 2011). While the very first sRNAs (MicF in *Escherichia coli*

and RNAIII in *Staphylococcus aureus*) were discovered in the early 80s, the impact and extent of their abundance in bacterial species have been only established in the last decade thanks to the introduction of RNA sequencing techniques (Hör et al., 2018). It is by now clear that *trans*-encoded sRNAs can be produced through several mechanisms: i) independent transcription from a dedicated transcription start site (TSS) (a.k.a. intergenic sRNAs); ii) derived from inside the coding sequence (CDS) of a certain gene (intragenic); iii) UTR-derived, which includes sRNAs that stem from the 5′ or the 3′ ends of an mRNA (Figure 1.1) (Hör et al., 2020).

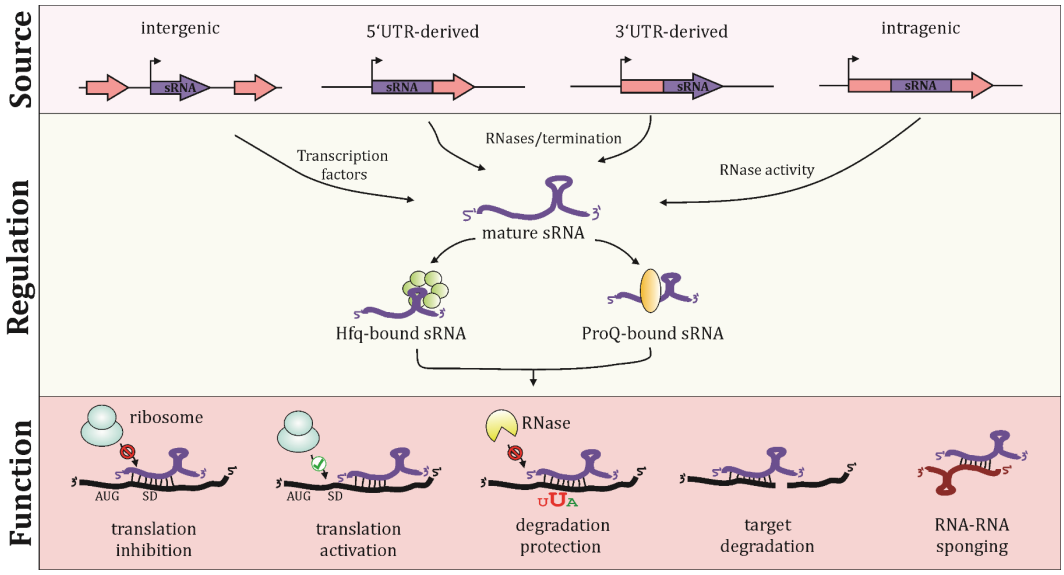

Figure 1.1 Sources of small RNAs in bacteria, their maturation and functions. Generally, bacterial sRNAs can be generated as independent transcripts from intergenic regions, or derived from parts of mRNAs (UTRs or coding sequences). Their maturation often relies on the activity of RNases and/or RBPs. Mature sRNAs exert their functions by inhibiting/activating their target mRNAs at the level of translation, stability or by inactivating other regulatory RNAs (RNA sponges). Adapted from Hör*, Matera* *et al.* 2020.

1.2 sRNAs of Gram-negatives are bound to RNA chaperones

As mentioned above, *trans*-encoded sRNAs (herein after "sRNAs") in Gram-negative bacteria usually function through interaction with proteins known as "RNA chaperones". The first protein of this class to be identified was Hfq (host factor for bacteriophage Qβ RNA replication), a homo-hexamer donut-

shaped protein belonging to the large eukaryotic Sm-like protein family (Møller et al., 2002). The Hfq hexamers contain several positively charged binding surfaces, providing them with the capacity to bind negatively charged nucleic acids (Schumacher et al., 2002). These are the proximal surface, engaged in interactions with the stretch of U residues, typical of bacterial Rho-independent terminators; the distal face, responsible for binding to ARN sequence repeats (where R stands for adenine or guanine and N any nucleotide); and the rim face, which binds to AU-rich sequences. Hfq-dependent sRNAs are all bound to the proximal face as they all feature a Rho-independent terminator with the typical U-stretch. Additionally, sRNAs can bind to the rim or to the distal face according to their nucleotide composition, which assigns them as Class I or Class II Hfq-dependent sRNAs, respectively (Schu et al., 2015; Zhang et al., 2013). It has been demonstrated that the positively charged residues (e.g. Arginine) on the Hfq rim face promotes annealing of the sRNA to its mRNA target (Panja et al., 2013, 2015).

A second class of RBP, recently established as RNA chaperone in bacteria, is the FinO domain-containing protein faimily. The best example of such a protein class is the one encoded by the *proQ* gene. The binding characteristics of ProQ are still not completely understood, although it has been shown that the monomer of this RBP tightly binds to double-stranded RNA structures, with very limited sequence preference towards A-stretch regions of the RNA (Bauriedl et al., 2020; Smirnov et al., 2016; Stein et al., 2020). Unlike Hfq, little is known about how ProQ performs post-transcriptional regulation in Gram-negative bacteria. It has been reported that this RNA chaperone binds RNAs (sRNAs and mRNAs) and stabilizes their half-lives through protection from RNase degradation (Holmqvist et al., 2018; Melamed et al., 2020). A few studies have suggested that, like Hfq, ProQ is able to promote duplex formation between a sRNA and its target mRNA in a classical RNA-RNA interaction-dependent manner (Melamed et al., 2020; Smirnov et al., 2017) (Figure 1.1).

In the last decades, *E. coli* and *Salmonella* have been key model organisms for the in depth characterization of RBP-bound sRNAs. Once transcribed or processed from a pre-mature transcript (Chao and Vogel, 2016; Updegrove et al., 2019), sRNAs can associate to Hfq (or ProQ). It has been shown that for most chaperone-bound sRNAs, the lack of binding to the RBP determines fast degradation of the sRNA and loss of regulatory function (Schu et al., 2015; Zhang et al., 2013). Given the limited copy number of the Hfq protein (30,000 to 60,000 molecules per bacterial cell, (Kajitani et al., 1994)), sRNAs compete for binding to this chaperone (Moon and Gottesman, 2011). Growth conditions as well as stress-response transcription programs guide sRNAs expression in bacterial cells, making the Hfq sRNA interactome a dynamic and constantly re-organizing network.

1.2.1 sRNA-mediated regulation occurs at different levels

Following maturation processes and binding to the RNA chaperone, Hfq-bound sRNAs can exert their function through several mechanisms (Figure 1.1). Generally, the outcome of a sRNA-mediated regulation depends on several factors: i) where the sRNA binds within the target mRNA; ii) the ability of a particular sRNA to directly or indirectly recruit the bacterial degradosome to the targeted site; iii) whether the targeted RNA is a coding mRNA. In summary, sRNAs are able to promote or inhibit translation initiation of an mRNA (Urban and Vogel, 2007), inhibit RNase activity by masking RNase recognition sites (Fröhlich et al., 2013), promote target degradation (Pfeiffer et al., 2009), or engage in RNA sponging mechanisms (Melamed et al., 2016, 2020; Miyakoshi et al., 2015).

The two main RNA elements that dictate sRNA regulation are the presence of stem-loop structures and complementarity regions between the sRNA and the target. Initially investigated for the asRNAs, it is now widely accepted that sRNA-target interaction is a multi-step process that initiates with base-pair formation involving a few high-affinity nucleotides, followed by structural rearrangements and extension of the base-pairing duplex (Brantl, 2007; Storz et al., 2011; Wagner et al., 2002). As other bacterial genes,

sRNAs are often conserved among closely related bacterial species. Interestingly, evolution preserves the function of such sRNAs by maintaining what is normally referred to as "seed sequence" of the sRNA, corresponding to the region engaged in base-pairing with the targets. For Hfq-dependent sRNAs, the rho-independent terminator at the 3'-end is often conserved too, as it is required for stable binding to the RNA chaperone.

1.2.2 State of the art methods to identify sRNAs

Historically, multiple methods have been employed to systematically identify sRNAs in bacteria and miRNAs in eukaryotes. In bacteria, global searches were based on the initial assumption that sRNAs have defined features: i) they are transcribed from intergenic regions (IGRs), which are parts of the genome that do not encode for mRNAs; ii) they are short in length (50-300 nts); iii) their transcription terminates through a Rho-independent terminator. The outstanding sequence conservation of sRNAs among bacterial species greatly expedited their annotation. These initial screens led to a first list of potential sRNAs and set the basis for a broader "sRNA search and characterization" (Argaman et al., 2001; Chen et al., 2002; Wassarman et al., 2001). Additional strategies to discover sRNAs employed the creation of cDNA libraries obtained by selecting RNAs based on their small size. These studies provided a first evidence that sRNAs can also originate from untranslated regions (UTRs) and not only from IGRs (Kawano, 2005; Vogel, 2003). The development of RNA-seq technology made sRNAs identification faster and more comprehensive. Particularly, the possibility to obtain a detailed annotation of transcription start sites (TSS) through differential RNA-seq (dRNA-seq) in certain species led to the identification of many more non-coding transcripts than appreciated before, discovering new sRNAs arising from the most diverse genomic locations (Kroger et al., 2012; Sharma et al., 2010; Thomason et al., 2015). Once a new putative regulatory RNA is identified, a deep functional characterization is needed to both confirm that the transcript encodes for a functional RNA and

to assign the biological role in the cell in terms of post-transcriptional regulation.

More recently, techniques that rely on immunoprecipitation (IP) of a known sRNA-binding protein have been widely employed (e.g. RIP-seq, CLIP-seq) in sRNA discovery studies. These methods (detailed in 1.2.2.1) contributed, particularly in *Salmonella* and *E. coli*, to the identification of a large number of 3'UTR-derived Hfq-dependent sRNAs (Chao and Vogel, 2016; Holmqvist et al., 2016; Miyakoshi et al., 2019; Wang et al., 2019), as well as to the discovery of novel RBPs, such as ProQ (Smirnov et al., 2016). A big advantage of these approaches over traditional sRNA-discovery methods is the ability to identify defined common features for the RNA binding properties of a specific protein. For instance, AUU-reach regions and U-stretches in the transcription terminator, together with stem-loop structures, precisely define an Hfq-dependent sRNA (Holmqvist et al., 2016). On the other hand, the ProQ protein seems to have high affinity to highly structured RNAs, and only very limited sequence specificity (Gonzalez et al., 2017; Smirnov et al., 2016; Stein et al., 2020).

In the following paragraphs, two of the most widely used RBP-centric techniques for sRNA identification will be discussed in detail.

1.2.2.1 RIP-seq

Techniques that employ the use of antibodies for protein co-immunoprecipitation (Co-IP) have been used for several decades. These biochemical methods complemented the more traditional genetic screens such as those performed with yeast two-hybrid assays (Hope and Struhl, 1986) to analyse protein complexes and protein-protein interactions. As anticipated previously, cellular complexes often involve interactions of RBPs with RNA molecules. Technical innovation permitting high-throughput analysis of the transcriptome (e.g. Chip-microarrays and RNA-seq) allowed techniques as Co-IP to be employed in the identification of the RNA ligands of a known RBP. Some of these methods include RIP-Chip (Keene et al., 2006; Sittka et al., 2008) and RIP-seq (Figure 1.2A, (Chao et al., 2012)). Both rely on

on-beads Co-IP of a tagged RBP followed by analyses of the bound RNAs on a microarray chip (RIP-Chip) or through RNA sequencing technology (RIP-seq). These approaches provide information on transcripts that co-purify with the RBP, albeit with poor genomic resolution. In fact, while sRNAs are fairly short transcripts, mRNA sizes can vary and reach several thousands of nt in length. For this reason, both RIP-based approaches failed to identify precise binding sites of an RBP. Knowing the RNA binding preferences of a protein provides insights of their cellular functions, apart from being a proxy for predicting the targets of similar proteins. Additionally, in RIP protocols very mild washing steps are performed to preserve the interaction between the RBP and their RNA ligands. This results in the detection of a greater amount of false positive interactions.

1.2.2.2 CLIP-seq

A RIP-based method with the addition of a UV-crosslinking step is named CLIP ((Ascano et al., 2012; Darnell, 2010); sometimes also named CRAC in eukaryotes (Granneman et al., 2009)). While several CLIP-like methods differ in the type of UV-crosslinking or in the strategies for cDNA library preparation, they all rely on the formation of a UV-dependent covalent link between the protein moiety and the RNA (Figure 1.2B). When compared to RIP-seq, this additional step provides several improvements. Not only does it enable a more stringent protein Co-IP under denaturing condition, but it also allows for a higher resolution when identifying RRE (RNA recognition elements). In fact, the necessary RT (reverse transcription) step for converting the short RNAs to cDNA before sequencing introduces mutations at the cross-linked nucleotides, therefore helping to pinpoint the exact binding site of the RBP (König et al., 2012). In bacteria, CLIP-seq has been successfully applied to several RBPs in very distant organisms. In *Salmonella* it has allowed the identification of the RNA ligands for Hfq, CsrA (Holmqvist et al., 2016) and ProQ (Holmqvist et al., 2018). In the human pathogen *Neisseria meningitidis,* CLIP-seq revealed the targetome of the minimal FinO-domain protein, ProQ (Bauriedl et al., 2020).

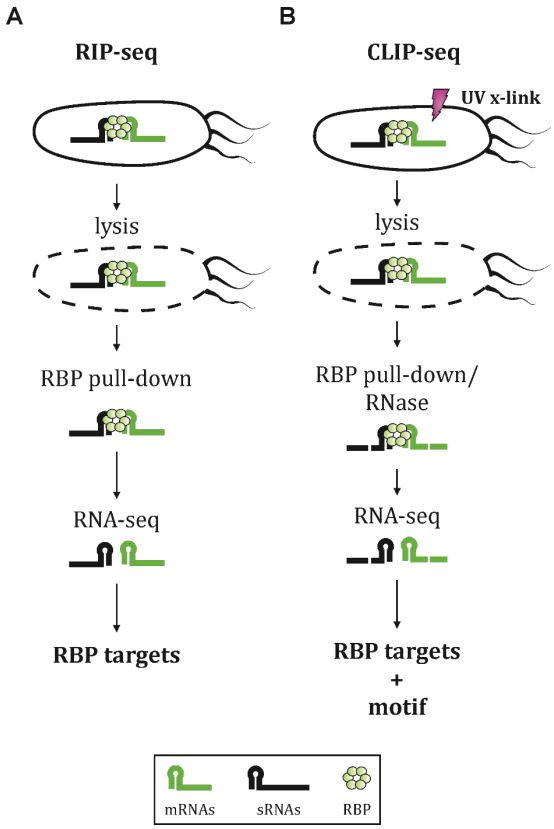

Figure 1.2 RBP-centric methods to identify sRNAs.
(**A**) RNA immunoprecipitation procedure (RIP-seq). Following lysing of the bacteria, the tagged protein of interest (RBP) is purified together with its bound RNA ligands. RNA extraction is performed and RNAs are sequenced. The protocol uses non-denaturing conditions to preserve the RBP-RNA interactions. (**B**) As in (**A**), but cells are UV-cross linked before being lysed. The Co-IP and washes steps are performed under denaturing conditions as the RNA ligands are covalently bound to the RBP. This method allows, apart from the identification of RNA targets, of specific sequence motif to which the RBP preferentially binds.

1.2.3 Traditional methods to identify sRNA targets and limitations

Although methods like RIP-seq and CLIP-seq allow for the identification of sRNAs bound to an RBP under specific conditions, identifying the RNA targets of a given sRNA remains a challenge.

One of the first established methods to characterise mRNAs targeted by a sRNA was by analysing the transcriptome (e.g. microarrays or RNA-seq later on) of a WT (wild type) strain of a bacterium of interest compared to a mutant carrying the "sRNA-of-interest" deletion allele (a.k.a. differential gene expression analysis; (Papenfort et al., 2008, 2009)). These

approaches, flanked by pulse expression of a sRNA followed by transcriptomic (Fröhlich et al., 2016; Massé et al., 2005; Papenfort et al., 2006), rely on changes occurring at the RNA level of the targeted mRNA (usually negative regulation). Although these methods proved quite useful for the discovery of the top targets of many sRNAs, in particular Hfq-dependent sRNAs of *E. coli* and *Salmonella*, they came along with disadvantages. Firstly, as just mentioned, de-regulation at the RNA level of the target mRNA must take place in order to detect differences in a deletion or pulse-expression experiment of a certain sRNA. For instance, sRNAs that regulate their targets only by translation control without inducing degradation could not be identified. Secondly, by affecting the gene expression of a multitude of genes, these approaches suffer from off-target effects. Nevertheless, this could be partially overcome by performing the same experiment with a "seed" mutant of the sRNA as in this scenario only specific targets would be affected by its expression (Sharma et al., 2011).

While research is constantly growing in the bioinformatics field for *in silico* predicting sRNA targets through computing their sequences with *ad hoc* algorithms (e.g., RNAhybrid, CopraRNA, IntaRNA, reviewed in (Pain et al., 2015)), *in vivo* techniques, aimed at a comprehensive identification of sRNA targets, would result in a better understanding of the complex network of RNA-RNA interactions in living organisms.

1.2.3.1 Methods to identify the targetome of a single sRNA: MAPS & GRIL-seq

When only a single sRNA is of interest, identifying with high confidence its target spectrum is nowadays fairly simple. Most current studies combine old (described in 1.2.3) and new techniques to address this question. The "old" usually refers to methods aimed at identifying de-regulated targets of the sRNA, while "new" methods rely on pull-down and/or proximity RNA ligation reaction to "physically" purify RNAs bound to the sRNA.

One way to purify RNAs (or proteins) bound to a sRNA is by tagging the RNA of interest with an aptamer (e.g., MS2, (Said et al., 2009)). The

widely used method MAPS (MS2 affinity purification coupled with RNA sequencing, Figure 1.3A) relies on this principle. In this scenario, the sRNA of interest is 5'- or 3'-tagged with the bacteriophage MS2 stem loops and expressed from a plasmid carrying an arabinose-inducible promoter (Lalaouna et al., 2015). Upon induction of the tagged sRNA, the cell lysate is ran through an affinity column whereby the MS2 coat protein is immobilized. Eventually, the tagged sRNA together with its RNA targets will be co-purified by elution from the column with maltose. Eluted and purified RNAs are then subjected to RNA sequencing.

Recently, a method with similar scopes has been developed in the bacterium *Pseudomonas aeruginosa,* namely GRIL-seq (Global small non-coding RNA target identification by ligation and sequencing; (Han et al., 2016)). In this specific case, bacteria carry an arabinose-inducible plasmid expressing the native sequence of the sRNA and a second plasmid carrying an IPTG-inducible gene encoding for the T4 RNA ligase I enzyme (Figure 1.3B). *In vivo* ligated interacting RNAs (chimeras) are subsequently purified through RNA pulldown with an on-beads-bound oligonucleotide carrying a complementary sequence to the sRNA of interest. Upon RNA sequencing, reads mapping to two different loci in the *Pseudomonas* genome are considered chimeric reads and therefore possibly resulting from an RNA-RNA interaction *in vivo*.

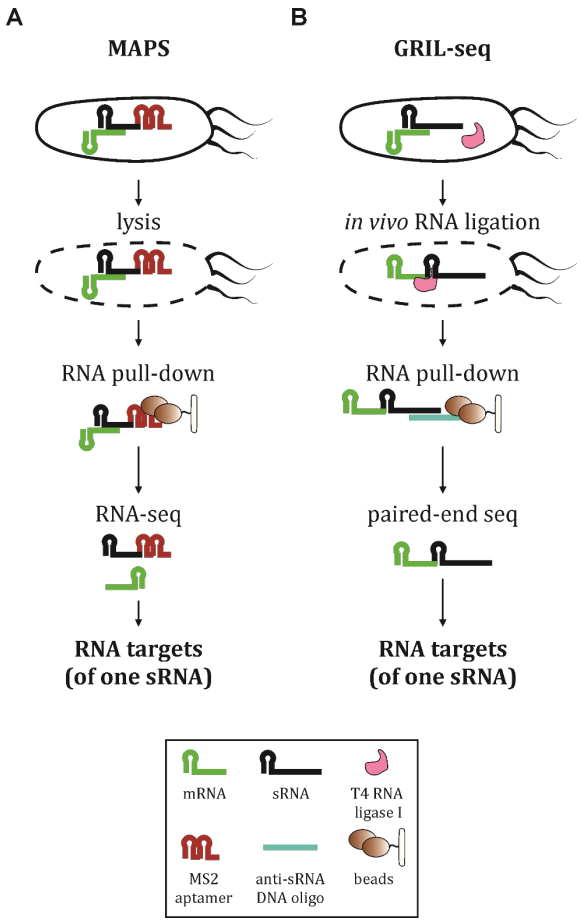

Figure 1.3 sRNA-centric methods to identify sRNA targets.
(**A**) MS2 pulldown of a sRNA followed by sequencing procedure (MAPS). Following lysis of a strain carry an MS2-tagged sRNA of interest (SOI), RNA complexes are purified with an MBP-bound resin. Elution of the RNAs is followed by RNA sequencing to identify the ligand of the SOI. (**B**) Bacteria carrying a plasmid-driven expression of the T4 ligase I and a plasmid expressing the SOI are lysed after growth. Chimeric ligated RNA pairs are purified with an oligo carrying the reverse-complement of the SOI sequence. Purified RNAs are subjected to paired-end sequencing to identify RNA hybrids (GRIL-seq).

Both MAPS and GRIL-seq proved successful in complementing pulse-expression RNA-seq dataset and expanding the knowledge on sRNA's biology. For instance, a MAPS study on the PinT sRNA in *Salmonella* (Correia Santos et al., 2021) highlighted new targets that could not be previously identified (Westermann et al., 2016).

1.2.4 Advanced methods to globally identify sRNA targets

However, most bacteria express several hundreds of sRNAs depending on the growth condition. Identifying the targetome of each sRNA with individual studies appears to be an arduous mission. This challenge translates into the need for developing techniques that aim at identifying all

sRNA targets on a global scale and in a more unbiased fashion. In the following paragraphs, three of very recently developed methods will be described and discussed: Hi-GRIL-seq (1.2.4.1), CLASH (1.2.4.1) and RIL-seq (1.2.4.3 and 1.2.4.5).

1.2.4.1 Hi-GRIL-seq (High-throughput GRIL-seq)

Briefly described in 1.2.3.1, GRIL-seq has been originally developed to capture *in vivo* ligated interacting RNAs of a specific sRNA. The advanced Hi-GRIL-seq (Figure 1.4A) simply omits the step of specific-sRNA-enrichment, in order to purify the total RNA of *Pseudomonas* and therefore capturing the whole pool of *in vivo* ligated RNAs (Zhang et al., 2017). With the possibility of adopting deeper RNA-seq technologies and the strategy of depleting the dominant rRNAs from the RNA pool, Hi-GRIL-seq has proved successful in capturing RNA pairs that do not depend on the function of an RBP.

1.2.4.2 CLASH (cross-linking, ligation, and sequencing of hybrids)

Originally developed in eukaryotic cells to map RNAs bound to hAGO1 in the RISC complex (Helwak et al., 2013) or to identify snoRNA (small nucleolar RNA)-rRNA interactions in the snoRNP complex (Kudla et al., 2011), CLASH allows the mapping of base-pairing RNAs whereby the interaction is mediated by an RNA chaperone (Figure 1.4B). Technically, this method results from combining the UV crosslinking principles of CLIP-seq and CRAC (described in 1.2.2.2) to capture stable ribonucleoprotein complexes *in vivo* together with the activity of a T4 RNA ligase I enzyme to allow hybrids formation upon complex isolation. CLASH has been recently extended to the bacterial field in an attempt to identify hybrids bound to the major endonuclease, RNase E, of the pathogenic *E. coli* (EHEC) (Waters et al., 2017). This study confirmed the hypothesis that many bacterial sRNAs recruit RNase E while binding to their target RNAs (Pfeiffer et al., 2009). Additionally, the same approach has been used to further characterize the targetome of Hfq in non-pathogenic *E. coli* K-12 strain (Iosub et al., 2020). An

important signature of the CLASH protocol is the high stringency purification step the complexes undergo before RNA ligation and hybrids isolation. As a consequence, the sequencing reads from chimeric RNA fragments result in not more that 1% of the total sequenced RNAs. While this feature ensures the detection of very abundant, stable and non-spurious interactions, it also denies the ability to identify RNAs that base-pair only transiently, with a lesser affinity or generally low abundant RNAs.

1.2.4.3 RIL-seq (RNA interaction by ligation and sequencing)

The development of RIL-seq occurred concomitantly to that of CLASH and Hi-GRIL-seq. As for the previously discussed approaches, RIL-seq aims at capturing a snapshot of the *in vivo* RNA interactome mediated by an RBP, first and foremost Hfq (Melamed et al., 2016). The concept of RIL-seq is based on a few crucial steps (Figure 1.4C):

- Co-IP of the chromosomally-tagged RBP of choice.
- RNase A/T1 treatment of the RNP complexes to shorten the RNA fragments.
- Polynucleotide Kinase (PNK) treatment to prepare the RNA ends for RNA ligation.
- On-beads proximity RNA ligation with the T4 RNA Ligase I enzyme.
- RNA isolation, cDNA library and paired-end sequencing of the single and chimeric fragments.

While the similarities with the CLASH approach are pronounced, a few differences can be highlighted. For instance, every step in RIL-seq (from Co-IP to RNA isolation) is performed in non-denaturing conditions (Melamed et al., 2018). Hence, given the lower stringency in detection of chimeras, the downstream steps require a more elaborate process of statistical analysis and interpretation of the data. Firstly, a Fisher's exact test is performed to preserve only chimeric fragments significantly enriched in the experiment (detailed in 4.5.1). Secondly, a threshold of statistically significant chimeras (S-Chimeras) is applied in order to retrieve interactions represented by a high enough number of fragments. The threshold is selected according to

sequencing depth and to the amount of chimeras detected in the control experiment (non–FLAG Hfq). This process, in addition to ensure a more confident repertoire of RNA-RNA interactions, also massively reduces the size of the dataset (from several 100,000 interactions before filtering to 1,000-2,000 after filtering), making it easier to explore.

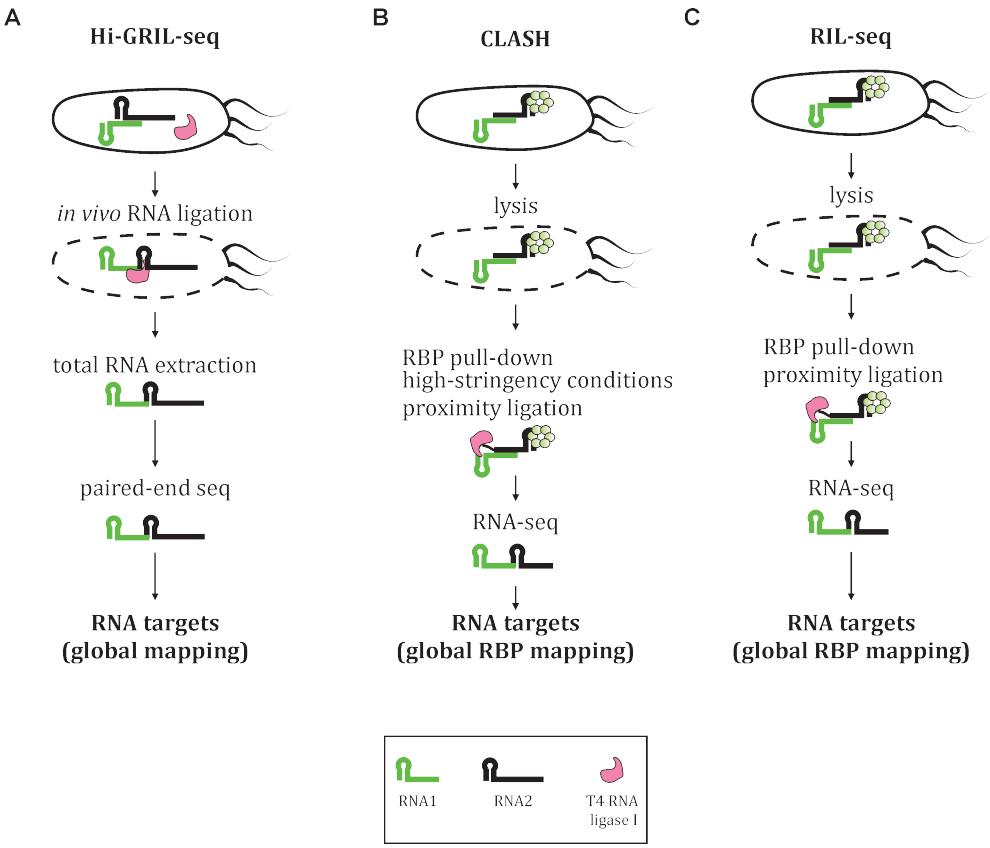

Figure 1.4 Global methods for deciphering RNA-RNA interactions *in vivo*.
(**A**) Procedure as described in Figure 1.3B, without sRNA-enrichment. Ligated RNAs are purified and directly subjected to paired-end sequencing. (**B**) RBP-centric mapping of RNA-RNA interactions (CLASH). UV-cross linked cells carrying a flagged RBP are lysed, followed by IP of the RBP with its ligands in denaturing stringent conditions. *In vitro* proximity ligation is performed, followed by paired-end sequencing of the hybrids. (**C**) Protocol similar to (**B**), where all the steps are performed in non-denaturing conditions.

1.2.4.4 Pros and cons of the discussed methods

In choosing the best method for mapping RNA-RNA interactions, a few considerations have to be made. Selecting whether to undertake an RBP-independent method is of primary importance. GRIL-seq, Hi-GRIL-seq and MAPS are clearly more suitable (and the only available) in cases where no traditional RNA chaperone is known (e.g. as it is the case for most Gram-positive bacteria). GRIL-based approaches require the bacterium of interest to be competent for plasmid uptake and appropriate for plasmid-based expression of the enzyme T4 RNA ligase I. Additionally, both GRIL-seq and MAPS require prior knowledge of the nucleotide sequence of the sRNA of interest.

When an RBP is known in the model organism, approaches such as CLASH or RIL-seq can be applied. While CLASH should be preferred in circumstances whereby priority is given to identifying only the top target interactions, as very stringent purification steps are employed, RIL-seq is more informative in terms of completeness of the interactome. Additionally, given the native conditions preserved throughout the protocol, RIL-seq could potentially be performed omitting the UV cross-linking step, therefore reducing the time and cost of the experiments, beside eliminating possible transcriptional perturbations due to UV-induced stress. On the other hand, the downstream statistics of the sequencing data is more straightforward for CLASH, removing the need for specialised bioinformatics expertise and reducing the overall time of data analysis.

1.2.4.5 Choice of RIL-seq over other global methods

The power and reliability of all the aforementioned methods notwithstanding, RIL-seq was selected as the strategy to enlarge our knowledge on sRNA regulation in *Salmonella.* This choice is based on several considerations: i) the availability of existing CLIP-seq (Holmqvist et al., 2016) and RIP-seq (Chao et al., 2012) datasets performed on Hfq in this bacterium triggered the interest in providing yet a further resource on this

RNA chaperone with respect to RNA-RNA interactions; ii) prior studies of Hfq had extensively shown that the majority of sRNAs in Gram-negative bacteria act via an RBP, making a protein-dependent method the most suitable in this scenario; iii) RIL-seq, as described in 1.2.4.3, preserves native conditions during the purification steps and employs a solid statistic approach, allowing for the identification of less stable or less abundant interactions, therefore a more faithful snapshot of the Hfq-mediated RNA interactome.

1.3 Are there more regulatory RNAs yet to be discovered?

A dozen of studies, published in the last five to six years, took global approaches in an attempt to answer the question: when shall we stop looking for regulatory RNAs? Initially, it was thought sRNAs could only originate from intergenic regions and be independently transcribed (Argaman et al., 2001; Wassarman et al., 2001). This dogma was immediately challenged by further studies that identified functional regulatory RNAs derived from UTRs of coding mRNAs (Chao et al., 2012; Kawano, 2005; Vogel, 2003) or cleaved from tRNA fragments (Lalaouna et al., 2015). More recent studies have shown regulatory RNAs can originate from internal regions of an mRNA coding sequence (Adams et al., 2021). Not only have new regulatory RNAs continuously been discovered, but new ways these sRNAs act as post-transcriptional regulators have been identified. Recently, a large amount of work has been dedicated to the identification of "RNA sponges", i.e., regulatory RNAs that act by sequestering the function of another sRNA (Denham, 2020), as their identification proved challenging without global methods.

Above all, global approaches have taught us that a lot is yet to be discovered in the sRNA world.

1.4 Used model organism

In the present thesis, the bacterial pathogen *Salmonella enterica* serovar Typhimurium was used to investigate the Hfq interactome through RIL-seq and to potentially identify new regulatory RNAs and means of regulation. Most of the presented results refer to a dataset performed during *in vitro* growing conditions. However, RIL-seq was eventually adapted, optimized and performed in *Salmonella*-infected macrophages to generate an *in vivo* map of RNA interactions.

1.4.1 *Salmonella enterica* Serovar Typhimurium

Salmonella enterica is a flagellated, rod-shaped, aerobic, Gram-negative bacterial species that can infect a broad range of vertebrate hosts. Several thousand (over 2,500) serovars (or serotypes) of this species have been identified and classified according to their lipopolysaccharide (LPS) and flagellar antigens. Among these are *Salmonella enterica* Typhi and Paratyphi, responsible for typhoid fever, which is typically characterized by abdominal pain and fever. *Salmonella enterica* Typhimurium and Enteritidis serotypes are classified as non-typhoidal *Salmonella* (NTS) and their pathogenicity can vary according to their host type. They are responsible for self-limiting gastroenteritis in several mammals (humans, poultry, swine and cattle) as well as for systemic infection in immunodepressed humans (LaRock et al., 2015). *Salmonella* infection begins with oral ingestion of contaminated water or food. Bacteria survive the acid environment of the stomach and reach the intestine where they cause inflammatory diarrhea (Harris et al., 1972). Inflammation is a necessary reaction to confer *Salmonella* a growth advantage over the members of the intestinal microbiota (Stecher et al., 2007; Winter et al., 2010).

Once in the intestinal lumen, *Salmonella* is able to invade and survive inside several host cell types, including epithelial cells and macrophages. While in the latter bacteria are naturally engulfed via phagocytosis, *Salmonella* can trigger its own uptake by epithelial cells. Survival inside the

host cells is achieved by remodelling of the bacterial surface upon sensing of the host innate immunity (Dalebroux and Miller, 2014). These events are orchestrated by transcriptional regulators, such as the two-component systems (TCS) PhoPQ, PmrAB, RcsBC and EnvZ-OmpR, but are also highly influenced by two different type III secretion systems (T3SSs) and the effectors they secrete, encoded by the *Salmonella* pathogenicity island 1 (SPI-1) and SPI-2 (Haraga et al., 2008). Analysis of the complete genome sequence of *Salmonella* has suggested these two genomic islands have been independently acquired through horizontal gene transfer (HGT) (McClelland et al., 2001). While the machinery encoded by the SPI-1 T3SS transports effector proteins that promote invasion across the cell membrane, the SPI-2 secretes proteins important for intracellular survival and formation of the *Salmonella*-containing vacuole (SCV). Nevertheless, it has been shown that many SPI-1 effectors are maintained in the host cells and could play a role in survival within the SCV (Brawn et al., 2007; Giacomodonato et al., 2007; Steele-Mortimer et al., 2002).

Salmonella, like many Gram-negative bacteria, possess a well-organized cell envelope composed of an outer membrane (OM) and an inner membrane (IM), separated by the periplasm (Silhavy et al., 2010; Vergalli et al., 2020). The primary function of such a thick shield is to protect bacteria from the surrounding environment, while concomitantly it must ensure sufficient permeability of nutrient molecules. OM porins are the primary players to mediate permeation through the envelope and exert partial solute-specificity due to their charge preferences and pore sizes. Both *E. coli* and *Salmonella* constitutively express three major porins, namely OmpA, OmpC and OmpF, in addition to the *Salmonella*-specific OmpD. Importantly, other specialized OM transporters are induced when specific nutrients are detected in the environment. For example, ChiP is specifically expressed when chitosugars are available in the surroundings (Figueroa-Bossi et al., 2009). However, most nutrient specificity is achieved at the IM due to the presence of a repertoire of selective transporters. The Dpp (dipeptides) and Opp (oligopeptides) systems represent the major peptide importers. Precise

regulation of envelope proteins is also a matter of cellular economy, as many of the OM and IM transporters are extraordinarily abundant, with porins reaching numbers as high as ribosomal proteins (10^5 copies/cell, (Li et al., 2014; Vergalli et al., 2020)). Regulation of the properties of the OM or the IM is achieved by several transcription factors and TCSs (Barchinger and Ades, 2013; Dalebroux and Miller, 2014; Egger et al., 1997; Grabowicz and Silhavy, 2017; Konovalova et al., 2016; Raivio, 2014). However, our understanding of how cross-communication between the two membrane compartments is accomplished is still in its infancy.

All known OM porins in *Salmonella* and *E. coli* are post-transcriptionally regulated by sRNAs that act by fine-tuning the expression of such bulky proteins in response to specific signals. Some of these respond to the broad envelope-stress regulator σ^E (Fröhlich and Gottesman, 2018; Klein and Raina, 2017). Additionally, many characterized sRNAs have been reported to regulate inner membrane proteins (IMPs), to either restrict the import of toxic sugars (Kawamoto et al., 2005; Papenfort et al., 2013; Rice and Vanderpool, 2011) or to generally buffer IM stress (Chao and Vogel, 2016).

While many examples of sRNAs targeting different envelope compartments exist, e.g., MicA and RybB (Gogol et al., 2011) and CpxQ (Chao and Vogel, 2016; Grabowicz et al., 2016), little is known about whether sRNAs can be the regulatory link between OM and IM and their cross-talk. Global RNA interactome studies promise a more comprehensive view of sRNA activity with regards to envelope homeostasis in bacteria (Hör et al., 2018).

1.5 Aims of this thesis

The importance of RNA regulation in bacteria, as well as in all other living organisms, is widely recognized. Several global approaches have been used in the last decade to decipher the role of miRNAs in eukaryotes and sRNAs in bacteria and their cognate targets. However, in bacteria, these methods have been prominently applied to the model organism *E. coli,* leaving a gap in the understanding of bacterial pathogens. This doctoral thesis aimed to establish and improve RIL-seq in the bacterium *Salmonella enterica,* providing a valuable resource for this pathogen as well as discovering new ways by which RNA regulation is performed in this pathogen. The steps undertaken can be summarized within the following points:

- Optimization and application of the existing RIL-seq protocol to *Salmonella* during *in vitro* standard growing conditions.
- Analysis of the generated dataset with focus on the characterization of a newly discovered RNA sponge involved in envelope homeostasis.
- Extension of the RIL-seq approach to infection-relevant growing conditions, as well as to *Salmonella*-infected macrophage cell lines.

Chapter 2

2 The Hfq RNA interactome of *Salmonella* revealed by RIL-seq

Large part of the work presented in this chapter has been published in:

- **Matera, G.**, Altuvia, Y., Gerovac, M., El Mouali, Y., Margalit, H., Vogel, J. (2021). Global RNA interactome of *Salmonella* discovers a 5'UTR sponge for the MicF small RNA that connects membrane permeability to transport capacity. Molecular Cell, *accepted*.

Additionally, part of the results obtained in this chapter were performed in collaboration with the following scientists:

- Dr. Yael Altuvia (Hebrew University of Jerusalem, bioinformatics support)
- Prof. Dr. Hanah Margalit (Hebrew University of Jerusalem, bioinformatics support)
- Dr. Milan Gerovac (Helmholtz Institute for RNA-based Infection research, network visualization curator)
- Dr. Youssef El Mouali (Helmholtz Institute for RNA-based Infection research, experimental support with porin mutants)
- Dr. Kotaro Chihara (Helmholtz Institute for RNA-based Infection research, experimental support with size-exclusion chromatography)
- Dr. Elisa Venturini (Institiute for Molecular Infection Biology, University of Würzburg, *in vivo* RIL-seq in 2.3)

2.1 Adapting RIL-seq to *Salmonella*

RIL-seq has been originally designed for investigating the Hfq interactome in *E. coli* (Melamed et al., 2016). Nevertheless, its application to other organisms that contain Hfq or related RBPs should be possible. To generate

a map of Hfq-mediated RNA-RNA interactions in *Salmonella*, the RIL-seq protocol was applied to bacteria growing in LB at early stationary phase (ESP, OD_{600} of 2.0), i.e., when *Salmonella* expresses most of its sRNAs (Chao et al., 2012), with minor modifications (Melamed et al., 2016). Specifically, the composition of the lysis buffer was adapted according to previous immunoprecipitation-based techniques developed for this bacterium of interest, e.g. RIP-seq and CLIP-seq (Chao et al., 2012; Holmqvist et al., 2016). Additionally, the duration of the RNase A/T1 treatment was reduced to 5 min instead of the 7 min used for *E. coli* (detailed in Materials & Methods). RNA-RNA pairs bound to Hfq were recovered through co-immunoprecipitation (Co-IP) in a lysate of an *hfq::*3×FLAG strain, followed by on-beads proximity RNA ligation and RNA sequencing of the chimeric fragments. As a control, the same experiment was simultaneously performed in a strain containing a non-FLAG version of the *hfq* gene. Albeit the sequencing reads of the FLAG vs. non-FLAG strains were not in any way normalized to one another, as established in the original protocol (Melamed et al., 2018), the control experiments served the purpose of setting the S-chimera threshold above which chimeras were considered as true RNA pairs with high confidence. The statistics (Fisher's exact test) performed in the RIL-seq data analysis are extensively described in the Materials & Methods section.

2.1.1 Quality control of the *Salmonella* RIL-seq (before sequencing)

To evaluate whether RIL-seq experiments have been successful, sequencing of the recovered RNA fragments and quantification of the amount and quality of chimeric fragments is a key step. It is nevertheless important to perform pilot experiments prior to sequencing, to assess the quality of the protein pulldown. These pre-sequencing quality control (QC) steps allow for reduction of the costs and time in the event of a failed experiment. For this reason, the samples (both the WT and *hfq::3*×FLAG) of two biological replicates were used for western blotting experiments after the IP step of the RIL-seq protocol. As shown in Figure 2.1 Hfq was successfully recovered (IP

lanes) in the strain carrying a flagged version of the *hfq* gene but not in the WT control strain, when compared to the respective input lanes (LYSATE). The "UNBOUND" lanes represent the pool of Hfq protein that did not successfully bind the beads upon the IP step.

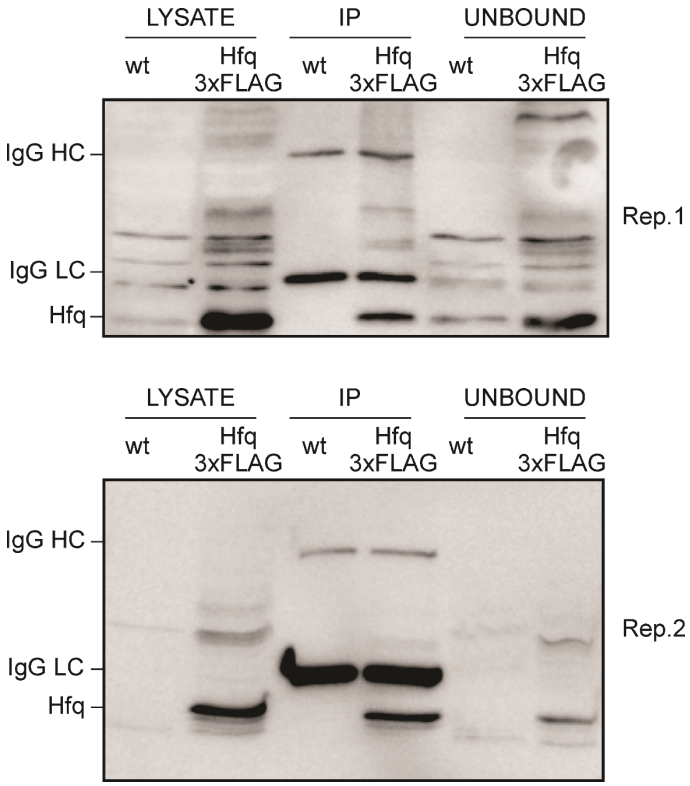

Figure 2.1 Western blot detection of Hfq after RIL-seq IP step.
Two replicates of pilot RIL-seq experiments are shown. For each blot, the LYSATE (input), the IP (eluted fraction) and UNBOUND (lysate fraction after IP), of the WT and the *hfq::3*×FLAG strains, respectively, were loaded onto an SDS-PAGE gel. The "Hfq" label indicates the bands corresponding to the monomer of the Hfq protein. "IgG HC" and "IgG LC" indicate the bands corresponding to heavy and light chains, respectively, of the anti-FLAG antibody used for the IP.

Additionally, following RNA purification after the ligation reaction, the samples of the *hfq::3*×FLAG strain and the respective WT controls were inspected on a bioanalyzer machine to determine both RNA concentration and size distribution. This step allows to assess whether excessive degradation of the RNA samples occurred at any step of the protocol. Furthermore, this QC step is crucial for determining the average length of the recovered RNA fragments, and whether a further RNA fragmentation

step would be necessary prior to cDNA library preparation. As shown in Figure 2.2, the typical profile of a successful RIL-seq shows RNA peaks at around 150-200 nt (Melamed et al., 2018). Importantly, the RNA profiles of the WT samples look similar to the *hfq*::3×FLAG ones, although, as expected, the concentration of RNA measured was 4 to 5 fold lower. In our specific case, no fragmentation was needed as only a very small portion of the RNA pool exceeded 200 nt in length.

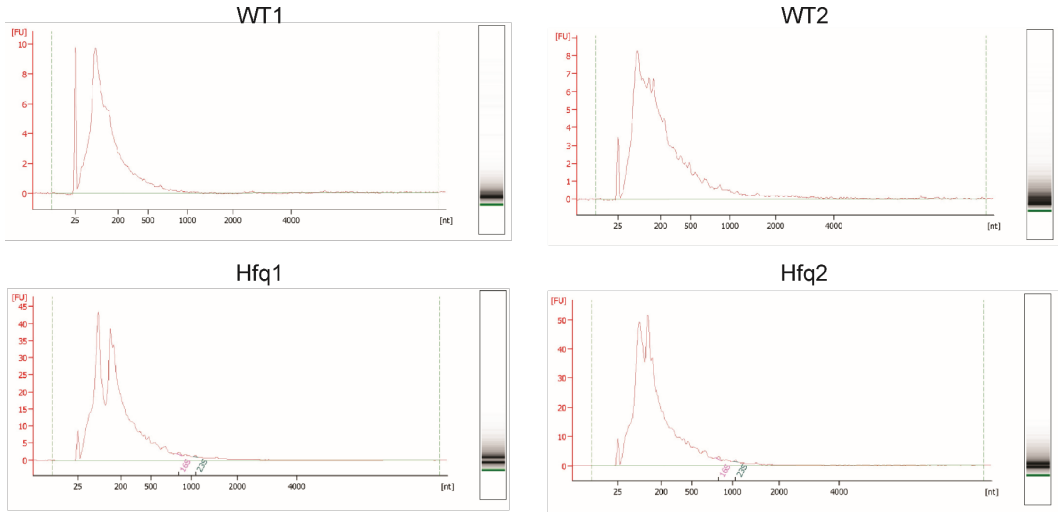

Figure 2.2 Bioanalyzer output of the two *Salmonella* RIL-seq replicates.
The RNA eluted after the last step of the RIL-seq protocol was subjected to an RNA pico Bioanalyzer QC run. The x-axis indicates the size of recovered fragments (expected peaks at 150-200 nts); the y-axis represents the FU (fluorescence units), indicating the intensity of the detected fragments. WT1: replicate 1 for the WT sample; WT2: replicate 2 for the WT sample; HFQ1: replicate 1 for the *hfq*::3×FLAG sample; HFQ2: replicate 2 for the *hfq*::3×FLAG sample.

2.1.2 Paired-end sequencing of RIL-seq libraries

To evaluate whether sequenced transcripts resulted from ligation events of proximal RNAs, the cDNA libraries prepared with the purified RNAs were submitted to paired-end sequencing. With this approach, both ends of the cDNA are read. This strategy allows for independent mapping of the two reads, derived from the two ends of the same cDNA molecule to the *Salmonella* genome. The procedure provides information on whether the cDNA fragment is chimeric or derived from a single RNA (detailed in Materials & Methods). Two independent experiments were performed and

sequenced at an average sequencing depth of ~30-40 million reads per sample (Table S1A). Of all the sequenced fragments for each library (two libraries for the WT and two libraries for the *hfq*::3×FLAG) an average of 25 million reads could be mapped to the *Salmonella* genome either as single or chimeric fragments. Specifically, for the two *hfq*::3×FLAG samples, H1 and H2, we recovered a total of 725,280 and 768,162 chimeras, respectively. The chimeric fragments represented ~3-4% of the total sequenced fragments per library. Whereas, for the two WT samples, W1 and W2, we only detected 3,060 and 1,053 chimeras, respectively (~0.05% of the total sequenced fragments per library). These statistics are a further proof that the experiments were successful as chimeras were almost exclusively detected in the Hfq samples.

The fragments successfully mapped (total, single and chimerics) were further analysed with respect to RNA type. For the first scope, a *Salmonella* annotation (manually improved for UTRs and sRNAs) was used (Table S4). The transcripts recovered with Hfq showed an abundance of sRNAs and mRNAs over the WT control sample, for both single and chimeric fragments (Figure 2.3A). To filter out spurious ligation events, a p value ≤0.05 as determined with a Fisher's exact test (S-chimeras, Statistically significant chimeras) was set as a requirement for all the chimeras. These S-chimeras were further filtered to include only RNA pairs represented by at least n=30 chimeric fragments. These stringent requirements ensure a low number of false-positives and therefore enrich for chimeras representing an *in vivo*-captured interaction. As expected, only the libraries generated from the *hfq*::3×FLAG samples yielded a substantial number of S-chimeras (Figure 6.1A, Table S1A). Most of them were formed by ligation of mRNA (specifically, CDSs and 5'UTRs) and sRNA fragments, with the 5'UTRs mostly represented by the first part (RNA1) and the sRNAs by the second part (RNA2) of the chimeric fragment (Figure 6.1B-C). The differential positioning of certain RNA types on either the first or second part of the chimera can be explained by the molecular binding features of Hfq to sRNAs. In fact, the 3' end of sRNAs is fully wrapped around the rim surface

of Hfq making it poorly accessible for a ligation reaction (Dimastrogiovanni et al., 2014). Additionally, this positional bias has the potential to be used as a proxy to discover previously not annotated sRNAs (Bar et al., 2021).

Salmonella RIL-seq, as did *E. coli* RIL-seq (Melamed et al., 2016, 2020), also recovered a great number of interactions between non-coding regions (e.g. sRNA-sRNA, sRNA-IGR, etc.), supporting the idea that the Hfq network extends far beyond classic mRNA regulation (Hör et al., 2020).

Collectively, our RIL-seq data detected 66 sRNAs within S-chimeras, 25 of which were known Hfq-dependent sRNAs (Hör et al., 2020). The relative proportion of abundance in S-chimeras of these sRNAs is shown in Figure 6.1D. In the growth conditions used to perform RIL-seq, the Hfq-mediated network is dominated by a handful of sRNAs, and many of these are known regulators of envelope proteins (Figure 2.3B and Figure 6.1D) (Chao and Vogel, 2016; Johansen et al., 2008; Papenfort et al., 2008; Pfeiffer et al., 2007; Yin et al., 2019).

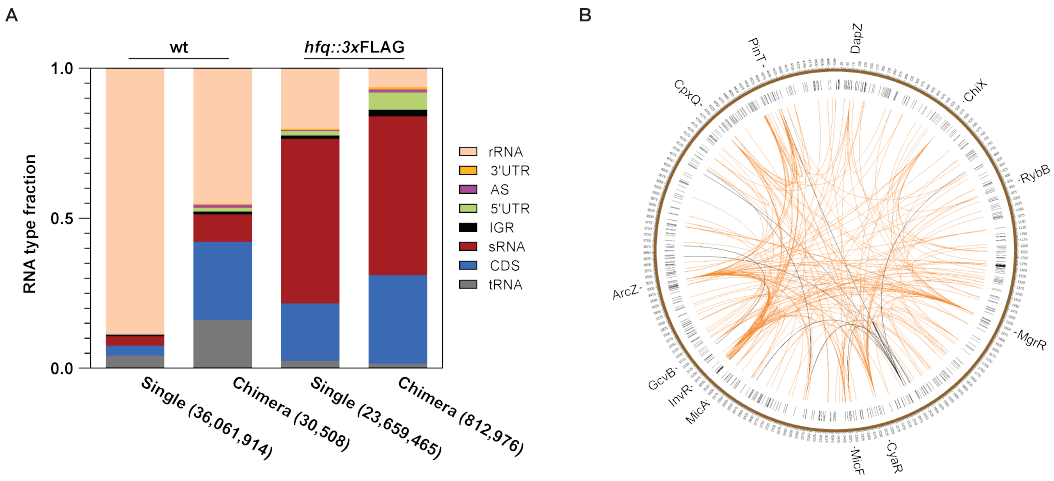

Figure 2.3 Distribution of RIL-seq data and global view of detected sRNAs.
(**A**) Relative frequency of each RNA type for chimeric and single fragments referred to one replicate. Libraries W1 and H1, for the WT and *hfq*::3×FLAG, respectively, were used as an example. Both single and chimeric reads mapping to *Salmonella* plasmid were excluded because of limitations in the available annotations. (**B**) Circos plot generated for library "H1" and drawn on the circular *Salmonella* chromosome. The plot shows only S-chimeras represented by at least 100 fragments. Interactions involving a sRNA in at least one of the two RNAs are marked in orange. The few interactions which do not involve sRNAs are marked in black. Labeled in the external part of the plot are sRNAs that dominate the network. The Circos plot was generated by Dr. Yael Altuvia.

2.1.3 Improved visualization of RIL-seq networks

One of the biggest challenges of global approaches like RIL-seq is the high complexity and large sizes of the generated datasets. The research community would benefit from interactome studies if data visualization was more interactive and intuitive.

To improve the visualization of these complex datasets, and to gain more insights into the Hfq network, we designed an online RIL-seq browser that allows all interactions to be simultaneously displayed (Figure 2.4, left part and Figure 6.2). A total of ~1,300 RNAs and ~1,770 S-chimeras were detected in the two replicates included in this study. The browser allows an interactive search of any kind of RNA type (e.g., sRNAs, CDS, UTRs, IGR, antisense, tRNAs) enabling a selective display of particular RNA regulons (Figure 2.4, right part and Figure 6.2). Where available, GO terms associated with each gene were included in the visualization tool, providing more biological information on the network. Finally, the online tool was developed to allocate any kind of RIL-seq dataset (e.g., from other bacterial species or RIL-seq experiments performed with a different RNA chaperone). As a proof of principle, we included in the current version of the network the previously published Hfq and ProQ *E. coli* RIL-seq datasets (Melamed et al., 2016, 2020) (https://resources.helmholtz-hiri.de/rilseqset/).

One of the advantages of including the *E. coli* Hfq datasets was to conduct a direct comparison to the *Salmonella* RIL-seq performed in similar growth conditions (Melamed et al., 2016, 2020). One of the main differences was statistics. While RIL-seq in *E. coli* detected sRNAs as RNA2 in ~90% of the S-chimeras (Melamed et al., 2016), this was less pronounced in the *Salmonella* data (~60%, Figure 6.1B). Although this parameter does not reflect directly on the quality of the dataset, it is nevertheless an indication of a functional Hfq hexamer. However, in the specific case of our *Salmonella* dataset, a very abundant pair (~12% of all sRNA interactions; Table S2B-C) between ArcZ sRNA and the *Salmonella*-specific *spaT* mRNA, could have been responsible for these altered statistics. For reasons unknown, this RNA-

RNA interaction dominanted the dataset and unexpectedly presented the sRNA as RNA1 in the chimera. While this aspect seemed to be the major difference between the *E. coli* and *Salmonella* datasets on a global scale, more similarities could be found with respect to a more biological question. In fact, examining the pool of interactions, a substantial overlap between the *Salmonella* and *E. coli* RIL-seq could be highlighted; disregarding the *Salmonella*-specific InvR and PinT sRNAs, the majority of the top-detected sRNAs in the *E. coli* RIL-seq data (stationary phase (Melamed et al., 2016)) matched those in the current study. This not only confirms that these two bacteria share a quite large amount of core sRNAs and therefore post-transcriptional control, but it proves that RIL-seq is a solid and reproducible method not only in *E. coli* but also within other Gram-negative bacterial models.

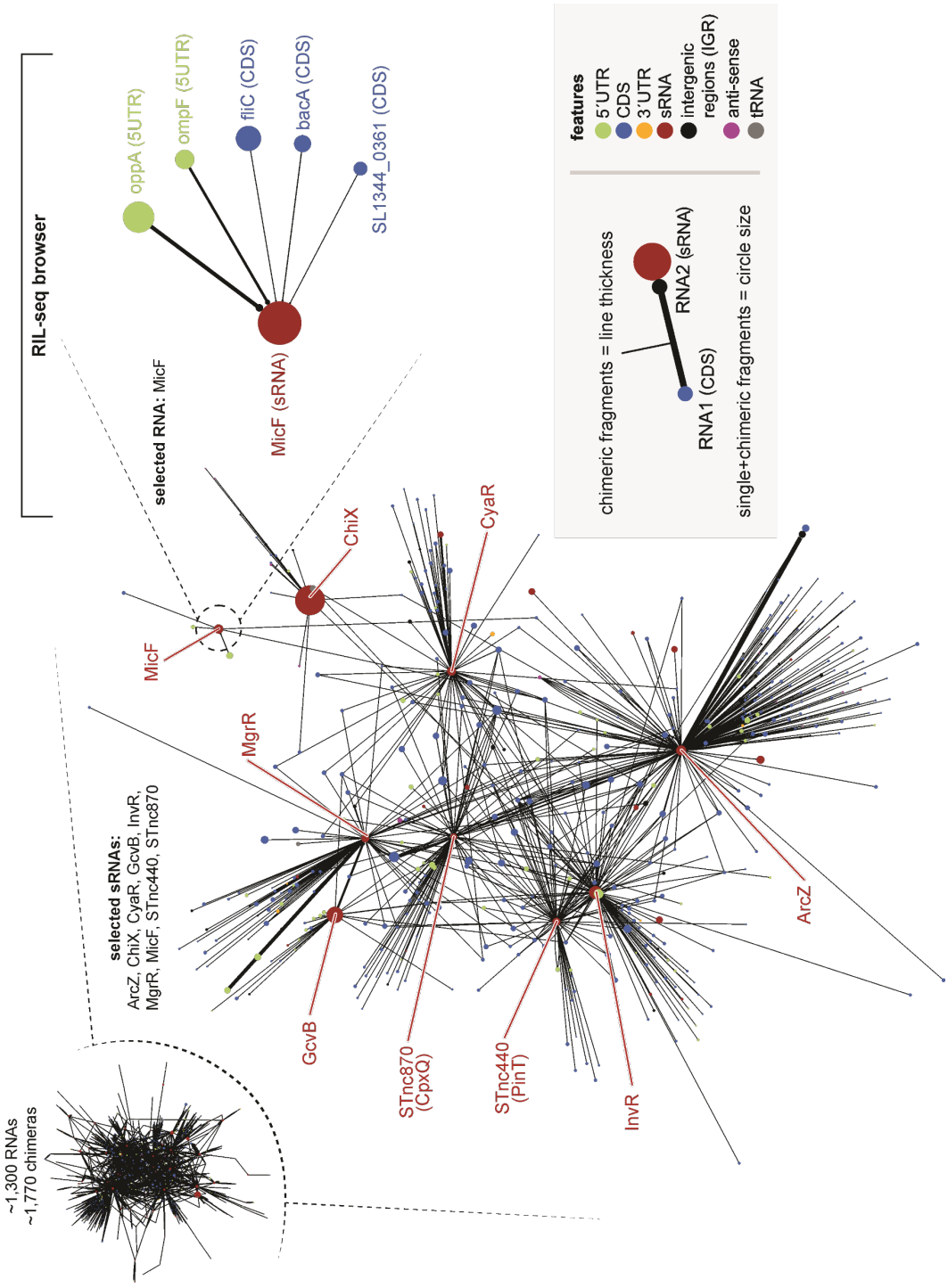

Figure 2.4 Visualization of RIL-seq data on the online browser.
Global representation of all interactions detected in the RIL-seq dataset (http://resources.helmholtz-hiri.de/rilseqset/). Highlighted in red are known *Salmonella* sRNAs, forming core regions in the network. A zoom-in view of the MicF detected targets is shown in the upper right part of the figure (this figure was created by Dr. Milan Gerovac).

2.2 sRNAs and RNA sponges linking OM and IM functions

2.2.1 sRNAs are involved in envelope cross-regulation

As previously stated, while our RIL-seq managed to detect RNA interactions involving 66 unique sRNAs, the most abundant and reliable interactions involved sRNAs and mRNAs encoding for genes with envelope-related functions. This gave us the possibility to further investigate new mechanisms of RNA regulation involving genes with membrane functions.

Therefore, with focus on envelope regulation, we analyzed our RIL-seq data filtering for new chimeras involving mRNAs of extracytosolic proteins. We created a list of *Salmonella* gene entries related to membrane when referred to 'location' in the cell or with a GO term 'membrane'. This list was mainly based on the BioCyc database (Karp et al., 2019) with additional manual amends to compensate for several missing membrane proteins. Overall, the final number of membrane-related genes was 1,119 (Table S3). By re-examining our RIL-seq data in view of this list, we identified 209 interactions formed by 29 unique sRNAs and 136 unique target mRNAs encoded by envelope-related genes (Figure 2.5A, Table S3). Not only did these extracted interactions include many, if not all, of the previously known sRNAs with major membrane-related functions, but they also contained new unexpected interactions. For instance, the *flgL* 3'UTR-derived sRNA FlgO (previously STnc840; (Chao et al., 2012; Hör et al., 2020)), was detected in chimeras with the mRNAs of *invG* (component of the *Salmonella* SPI-1 type 3 secretion system), *ompD* (the major *Salmonella* porin), and *napA* (periplasmic nitrate reductase) (Table S3). Interestingly, FlgO has been recently proposed to enhance bacterial invasion (Bomjan et al., 2019). Given the structural and functional similarities between flagella and the T3SS (Blocker et al., 2003), it is tempting to speculate that FlgO connects expression of flaggelar components (*flgL* mRNA) to regulation of specialized secretion systems (*invG*). Additionally, downregulation of the major *Salmonella* porin OmpD by sRNAs has been shown to be key for building up the bulky T3SS into the bacterial membrane (Pfeiffer et al., 2007).

Surprisingly, our chimeras analysis highlighted abundant envelope cross-regulation, i.e., where a sRNA previously known to target mRNAs of genes encoding one membrane also targeted another envelope compartment (Figure 2.5). For instance, the sRNAs MicA and RyhB, well-known repressors of OM major porins (Coornaert et al., 2010; Gogol et al., 2011; Udekwu et al., 2005), were detected in chimeras with additional mRNA targets encoding IM-proteins. MicA showed abundant chimeric fragments with the mRNA of HtpX, an IM-located protease, with the *tolB* mRNA, encoding the periplasmic part of the envelope-spanning Tol-Pal system, and with the well-known target *ompA* mRNA (Figure 2.5B, Figure 6.3A). In support of our RIL-seq data, the IntaRNA algorithm (Raden et al., 2018) predicted a stable RNA duplex between the seed of MicA and a conserved region in the 3' part of the coding sequence (CDS) of *tolB*, ~80 nt upstream of the *pal* start codon (Figure 6.3B-C). Altogether, these evidences could, at least in part, explain a previously reported downregulation of the *pal* mRNA by MicA in *E. coli* (Gogol et al., 2011). Similar *in silico* predictions were identified for other major MicA-containing chimeras (Figure 6.3D-G). The RybB sRNA showed abundant interactions with the mRNA of the IM-located sodium/proline symporter PutP, in addition to its previously identified targets, i.e., *ompA*, *ompC, ompD,* and *tsx* mRNAs (Figure 2.5C). In addition to this, sRNAs such as the 3'UTR-derived CpxQ and CyaR, well-established to regulate targets in the IM/periplasm (e.g., *skp*) or OM (e.g., *ompX*), respectively, each showed chimeras with additional membrane compartments (Figure 2.5D-E). Finally, our results confirmed the major involvement of the PhoP/Q-activated sRNA MgrR in IM regulation, detecting this sRNA in chimeras with the *pitA* and *ygdQ* mRNAs both encoding for IM proteins (Figure 2.5F). Both targets were identified in the *E. coli* RIL-seq (Melamed et al., 2016) as well as in a previous study focused on deciphering MgrR function in *E. coli* (Moon and Gottesman, 2009; Yin et al., 2019).

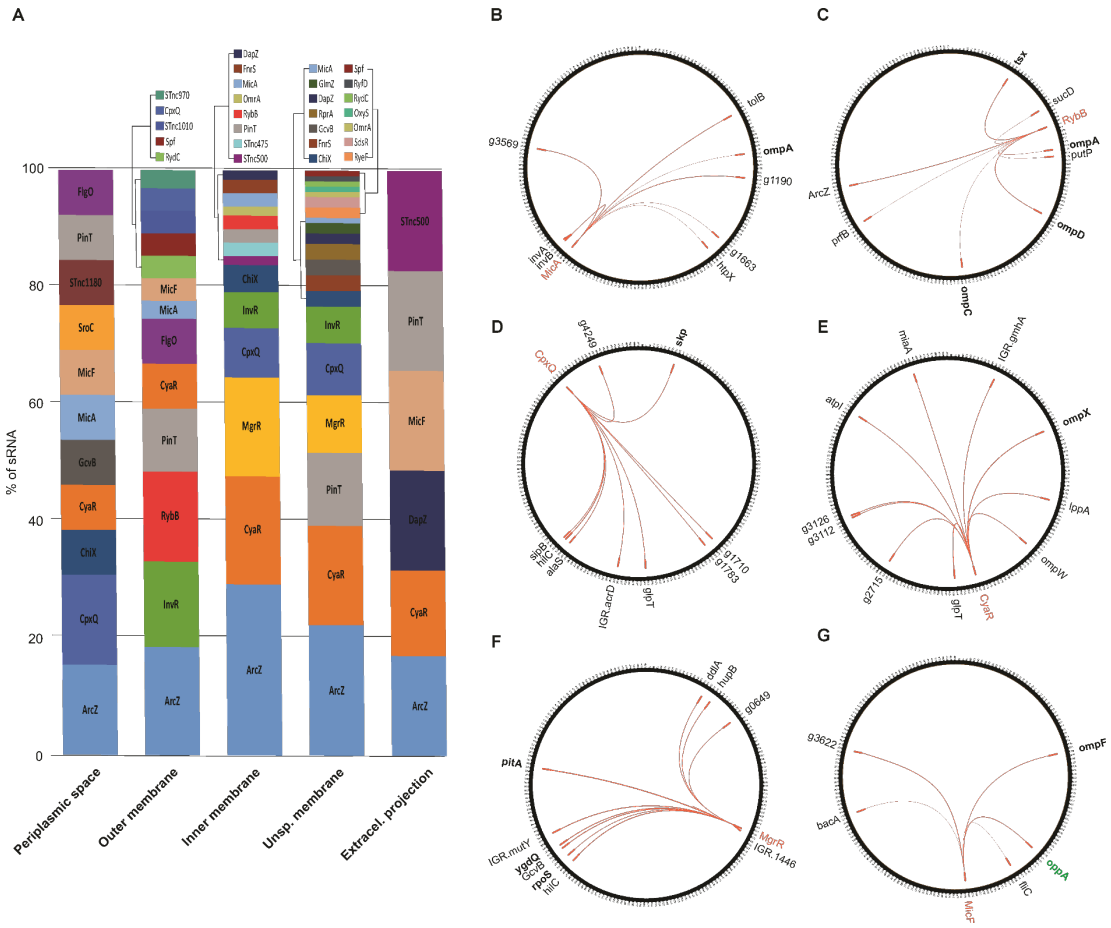

Figure 2.5 RIL-seq recapitulates envelope regulation by known sRNAs.
(**A**) Distribution of sRNAs forming S-chimeras with at least one protein coding gene annotated as component of the *Salmonella* envelope (Table S3 and Karp et al., 2019). Interactions from both replicates were pooled and included in the plot. sRNAs STnc870, STnc840 and STnc440 were here renamed to CpxQ, FlgO and PinT, respectively (Hör et al., 2020). (**B**, **C**, **D**, **E**, **F**, **G**) Circos plot representation of interactions involving known envelope-related sRNAs. The sRNAs are labled in orange in each plot, and previously known targets of each sRNA are highlighted in bold. The interaction further studied in this thesis is highlighted in green (**G**).

2.2.2 MicF-*oppA* interaction is conserved among *Enterobacteriacae*

A striking observation when digging into the *Salmonella* RIL-seq data regarded the chimeras formed by the classic porin repressor, MicF (Figure 2.5G and Figure 2.6A). As expected, a major interaction partner was the *ompF* mRNA (20-30%), historically representing the first example of sRNA regulation in bacteria (Mizuno et al., 1984). However, 40-60% of the total MicF S-chimeras (twice as much as with *ompF*) were obtained by ligation with the *oppA* mRNA, encoding for the first gene of the major IM transporter of oligopeptides operon *oppABCDF* (Figure 2.6A). Previous MicF target searches in both *E. coli* and *Salmonella* (Corcoran et al., 2012; Holmqvist et al.,

2012) had not seen any MicF effects on this operon, making this predicted interaction worthy of a deeper investigation. Interestingly, our RIL-seq data suggested strong similarity between the MicF hybrids with *oppA* and *ompF*. In fact, both formed extended duplexes with the entire MicF sequence. By exploring a detailed representation of the RIL-seq reads, we could appreciate that both the 5′ and the 3′ ends of MicF were recovered in the chimeric fragments, reflecting the fact that upon RNase treatment short RNAs integrity is largely preserved (Figure 2.6B, blue reads, Table S2B-C). Conversely, only a restricted region of the respective 5′UTRs was retrieved in the hybrids (Figure 2.6B, red reads). Nevertheless, this suggested that both 5′UTRs are likely engaging in an interaction with the outstandingly conserved seed sequence of MicF (Figure 6.4A and (Corcoran et al., 2012)). Moreover, with respect to *in silico*-predicted hybridization energy, these interactions were similar (*oppA* -22 kcal/mol; *ompF* -25.48 kcal/mol; Figure 2.6C-D). However, a previous in-depth study of MicF identified only the first 12 nt of the seed sequence as essential for *ompF* regulation (Corcoran et al., 2012).

A sequence alignment of different *oppA* genes was performed to assess the degree of conservation within the Enterobacteriacae family. The results suggest marked local conservation of the putative MicF binding site within the *oppA* 5′ UTR (Figure 6.4B), spanning the complement of the entire seed region of MicF. Apart from the predicted MicF-binding site and another highly conserved region overlapping with the RBS of *oppA,* this 5′UTR does not appear to be highly conserved. Collectively, these data suggest that MicF might need the whole 25 nt-long seed sequence to interact with *oppA*, contrary to the first 12 nt needed to regulate *ompF*.

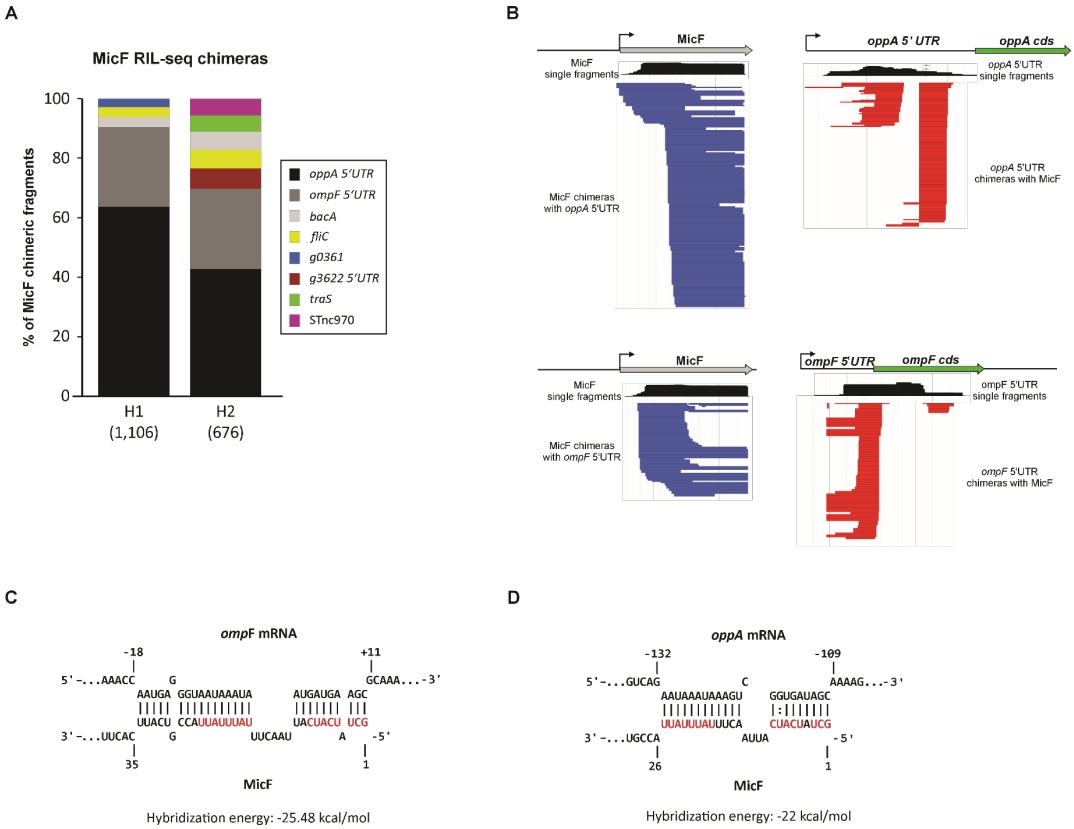

Figure 2.6 MicF predominantly forms chimeras with the *oppA* 5'UTR.
(**A**) S-chimeras involving MicF sRNA in the two RIL-seq replicates (H1 and H2) are plotted relative to the total number of chimeras involving MicF. 40-60% of the total number involves the *oppA* 5'UTR, while the known target *ompF* takes up 20-30% of the total. g0361 and g3622 stand for the genes annotated as SL1344_0361 and SL1344_3622, respectively. (**B**) Genome browser-based screenshot of the MicF, *ompF* and *oppA* 5'UTR chimeric-reads shown in stacked mode. The upper panel shows reads belonging to the MicF and *oppA* 5'UTR chimeras, the lower panel shows reads belonging to the MicF and *ompF* chimeras. Reads extracted from the first or second positions in the chimeras are marked in red or blue, respectively. For visualization reasons, only one fourth of the total pool of chimeric reads were randomly selected and shown. The black plot above each chimeric reads region represents coverage of the respective RNA derived from the RIL-seq single fragments counts. This information recapitulates binding of the RNA to the Hfq chaperone. The plot was generated using JBrowse. (**C**) Prediction of RNA duplex formation between the MicF sRNA with the *ompF* mRNA or with the 5'UTR of *oppA* (**D**).

2.2.2.1 MicF does not affect OppA synthesis

The typical means of regulation of an Hfq-dependent sRNA such as MicF is to repress or activate mRNA translation by binding near the AUG start codon (Hör et al., 2020). However, the predicted MicF site in *oppA* lies ~160 nts upstream of the respective start codon and outside of the ribosome binding site (RBS) window, a less likely mRNA region for effective translational control (Figure 2.7A). To determine whether MicF regulated

oppA, we used a translational reporter system in which the 5′ region of either *ompF* (used as a positive control) or *oppA* was fused to the CDS of the green fluorescent protein (*gfp*), and co-transformed with a plasmid driving the expression of the MicF sRNA (Figure 2.7B, left cartoon). MicF expression led to a complete loss of fluorescence of strains carrying an *ompF::gfp* fusion (Figure 2.7B, right panel). By measuring GFP protein levels with a specific antibody we could assess that MicF downregulates *ompF* translation by 12-fold, as previously reported (Figure 2.7C, (Corcoran et al., 2012)). In contrast, MicF did not have any influence on the fluorescence of strains carrying an *oppA::gfp* fusion, nor on the GFP levels as measured through western blot (Figure 2.7B-C).

Alongside this, we could take advantage of an OppA-specific antibody for immunoblotting to measure the endogenous levels of the *Salmonella* OppA protein. The western blot showed that the levels of OppA were unaffected by deletion of the chromosomal *micF* gene or by a plasmid-driven MicF overexpression (Figure 2.7D, lane 4-5). Additionally, transcript levels of the *oppA* mRNA were unaffected by MicF, as determined by RT-qPCR-based quantification with primers binding in the *oppA* CDS or 5′ UTR. In contrast, *ompF* mRNA levels were clearly influenced by MicF knockout or overexpression, proving that the sRNA was active in this growth condition (Figure 2.7E). Finally, OppA levels were expectedly increased in a Δ*hfq* background (Figure 2.7D, lane 3), indicating that Hfq-dependent mRNA regulation of *oppA* by other endogenous Hfq-dependent sRNAs (DapZ, GcvB (Chao et al., 2012; Sharma et al., 2007)) was functional in the growth condition used here. Altogether, these results discredited the traditional model in which MicF is the regulatory RNA and the *oppA* mRNA the target.

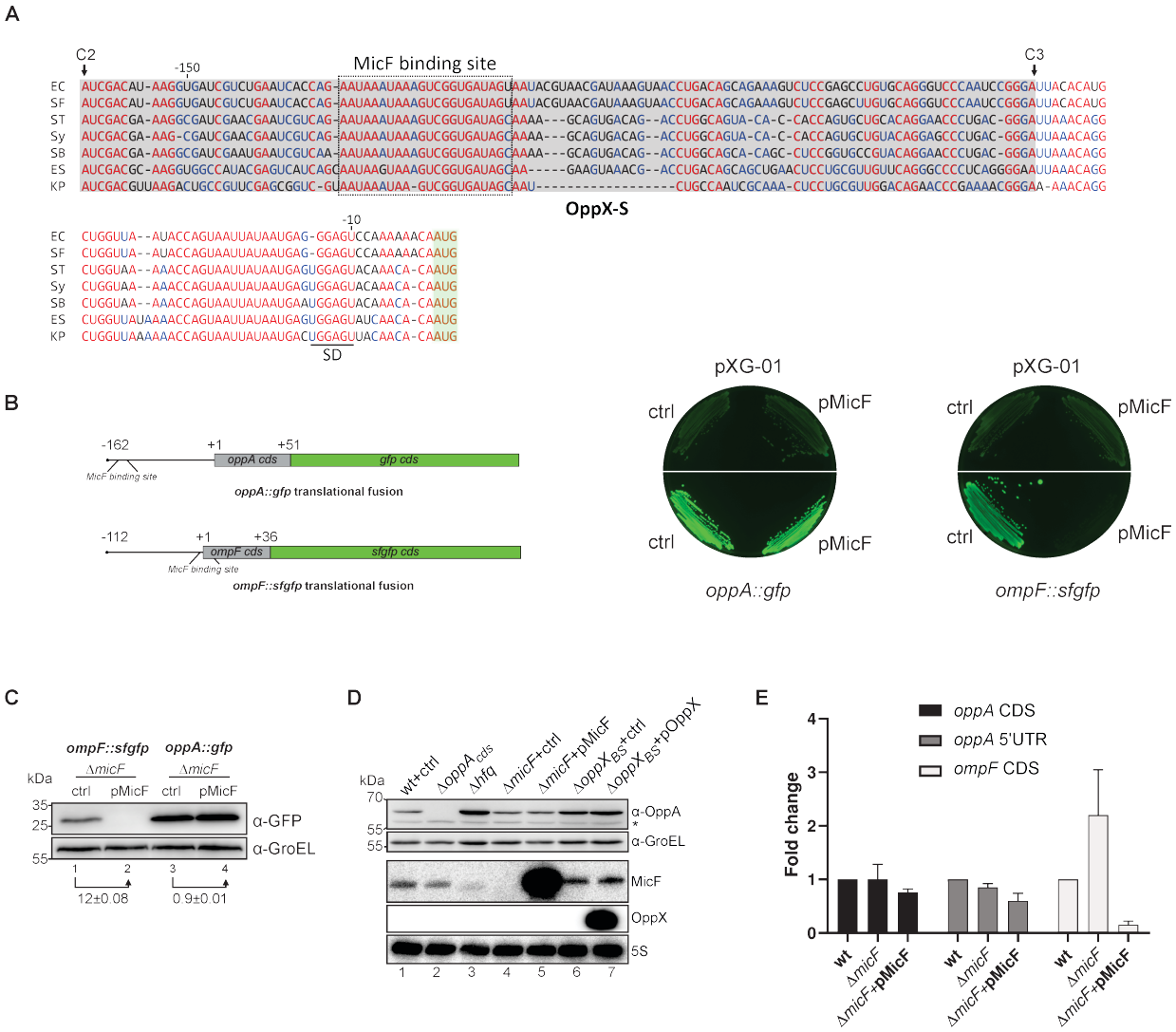

Figure 2.7 MicF does not regulate *oppA* at the RNA or at the protein levels.
(**A**) Multalin representation of the *oppA* alignment from the -162 position until the AUG, among Enterobacteriacae. EC= *Escherichia coli;* SF= *Shigella flexneri;* ST= *Salmonella* Typhimurium; SY= *Salmonella* Typhi; SB= *Salmonella* bongori; ES= *Enterococcus sp.* 638; KP= *Klebsiella pneumoniae.* (**B**) On the left panel, a schematic representation of the *gfp* translational fusion constructs of the *ompF* and *oppA* 5'UTRs, including the first 12 and 17 amino acids, respectively. Highlighted between black lines are the predicted MicF binding sites on both constructs. On the right panel, LB Agar plates showing fluorescence upon exposure to UV light. Strains containing the control plasmid (pXG-01) with or without a background expression of MicF (pMicF) are shown in the upper part; strains containing the two *gfp* reporter plasmids (*oppA* on the left, *ompF* on the right) with or without MicF overexpression. (**C**) Western blot analysis of total protein samples taken at OD 2.0 in LB media of the indicated strains carrying either the *ompF* or the *oppA* GFP fusion constructs co-expressed with a control plasmid (pJV300) or with a plasmid expressing MicF (pMicF). Quantification of individual experiments are shown below the gels as MEAN and ±SD values. (**D**) Western blot detection of endogenous OppA protein in *Salmonella* in several genetic background confirms no regulation upon MicF deletion or plasmid-driven expression. Detection of MicF and *oppA* UTR on a northern blot was used as a control. (**E**) qPCR measurements of a WT, Δ*micF* + pJV300 and Δ*micF* + pMicF strains grown to OD_{600} 2.0 in LB. DNA oligos binding in the coding sequence of *oppA* (*oppA* CDS, black histograms), in the 5'UTR of *oppA* (*oppA* 5'UTR, grey histograms) or in the coding sequence of *ompF* (*ompF* CDS, light grey histograms) were used to amplify and detect the respective RNA fragments. 5S rRNA was used to normalize the raw values. Three biological replicates were performed. For the northern blot and the western blot, 5S and GroEL detection was used as a loading control, respectively.

2.2.2.2 The long *oppA* 5'UTR generates stable RNA fragments

The *oppA* 5'UTR is longer than typical bacterial 5'UTRs, measuring on average 20-60 nts in length. A previous study using 5'RACE proposed the *oppA* transcription to start 162 nt upstream of the *oppA* start codon (Sharma et al., 2007). However, the SalCom database (Kröger et al., 2013) and dRNA-seq experiments suggested the presence of another TSS ~500 nts upstream of the ATG, as well as several possible internal cleavage sites to which we will refer to as C1, C2 and C3 (Figure 2.8A).

In support of the processing events predicted by SalCom and dRNA-seq, a northern blot probing of the *oppA* 5'UTR using a labeled DNA oligo antisense to the -132 to -109 region (relative to AUG) revealed the existence of stable RNA fragments (Figure 2.8B-C). For simplicity, we will refer to these detected RNA species from the *oppA* 5'UTR as OppX. The band above the 404-nt marker band matches the size of the full-length 5'UTR (OppX-L, 460 nts), spanning from the more upstream TSS to the C3 cleavage site. The intermediate (OppX-M, 190 nts) and the smallest (OppX-S, 109 nts) bands would correspond to fragments spanning from C1 to C3, and C2 to C3, respectively. We reasoned that the *oppA*-MicF chimeras detected in RIL-seq occurred with these independent 5'UTR fragments rather than with the full-length *oppA* mRNA, given their abundance. Interestingly, all these OppX fragments were also detected in a Δ*micF* strain, suggesting that the MicF sRNA is neither involved nor required for correct processing of the transcribed *oppA* 5'UTR (Figure 2.8B-C, lanes 2).

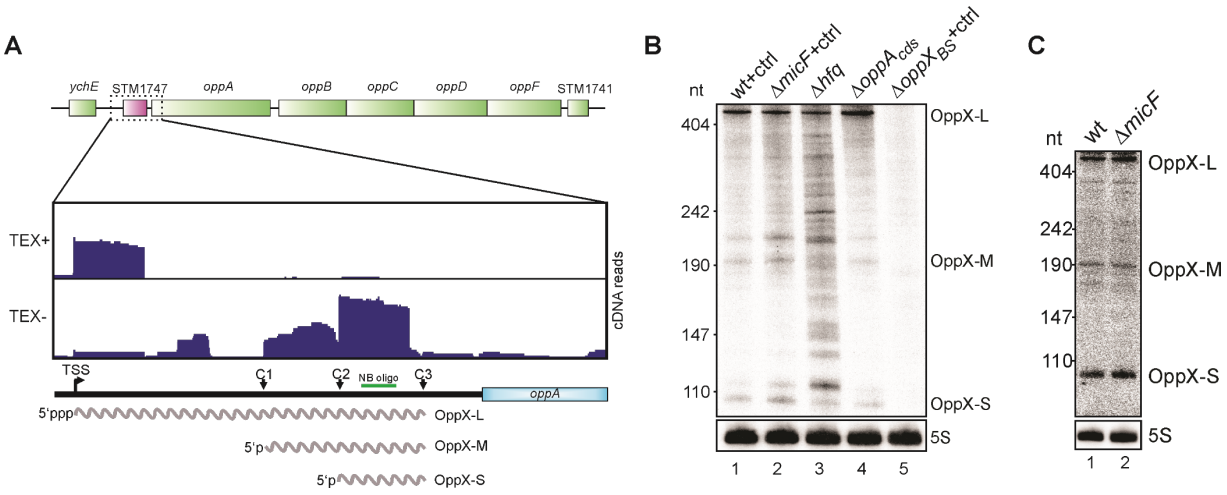

Figure 2.8 *OppA* generates RNA species *in vivo* in a MicF-independent manner.
(**A**) dRNA-seq data shows a new TSS for the *oppA* locus (highlighted with a black arrow in the schematic view). TSS are predicted by enrichment of read coverage in the TEX+ library compared to the TEX-. Enrichment of TEX- compared to TEX+ suggests the presence of a processing site (indicated as C1, C2 and C3 in the cartoon). These sites were as well predicted in the published TIER-seq study (Chao et al., 2017). (**B**,**C**) Northern blot membranes probed with a labeled DNA oligo (JVO-16463) binding inside the 5'UTR of *oppA*. 5S rRNA was used as a loading control.

2.2.2.3 RNases are involved in the maturation and turnover of OppX

RNase E represents the main nuclease responsible for mRNAs turnover in *Salmonella*. A map of transcriptome-wide RNase E cleavage sites (TIER-seq (Chao et al., 2017)) predicted multiple recognition sites within the *opp* operon including the 5'UTR of *oppA*. Interestingly, the most prominent cleavage sites corresponded to the C1, C2 and C3 positions described above. Additionally, these sites contain AUU sequences, a well-known consensus able to recruit RNase E (Chao et al., 2017).

As deletion of the gene coding for RNase E (*rne*) is lethal in *Salmonella,* we took advantage of an *rne* allele that encodes a temperature sensitive RNase E to experimentally investigate the steps of OppX processing (Figueroa-Bossi et al., 2009). Surprisingly, the necessary shift to higher growth temperature to inactivate RNase E alone affected the stability of the OppX fragments, i.e., regardless of whether *Salmonella* expressed the wild-type or the temperature-sensitive RNase E protein (Figure 2.9A; compare the 28°C to the 44°C samples). Nonetheless, the OppX-L form only accumulated in the *rne*-TS strain at 44°C, arguing that RNase E is required for the 5'UTR

processing of *oppA*. In an attempt to overcome this heat instability of OppX fragments, we expressed the *oppA* UTR from a plasmid-driven constitutive promoter. Even in this case, despite the high expression levels of OppX in this setup, the shift to 44°C was generally deleterious for the short OppX isoform. Particularly, we noticed that those isoforms that lost the plasmid-borne terminator upon processing at the C3 site were mostly affected by temperature. In fact, the remaining bands stayed stable at high temperature (Figure 2.9B). In this scenario, however, inactivation of RNase E clearly abolished correct 5'UTR processing. This would support a model where RNase E is implicated in OppX biogenesis as a consequence of processing of the *oppA* mRNA. Additionally, deletion of the *hfq* gene also affected processing of the *oppA* 5'UTR (Figure 2.8B, lane 3). While not followed up here, we speculate that Hfq—alone or in concert with sRNAs—remodels this 5'UTR for correct processing by RNase E (Chao et al., 2017).

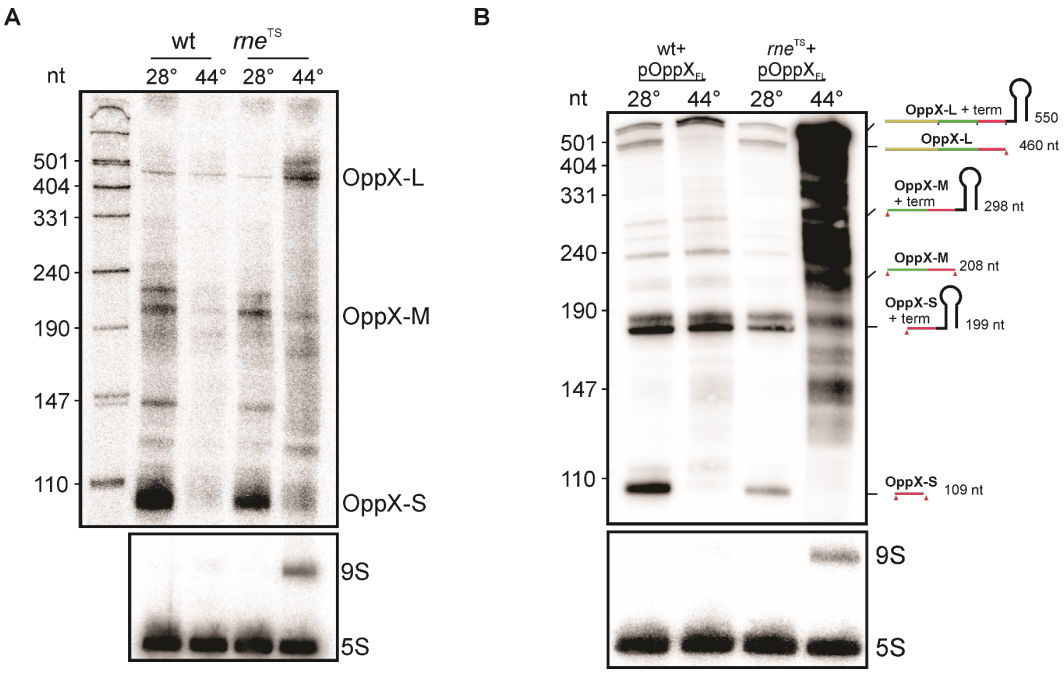

Figure 2.9 The *oppA* 5'UTR undergoes RNase E-dependent biogenesis.
(**A**) Northern blot probing of OppX in an isogenic wt and an *rne*^{TS} strains grown at 28°C (permissive temperature) to an OD_{600} of 2.0 and shifted to 44°C (non-permissive temperature) for 30 min. 5S and the incorrectly-processed 9S rRNAs were used as loading control and internal control for a correct RNase E inactivation, respectively. (**B**) As in (**A**) but strains were transformed with a high-copy plasmid encoding the full-length OppX (550 nt from the TSS including the plasmid-borne terminator).

We further assayed whether the temperature-dependent lability of OppX was promoted by the activity of a different nuclease. Therefore, we repeated the 28°C to 44°C shift experiment and measured the plasmid-driven OppX expression in several deletion backgrounds of known *Salmonella* nucleases. Interestingly, we could identify the exoribonuclease PNPase as a major player in the degradation of the OppX-S transcript (Figure 2.10). Nevertheless, a partial temperature-dependent lability was still detectable, indicating that this RNA species is either intrinsically less stable at higher temperatures or the activity of an additional nuclease is deployed.

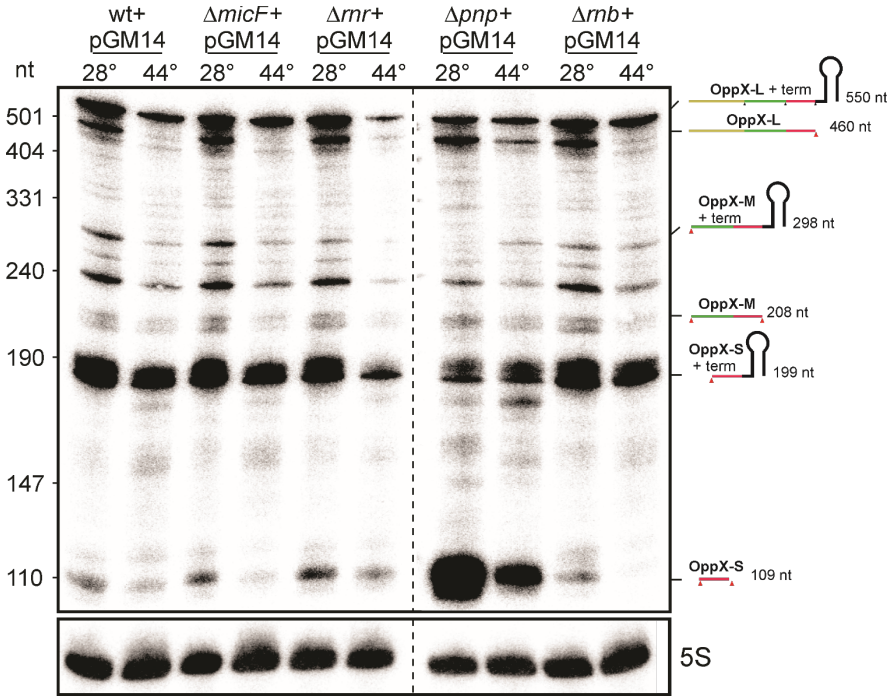

Figure 2.10 PNPase participates in the turnover and stability of OppX-S.
Northern blot probing of OppX in a *Salmonella* WT strain, Δ*micF* or several RNases single deletion backgrounds (Δ*rnr:* RNase R; Δ*pnp*: PNPase; Δ*rnb*: RNase II). All the strains contained a plasmid overexpressing a full-length version of OppX (550 nt from the TSS, including the plasmid-borne terminator). Strains were grown and treated as described in Figure 2.8. Right side: schematic representation of RNA species detected in the blot, according to expected length in nt. 5S rRNA was used as loading control.

2.2.2.4 OppX and MicF form stable complexes *in vitro*

Despite the fact that MicF RNA did not affect the expression (at the RNA and protein levels) of *oppA*, our RIL-seq data clearly suggested these two RNAs base-pair *in vivo.* To support this assumption, we investigated whether MicF and OppX could form stable complexes *in vitro* by performing electrophoretic mobility shift assays (EMSA). Indeed, 5'end-labeled *in vitro* transcribed MicF or OppX shifted in a native polyacrylamide gel when incubated with increasing concentrations of cold (unlabeled) RNA partner (Figure 2.11A, upper and lower gel, respectively). By plotting the percentage of bound RNA (y-axis) against the concentration of the unlabeled MicF RNA (x-axis) we could extrapolate a binding curve through which we estimated an apparent dissociation constant (k_d) of 25 nM (Figure 2.11A, lower graph). Importantly, these affinity values between RNAs are very typical of regulatory interactions in bacteria. The high affinity of binding could be partially explained by the evidence that both regions of MicF and OppX involved in the interaction are embedded in single-strand regions or poorly structured loops (Figure 2.11B-C).

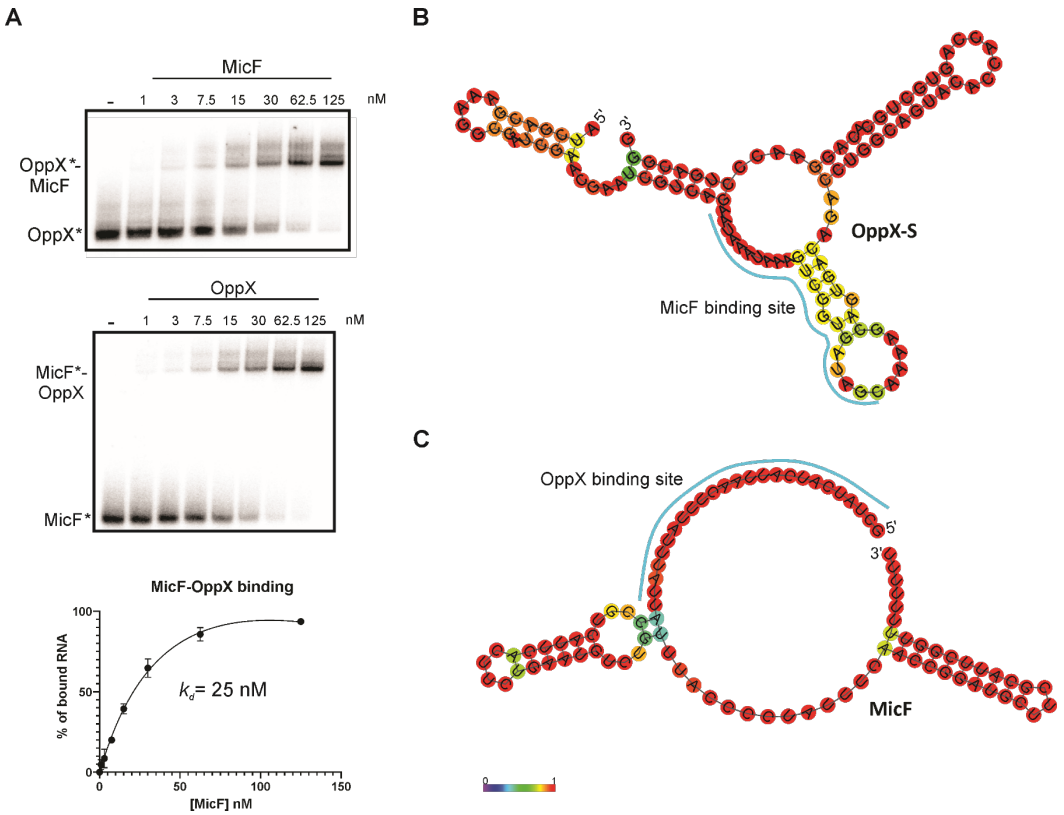

Figure 2.11 MicF and OppX-S form complexes *in vitro*.
(**A**) Band shift assays between 5′ end-labeled MicF and unlabelled OppX, or vice versa (top and bottom gel, respectively). Unbound labelled RNAs and in complex are indicated on the left of the gel. Increasing concentration of unlabelled RNA are indicated above the gel (0, 1, 3, 7.5, 15, 30, 62.5, 125 nM). In the lowest part, a binding curve was drawn with GraphPad Prism software whereby the concentration of unlabelled MicF (x-axis) was plotted against the bound fraction of labelled OppX (y-axis). The dissociation constant k_d was determined by estimating the concentration of MicF at which 50% of labelled OppX was in a bound state. (**B**) Secondary structure prediction of OppX-S and MicF (**C**) computed with the online tool RNA*fold* (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi). Indicated with a blue line is the respective region engaged in the duplex formation between OppX and MicF. The color coding of the nucleotides indicate probabilities of that specific nucleotide to be in a single strand or in a stem-loop (score: 0, blue; 1, red).

Given the high binding strength between MicF and OppX, we tested these two RNA species for lead(II)-based structure probing, alone or in complex with each other. This method allows for a precise mapping of the binding sites of RNA species at a single nucleotide resolution. When probing the 5′ end-labeled MicF, we observed protection in its 5′ single-stranded seed region when OppX-S was present (Figure 2.12A, left gel). Reciprocally, MicF protected a >20-nt internal region of 5′ end-labeled OppX-S (Figure 2.12A, right gel). Based on these results and high nucleotide conservation in related species (Figure 2.7A, Figure 6.4A-B), we conclude that MicF and OppX-S form a 21-bp imperfect RNA duplex (Figure 2.12B).

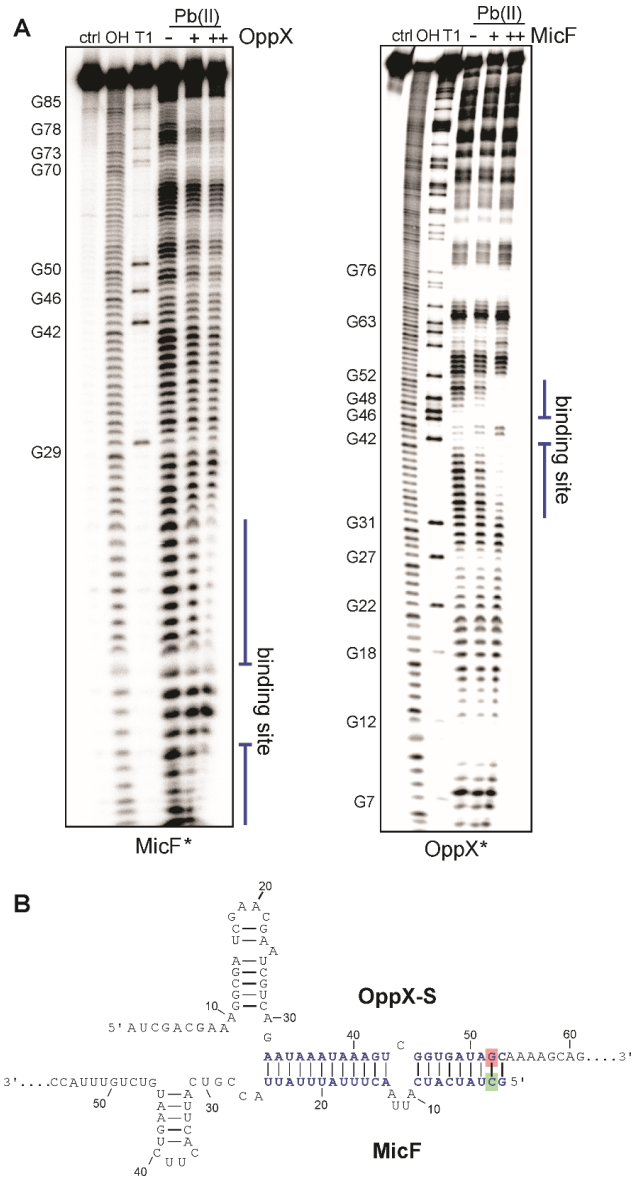

Figure 2.12 Binding between the seed of MicF and an internal region of OppX.
(**A**) RNA footprint using lead-acetate as cleavage chemical. The left panel shows the cleavage pattern of labelled MicF RNA in the presence of Pb2+ incubated with none (-) or increasing concentration of cold OppX (30 mM (+) or 500 mM (+)). The right panel shows the reverse experiment performed as described for MicF. The "ctrl" lanes indicate untreated labelled RNAs; "OH" lanes indicate alkaline treatment of RNAs to generate a single nucleotide ladder; "T1" lanes indicate RNase T1 digested RNAs for the G-ladder. Blue lines indicate regions protected by the unlabelled RNA. (**B**) Representation of the binding between MicF and OppX as confirmed from the structure probing and as suggested from IntaRNA predictions (Figure 2.6D). RNA structure representations are based on RNAfold predictions (Figure 2.11B-C).

A previous study showed hypersensitivity to mutations of the MicF seed sequence. Particularly, a C_2->G change was reported to abrogate the repression of several MicF targets (Holmqvist et al., 2012). This same single point mutation also impaired complex formation of MicF with OppX-S (Figure 2.13A, upper panel) in our *in vitro* system whereas a compensatory point mutation (G_{52}->C) in OppX-S would restore the stable complex (Figure 2.13A, lower panel). Having demonstrated that OppX is able to sequester the seed sequence of MicF, we sought to investigate whether this same RNA could compete with the formation of other MicF-RNA complexes. To do so, we aimed to challenge a MicF-*ompF* RNA complex, given their similar hybridization energy and k_d to the MicF-OppX complex (Figure 2.6C-D, Figure 2.11A and Figure 2.13B, upper gel). Indeed, a pre-formed MicF-*ompF* complex gradually collapsed as increasing concentrations of (unlabeled) OppX-S RNA were added to the reaction (Figure 2.13B, lower gel). The ability of OppX-S to interfere with the base pairing of MicF with *ompF* suggested that OppX-S could function as an RNA sponge, buffering the activity of MicF by sequestering this sRNA's seed region.

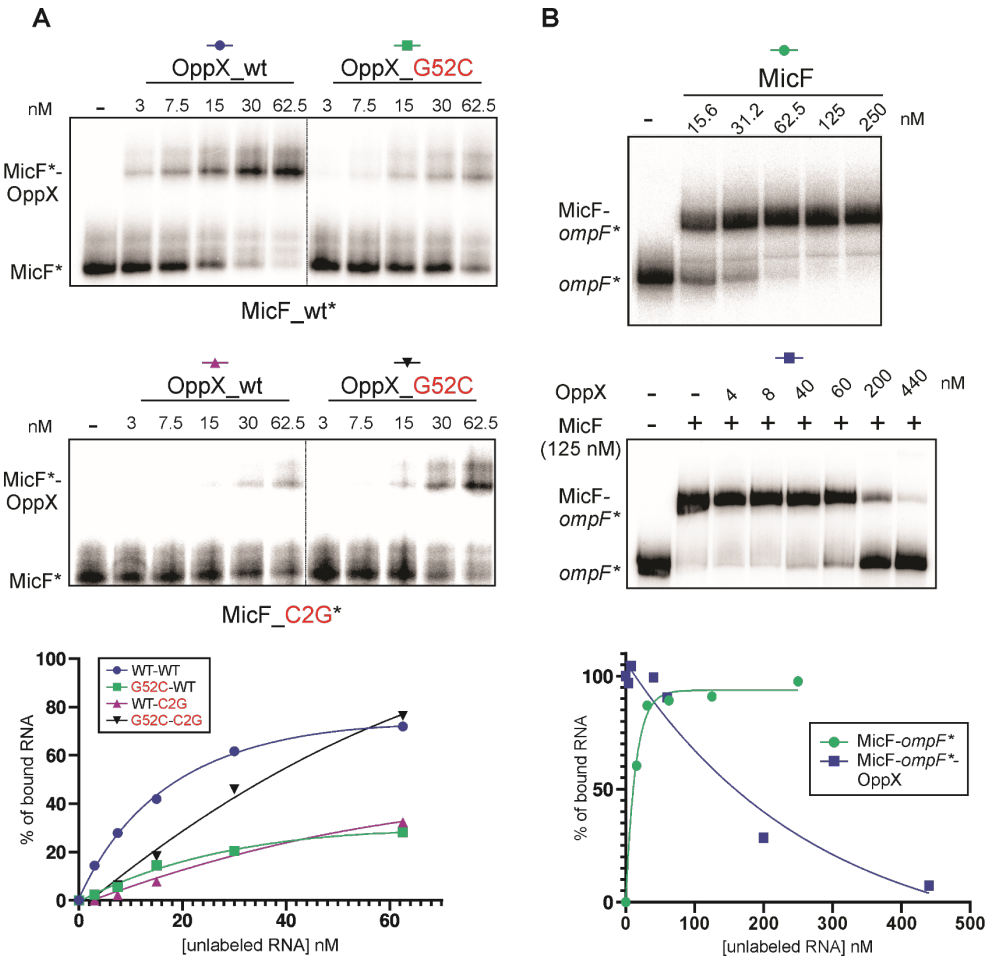

Figure 2.13 Mutations disrupt MicF-OppX binding and competition of OppX-MicF-ompF complex.
(**A**) EMSA gels of labeled WT (top gels) or C2G mutant (bottom gels) of labeled MicF RNAs with increasing concentrations of WT and G52C mutant of unlabeled OppX RNA (left and right gels, respectively). Quantification of the complexes is shown for these representative gels in the bottom plot. (**B**) EMSA of labeled ompF 5'UTR fragment with increasing concentration of MicF (top gel). Competition EMSA of labeled ompF 5'UTR with fixed concentration of MicF (125 nM) and increasing concentration of unlabeled OppX RNA (bottom gel). Quantification of the complexes in the presented gels is shown in the plot below (n=2 independent experiments).

2.2.2.5 OppX-MicF form stable complexes *in vivo*

MicF and OppX stably interact *in vitro* (Figure 2.11, Figure 2.12 and Figure 2.13) and OppX expression does not affect MicF stability (Figure 2.7). We then searched for evidence proving the existence of an inhibitory OppX-MicF complex *in vivo*. We performed fractionation of cellular complexes present in *Salmonella* lysates by centrifugation on a glycerol gradient (10-40% w/v) (Smirnov et al., 2016). To assess the quality of the fractionation experiments, an equal volume for each fraction was loaded on a 6% PAA gel and stained with ethidium bromide to highlight abundant RNA species. Particularly, this QC step is necessary to evaluate the correct sedimentation and the integrity of known RNA-protein complexes. For instance, tRNAs sedimentation profile peaks within the first 5 fractions, while ribosomal RNAs in the higher molecular weight fractions (10-20) (Figure 6.5).

Additionally, an RNA northern blot gel was probed with end-labeled oligonucleotides. An oligo binding within MicF primarily detected this sRNA in fractions #4 and #5, while OppX was detected in fractions #5 and #6 (Figure 2.14A), consistent with previous high-throughput Grad-seq data obtained in the same growth condition (Smirnov et al., 2016). The same experiment was performed with the lysate of a Δ*oppX* strain, showing a slightly extended sedimentation profile for MicF, towards higher molecular weight fractions, #6 and #7. As a control, we measured the differential sedimentation profile of the functionally unrelated Hfq-dependent ChiX sRNA (Figure 2.14B) or $tRNA^{Pro}_{CGG}$ (Figure 2.14C). Neither of the two RNAs showed an OppX-dependent change in their in-gradient distribution when probed for on the same blot membrane.

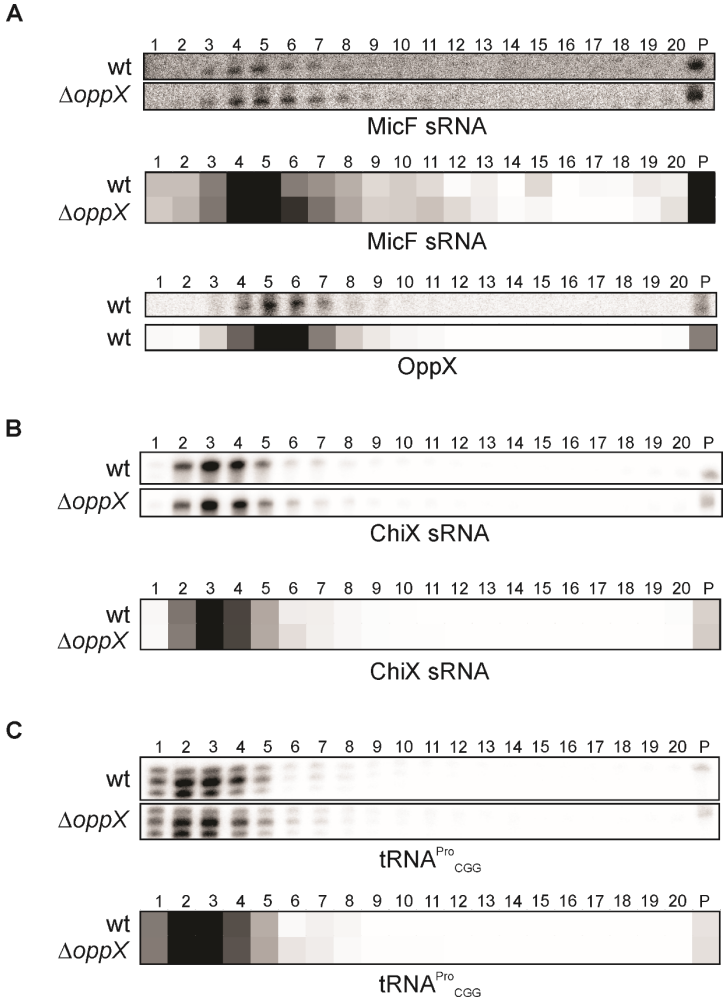

Figure 2.14 OppX deletion affects MicF sedimentation in a glycerol gradient.
(**A**)Northern blot probing of MicF (top) and OppX (bottom) of membranes showing glycerol gradient fractionation of a WT and a Δ*oppX* strains. (**B**, **C**) Same membranes of panel (**A**) probed for ChiX (top) or $tRNA^{Pro}_{CGG}$ (bottom) as a control. Densitometry plots of each hybridization is shown below the gels. The intensity of each band was quantified and normalized to the sum of intensities of all bands in the blot. Numbering above the gels indicate the in-gradient fractions (1-20). "P" represents the pellet fraction.

To confirm the observed OppX-dependent sedimentation of MicF we performed a higher resolution fractionation experiment using size exclusion chromatography (SEC, detailed protocol in Materials & Methods). *Salmonella* lysate of a WT or a *ΔoppX* strains were ran through a SEC column and fractions were collected according to their UV_{254} profile. Specifically, we retained fractions 29 to 48, as this was previously established to be the range cellular complexes sediment to (previous experiments performed in *E. coli* by Dr. Kotaro Chihara, personal communication). Unlike glycerol gradients, SEC fractions eluting as last (higher fraction number) represent lower molecular weight complexes. As for the gradient, we ran PAA gels and

stained them with ethidium bromide to check the quality and distribution of the RNA (Figure 6.6). With the same RNA samples two northern blot gels were performed and hybridized with the MicF probe and the OppX probe. In agreement with the gradient results, the sedimentation profile of MicF was altered by the deletion of the OppX RNA. Specifically, in the WT strain MicF showed a peak from fraction #39 to #35 and a second minor peak from fraction #31 to #29 (Figure 2.15A, upper gel), whereas in the *ΔoppX* the first main peak partially disappeared while the second peak increased relative to the total distribution (Figure 2.15A, lower gel). As a control, ChiX sedimentation was not affected by the presence or absence of OppX (Figure 2.15B). This suggests that when not sequestered by OppX, the MicF sRNA associates with other partners, either in the process of or in preparation for repressing the *ompF* mRNA.

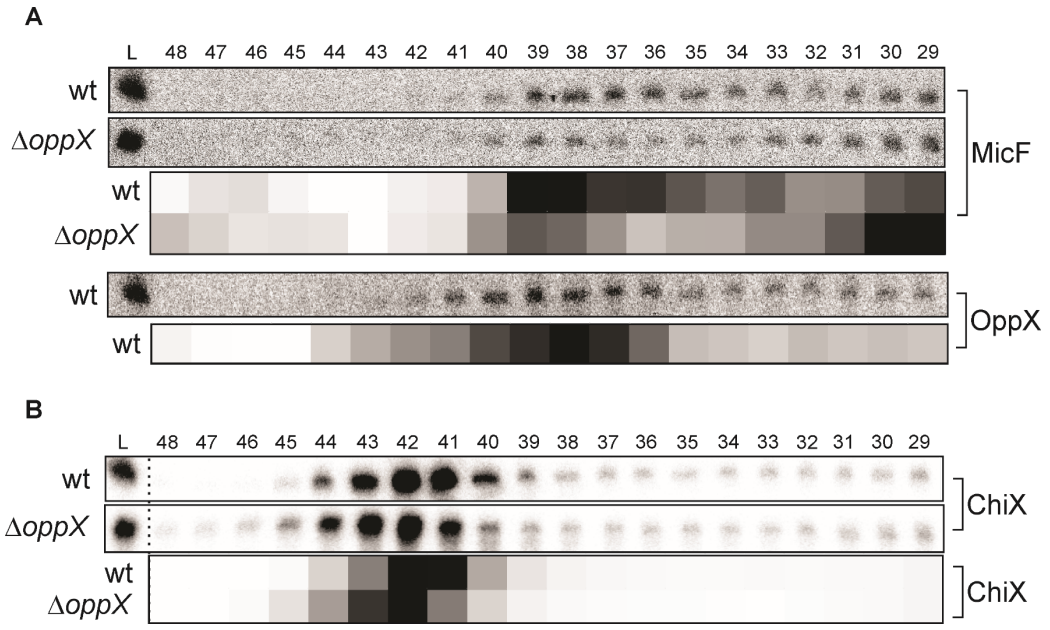

Figure 2.15 SEC confirms the glycerol gradient observations.
SEC experiments of a WT and a *ΔoppX* strains. Fractions #29 to #48 were collected, and RNA samples separated on a northern blot gel. (**A**) Hybridization of the membranes with a MicF or an OppX labeled DNA oligonucleotide. (**B**) Hybridization of the membranes with a ChiX probe was used as a control. Densitometry plots were performed as described for Figure 2.14.

2.2.2.6 Further characterization of the MicF-OppX complex

Having established that MicF and OppX form a stable complex *in vivo* we attempted to characterize other protein component of such a complex. We reasoned that, given the discovery of this interaction was provided by our RIL-seq data, Hfq must represent a central part of it.

We therefore performed MS2-pulldown of a strain carrying a plasmid-encoded MS2-tagged MicF sRNA or an MS2-tagged OppX RNA. The experiments were designed and performed according to the MAPS protocols established in *Salmonella* (Correia Santos et al., 2021). Upon binding of the lysates to an MBP-carrying purification column, specific complexes were eluted with maltose. Eluted RNAs and proteins were separately isolated and separated onto a silver-stained SDS-PAGE gel or a northern blot gel, respectively. The silver-stained gels show, in the eluate lanes (5-6 and 7-8), two major bands between the 35 and 55 kDa mark (Figure 2.16A). The higher molecular weight band (~55 kDa) corresponds to the maltose binding protein (MBP) immobilized on the column and used for binding the MS2 aptamer; the lower molecular weight band (~40 kDa) corresponds to MalE, a known *Salmonella* maltose-binding protein. Additionally, in the MS2-MicF eluate, a smaller band was specifically enriched when compared to the control sample (Figure 2.16A, lane 5-6). We reasoned this band to likely correspond to the Hfq monomer, given the typical ~10 kDa size mark. No other band could be specifically identified in this lane, nor could any band (except for the MBP and MalE) be detected in the MS2-OppX sample (Figure 2.16A, lane 8 vs. lane 7). As a control, northern blot gels hybridized with a probe binding to OppX (Figure 2.16B, upper gel) or to MicF (Figure 2.16B, lower gel) showed the two tagged RNAs are functional as each RNA could co-purify with the other.

While we could not specifically detect any protein other than Hfq in our silver-stained gel, a more sensitive method (such as mass spectrometry) could be ideal in identifying binding partners that are not as abundant as Hfq in a *Salmonella* lysate.

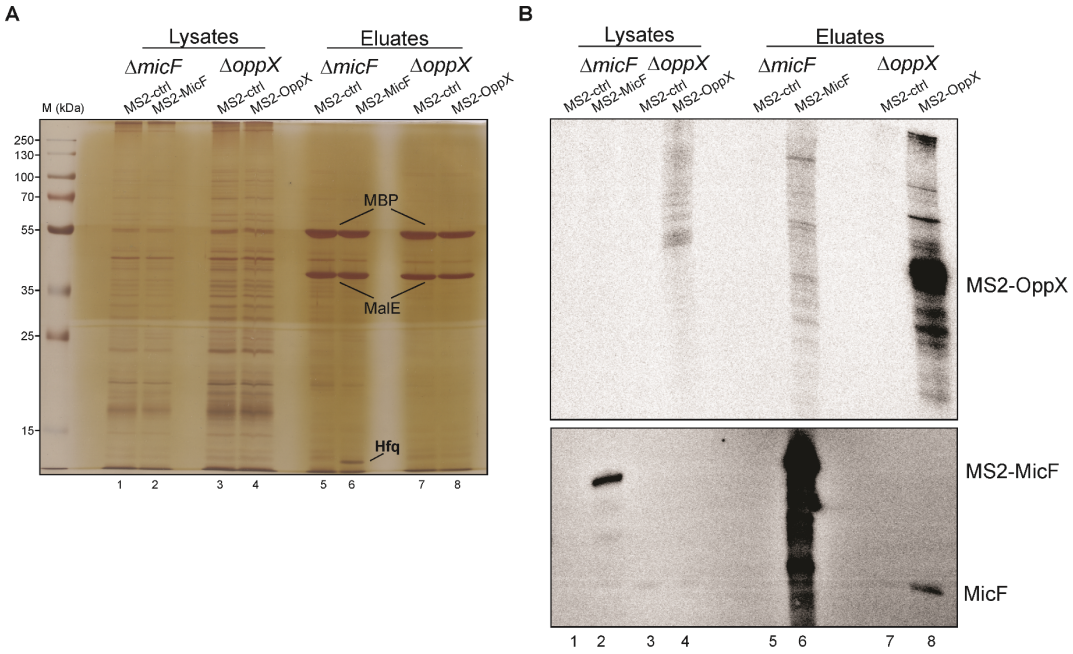

Figure 2.16 MS2-pulldown of MicF and OppX.
(**A**) Silver staining of protein samples isolated from *Salmonella* lysates and eluates of each strain indicated above the gel. The MicF pulldowns were performed in a Δ*micF* strain transformed with an MS2 empty plasmid (MS2-ctrl) or an MS2-tagged MicF constitutively expressing plasmid (MS2-MicF). The OppX pulldowns were performed similarly to the MicF ones. Relevant protein bands are indicated with black lines on the gel. (**B**) RNA samples extracted from the same samples indicated in (**A**) and loaded on a 6% PAA northern blot gel. Hybridization of the OppX probe (upper gel) and of the MicF probe (lower gel).

2.2.2.7 OppX indirectly modulates the expression of *ompF*

Our experiments suggested that OppX acted as a selective RNA sponge of the MicF sRNA. To prove this, we would expect to see OppX-dependent altered expression of MicF targets. We tested this hypothesis in the early stationary phase of growth, when absence of MicF (Δ*micF* strain) causes a pronounced upregulation of the major *ompF* target at both mRNA and protein levels, as compared to a WT *Salmonella* strain (Figure 2.17A, lane 1-4). In support of our model, genetic inactivation of OppX (by a chromosomal deletion of the 25 nt-long MicF binding site on the *oppA* 5'UTR) resulted in decreased levels of *ompF* mRNA and OmpF protein (Figure 2.17A, lanes 5). This result suggests that in this genetic background MicF transcripts are no longer sequestered by OppX. Reciprocally, a plasmid-driven overexpression of OppX-S increased *ompF* mRNA and protein levels (Figure 2.17A, lanes 6).

Importantly, the OppX-dependent effect was lost when the same experiments were performed in a Δ*micF* background, firmly reinforcing the idea that the observed effect originates from modulation of MicF activity (Figure 2.17B).

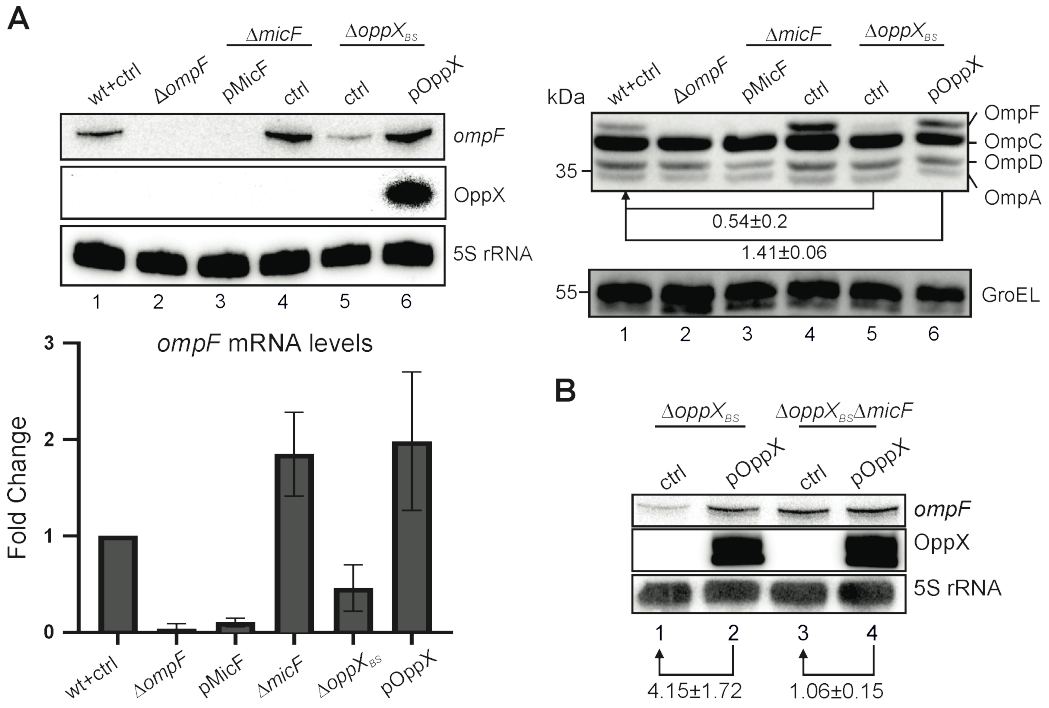

Figure 2.17 OppX modulates *ompF* expression in a MicF-dependent manner.
(**A**) Northern and western blot detection of *ompF* mRNA and protein expression upon chromosomal deletion of OppX (Δ$oppX_{BS}$) or overexpression from a high-copy plasmid (pOppX). Δ*ompF*, Δ*micF* +/- MicF overexpression strains were used as controls. Mean values and ±SD of three independent experiments are plotted below the gel. (**B**) Northern blot detection of *ompF* mRNA in a Δ$oppX_{BS}$ or Δ$oppX_{BS}$Δ*micF* background carrying a control plasmid (ctrl) or the pOppX. Mean values and ±SD of three independent experiments are indicated below the gel with arrows. 5S rRNA and GroEL were detected as loading control for the northerns and for the westerns, respectively.

Overexpression of OppX did not affect the MicF RNA levels (Figure 2.7D, lane 7). However, previously reported sponges in the Hfq-mediated network antagonize their target sRNAs by decreasing their steady-state level and stability (Figueroa-Bossi et al., 2009; Melamed et al., 2016; Miyakoshi et al., 2015). Nevertheless, alteration in sRNA stability can be masked by regulatory feedback of the same sRNA expression (Hoyos et al., 2020). Thus, we measured RNA stability through a rifampicin assay to assess potential effects of OppX and MicF on each other's cellular half-life. In this setting, neither deletion of the sRNA nor of the sponge affected the half-life of the

other (Figure 2.18A, upper and middle blot). For control, we measured the *ompF* mRNA half-life, and observed a reproducible reduction of its stability in the absence of OppX, i.e., when more free MicF is available for repressing this mRNA target (Figure 2.18A, lower blot). This suggests that at steady state conditions, a considerable fraction of MicF is sequestered and stored as an RNA complex with OppX, in order to titrate the MicF sRNA away from its main target, the *ompF* mRNA.

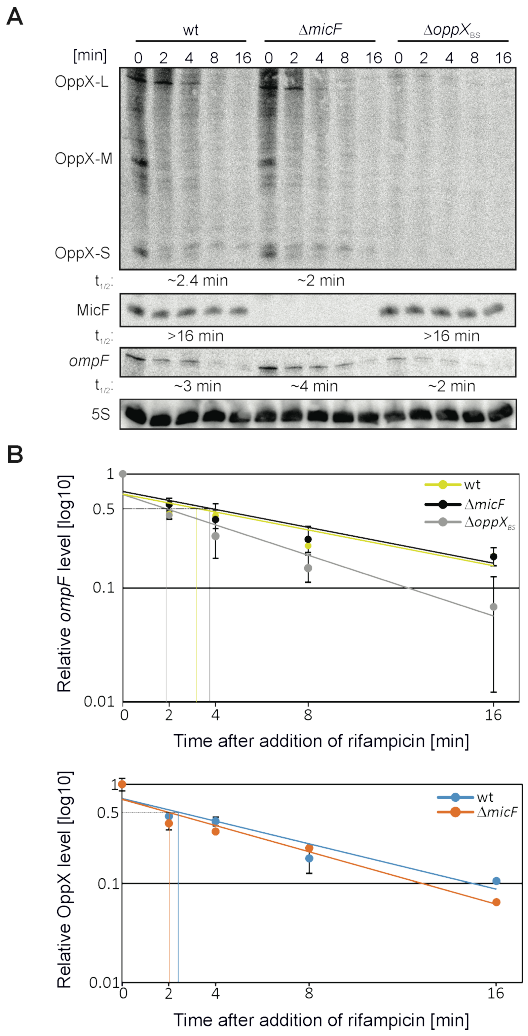

Figure 2.18 Half-lives of MicF, OppX and *ompF* with rifampicin assay.
(**A**) Northern blot analysis of a rifampicin experiment with hybridizations shown for OppX, MicF and *ompF* RNAs. Blot representative of three independent experiments. Mean values and ±SD are plotted in (**B**) for *ompF* (upper graph) and OppX (lower graph). Half-lives were estimated by fitting trend lines and extrapolating the x-values corresponding to 50% of the remaining RNA. Values are indicated below each gel.

2.2.2.8 OppX provides growth advantage in nutrient-poor media

The regulatory system involving OppX, MicF and *ompF* was identified in RIL-seq and supported by validation experiments in ESP (OD_{600} 2.0 in LB), suggesting the axis might fine-tune nutrient intake when bacterial cells need to optimize their uptake capacity in a nutrient-scarce environment. To determine whether deletion of OppX could confer a growth disadvantage in nutrient-poor media, we measured the OD_{600} over 24 hours in a WT and a Δ*oppX*$_{BS}$ strains. Strains deleted for MicF (Δ*micF*) and *ompF* (Δ*ompF*) were used as controls. All the strains were grown in rich medium (LB) and in mineral minimal medium (M9) supplemented with 0.2% casamino acids as the solely carbon source. As expected, no growth phenotype could be detected when strains were grown in LB. However, this was also the case for growth in the nutrient-poor media (Figure 6.7A). We hypothesized that the absence of a visible phenotype could be caused by redundancy among OMPs. Therefore, we sequentially deleted all the other OMPs of *Salmonella* (except for OmpD, given that intermediate deletion strains of this porin could not be generated) and measured growth. Neither single nor combined deletions of any other OMP led to a measurable growth defect with or without *ompF* or MicF (Figure 6.7B-C). We did not attempt to measure growth with the additional *oppX* deletion background, as an OppX-related growth defect would only be measurable if an *ompF* deletion resulted in a growth defect itself.

Nevertheless, we generated a *Salmonella* strain carrying simultaneous deletions of all the OMPs except for *ompF* (named as Δ*omps*) and used it as an isogenic control for strains carrying the Δ*omps* background together with a Δ*ompF, ΔmicF or* Δ*oppX*$_{BS}$ allele. Interestingly, a MicF deletion in Δ*omps* background provided a growth advantage in LB when compared to the Δ*omps* isogenic control (Figure 2.19A). None of the other mutants revealed any growth differences under these conditions. However, when the same strains were grown in nutrient poor media, both Δ*ompF* and Δ*oppX*$_{BS}$ in the Δ*omps* background showed a growth defect (Figure 2.19B). Collectively,

these results prove that when bacteria have OmpF as the solely available porin for nutrient uptake, the regulatory roles of OppX and MicF become important for assisting *Salmonella* growth in a nutrient-poor environment.

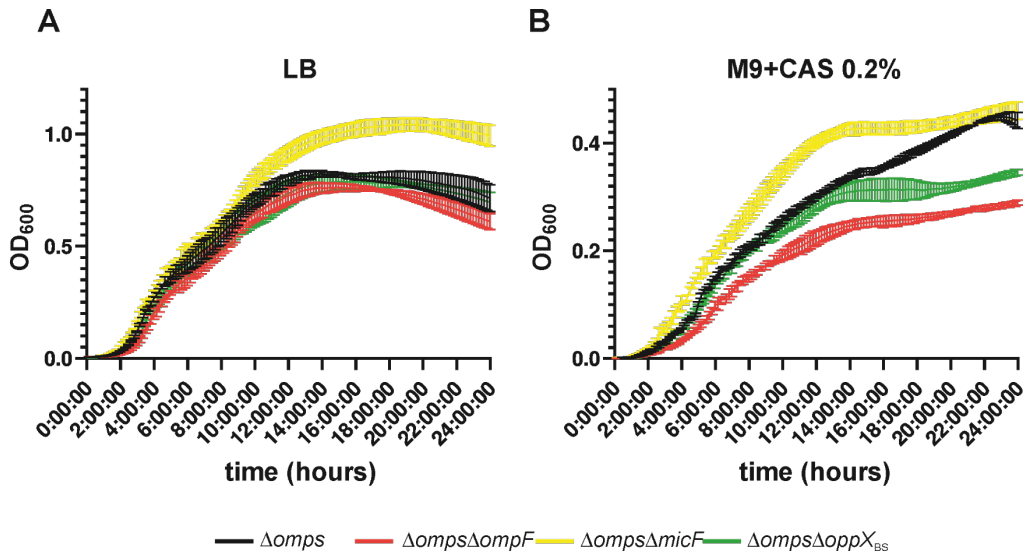

Figure 2.19 Growth phenotype of OppX in nutrient-scarce media.
(**A**) Growth curves of Δ*omps* strains with or without deletion of *ompF, micF* or $oppX_{BS}$ in LB media or in M9+0.2% casminoacids (**B**). Mean values and ±SD of six independent experiments (LB) or four independent experiments (M9+CAS 0.2%) are plotted.

2.2.2.9 The OppX axis is conserved beyond Salmonellae

First evidence that this regulatory axis could exist beyond *Salmonella* was highlighted by the remarkable base conservation of both MicF sRNA and the region MicF is predicted to bind on the *oppA* 5'UTR (Figure 2.7A and Figure 6.4). To prove the functionality of this axis in another member of the Enterobacteriaceae family, we expressed the OppX sequence of *E. coli* ($OppX_{EC}$) from a high-copy plasmid in the MG1655 K-12 strain. By measuring *ompF* levels by northern and western blots we could conclude that, in line with the results obtained in *Salmonella,* OppX expression results in increased *ompF* RNA and protein levels, respectively, when compared to an empty plasmid control (Figure 2.20A). We additionally extended our search for the *opp*-OppX-MicF-*ompF* axis to other families of the Enterobacterales, of which Enterobacteriacea is a member. Conservation in

the families of Erwiniaceae and Yersiniaceae was found (Figure 2.20B). Interestingly, similarly to what was observed for *Salmonella,* IntaRNA could predict potential base pairing between MicF and both *oppA* and *ompF* in the genomes of those Enterobacterales simultaneously bearing all three genes (Figure 2.20B, right hand part).

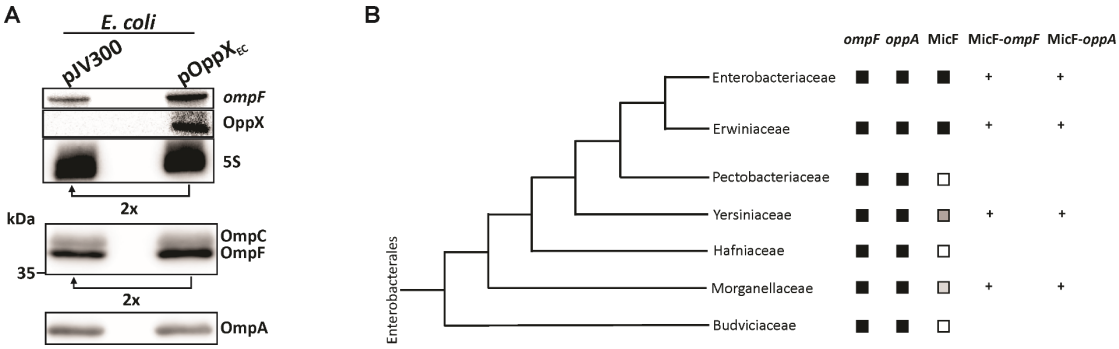

Figure 2.20 Conservation of the *opp*-OppX-MicF-*ompF* axis in Enterobacterales.
(**A**) Northern and western blots (top and bottom, respectively) of lysates of *E. coli* expressing a control plasmid (pJV300) or an overexpression of the *E. coli* OppX sequence (pOppX$_{EC}$). 5S rRNA and OmpA were used as loading controls. (**B**) Conservation of the OppA and OmpF coding sequences and of MicF sRNA within Enterobacterales. Potential base-pairing (+) between MicF and *ompF* (left hand side) or *oppA* (right hand side) was predicted by IntaRNA. The whole length of the mRNA sequences was used to investigate potential base pairing areas.

2.2.3 Discussion

As one of the most investigated bacterial model organisms, a great multitude of studies has focused on the identification and characterization of the targetome of regulatory RNAs in *E. coli* and *Salmonella*. However, in *Salmonella*, these studies mostly employed traditional techniques, excluding the possibility to identify RNAs and RNA interactions that did not match the discovery criteria for such "old" methods.

In this sub-chapter, we employed RIL-seq in *Salmonella* Hfq to draft a global map of RNA-RNA interactions in this bacterium. After adaptation and optimization of the published RIL-seq protocol (Figure 2.1 and Figure 2.2, (Melamed et al., 2018)), we obtained successful and reproducible interactome datasets (Figure 2.3 and Figure 2.4, Table S2). From a first hazy analysis of the data, a few surprising observations could be made: while the

most abundantly identified hybrids involved an sRNA and different regions of an mRNA, a high number of interactions between two sRNAs populated the dataset (Figure 6.1, Table S2B-C). This high representation of sRNA-sRNA interactions was unexpected, as this type of base-pairing often leads to degradation of either of the two RNAs, making them easily discoverable with traditional approaches. Several scenarios could lead to the formation of sRNA-sRNA duplexes in bacteria:

- The two sRNAs bear complementary sequence, they bind to each other which triggers degradation of one or both of them. Several examples belong to this scenario: in *E. coli* the *rbsB*-derived sponge, RbsZ, binds and induces RNase III-mediated degradation of the RybB sRNA, causing upregulation of the RybB targetome (Melamed et al., 2020); in most Enterobacteriaceae the *gltIJKL*-derived RNA SroC is an independent Hfq-bound sRNAs that base pairs with GcvB resulting in RNase E degradation of the latter and deregulation of the entire GcvB regulon (Miyakoshi et al., 2015); in *E. coli,* CLASH showed that the two Hfq-dependent sRNAs, CyaR and ArcZ, bind to each other causing degradation of CyaR and consequent deregulation of its targetome (Iosub et al., 2020).
- The two sRNAs bear complementary sequence, they bind to each other without inducing any degradation. The EHEC GcvB-sponge, AgvB, (Tree et al., 2014) and the 5'UTR-derived ChiZ, sponging the ChiX sRNA (Adams et al., 2021), both belong to this class. However, a tRNA-derived sponge (not considered an Hfq-binding sRNA), namely 3'ETS^{leuZ} binds both RyhB and RybB sRNAs affecting their targetome translation without evidence for sRNA degradation (Lalaouna et al., 2015).
- They simultaneously bind to the Hfq hexamer without forming a duplex, but their ends are close enough to allow proximity ligation. This scenario remains hypothetical as it has not yet been investigated before. However, in all Hfq RIL-seq datasets included in the present chapter, as well as in the published *E. coli* RIL-seq (Melamed et al., 2016), very abundant hybrids could be detected between the MgrR and the GcvB sRNAs (Table

S2B-C). These two sRNAs do not show any complementarity according to IntaRNA predictions. Nevertheless, it has been shown that MgrR is a target of the GcvB-sponge, SroC (Acuña et al., 2016). It is tempting to speculate that despite not interacting, MgrR and GcvB could be brought close to each other from simultaneous interactions with SroC. Further analysis of other unusual sRNA-sRNA chimeras is needed to establish whether this observation is a stand-alone event or if it underlines a widespread behaviour of sRNAs that take part to a common regulon.

A second obvious observation was the high representation in the dataset of interactions involving sRNAs and/or targets known to be part of the envelope regulon (Figure 2.5, Table S2B-C and S3). Maintenance of envelope homeostasis by the largest class of riboregulators, the Hfq-associated sRNAs, is widely recognized (Fröhlich and Gottesman, 2018). A great part of these regulatory networks was discovered through combined traditional approaches of sRNA-deletion/overexpression followed by global transcriptomics. With focus on interactions involving envelope-related transcripts, this chapter showcases the potential of RIL-seq in the identification of a regulatory RNA involved in envelope control overlooked by traditional methods. This discovery underscores the importance of studying bacterial RNA regulons at a global scale. The regulatory RNA this sub-chapter has focused on, OppX, was selected because of its outstanding prevalence in the total detected chimeras of the well-characterized MicF sRNA (Figure 2.5A). OppX encompasses the relatively long 5'UTR of the *oppA* gene, the first one transcribed in the *opp* operon (Figure 2.7A).

The MicF sRNA is considered one of the first regulatory RNAs to be discovered and characterized (Andersen et al., 1987; Coleman et al., 1984). Both its own regulation (Aiba et al., 1987; Andersen et al., 1987; Chou et al., 1993; Deighan et al., 2000; Takayanagi et al., 1991) and its targetome, beyond the classic *ompF* mRNA, have been studied in detail (Corcoran et al., 2012; Holmqvist et al., 2012). Several transcription regulators, including the global regulator Lrp, controlling MicF transcription have been reported (Gama-Castro et al., 2007; Holmqvist et al., 2012; Ramani et al., 1994; Takayanagi et

al., 1991). However, factors that regulate MicF activity at the post-transcriptional level remained unknown. In this sub-chapter, the first such regulator of MicF activity, OppX, was identified and described. Being the primary RNA interactor of MicF in the studied conditions, without the knowledge of OppX activity, MicF biological function could have not been fully uncovered.

sRNAs that bind the 5'UTR of a coding gene traditionally cause alteration in the expression of the downstream gene, at either the level of translation or RNA stability (Hör et al., 2020). The outcome of MicF-OppX interaction was investigated in order to determine whether such far upstream base pairing (Figure 2.7A) could produce any significant effect on the *oppA* CDS. Combined molecular and *in vivo* approaches were used (Figure 2.7B-E) and confirmed that binding of MicF to the *oppA* 5'UTR encompassing OppX did not produce any alteration in translation, nor did it cause degradation of the *oppA* coding arm, unlike with the known *ompF* target. The discovery of OppX from a region far upstream of a coding gene start codon indicates that 5'UTR-derived regulatory RNAs can be easily overlooked due to insufficient transcriptome annotation, particularly for understudied organisms. Additionally, OppX identification challenges the dogma that sRNAs that pair in the 5' region of an mRNA denotes the latter as the regulated target. Interactome data like RIL-seq are more likely to enable the discovery of more potential regulatory RNAs arising from 5'UTRs.

An important question when studying UTR-derived non-coding RNAs is how they are generated as independent transcripts. The evidence that the *oppA* 5'UTR results in stable fragments *in vivo* (Figure 2.8) prompted us to ask whether the OppX isoforms were transcribed from independent promoters (Guo et al., 2014) or the result of a ribonuclease activity (Chao and Vogel, 2016). The combination of dRNA-seq (data generated from the lab, unpublished), TIER-seq (Chao et al., 2017) and *in vivo* experiments using an RNase E defective *Salmonella* strain suggested that the long *oppA* 5'UTR is subjected to extensive processing from this major ribonuclease (Figure 2.9).

Additionally, preliminary experiments highlighted two more factors involved in the biogenesis of OppX: i) a second enzyme, the 3'exoribonuclease PNPase seems to play a role in the turnover of some, but not all, OppX isoforms (Figure 2.10); ii) high temperature makes OppX heat-labile when a terminator structure at the 3'-end is missing (Figure 2.9B and Figure 2.10).

What are the binding features of the MicF-OppX interaction? Being the most abundant pair in the RIL-seq suggested that a possible duplex between these two RNA species must be robust enough to be recovered from a multi-step protocol. *In silico* predictions (Figure 2.6D) as well as *in vitro* binding assays (Figure 2.11, Figure 2.12 and Figure 2.13) demonstrated MicF and OppX specifically interacted with an affinity within the nanomolar range (k_d= 25 nM), typical for bacterial RNA regulators. Given the comparable affinity of MicF to its *ompF* target (Figure 2.13B) and the similar abundance of these three RNAs in ESP-growing *Salmonella* (Kröger et al., 2013), we reasoned that fluctuation of the expression of either of the three players would result in a biological outcome with respect to MicF-targets expression, making OppX an authentic competing RNA (or "RNA sponge") of MicF. The evidence for the formation of a stable MicF-OppX complex *in vivo* (Figure 2.14 and Figure 2.15), together with lack of degradation of MicF (Figure 2.7D and Figure 2.18), makes OppX an unusual RNA sponge that does not alter the stability of its target sRNA (point #2 described in second paragraph of the Discussion). Second only to ChiZ (Adams et al., 2021), OppX is a stand-alone example of 5'UTR-derived sponge identified so far in bacteria.

The outstanding sequence conservation of MicF, *ompF* and OppX (Figure 2.7A and Figure 6.4) suggested that this regulatory axis served an important physiological role with respect to envelope crosstalk, adjusting OM permeability to oligopeptide transport across the IM. Genetic screens have addressed the essentiality of porins, in particular OmpC and OmpF, for transport of small oligopeptides (Andrews and Short, 1985). Given the difference in pore size among porins (e.g. OmpF has a larger pore than

OmpC (Vergalli et al., 2020)) and the co-production of OppX with the Opp oligopeptide importer, this axis would allow the cell to rapidly increase translation of *ompF* when larger permeability is needed, for instance during starvation (Andrews and Short, 1986; Calvo and Matthews, 1994). Microenvironments through the gastrointestinal tract can be marked by differences in osmolarity (due to bile salts) and availability of peptides as the primary nutrient source. In enteric bacteria exposed to such dynamic conditions, OppX could serve as a supplementary regulator to revert MicF-mediated inhibition of OmpF and promote a more efficient nutrients uptake. Experimentally, we could demonstrate that chromosomal deletion or plasmid-driven overexpression of the OppX sequence when *Salmonella* was grown in early stationary phase (when nutrients availability begins to drop) adjust the expression of *ompF* at both the RNA and protein levels (Figure 2.17A). This effect was abolished with MicF deletion, confirming that OppX functions through MicF (Figure 2.17B). Additionally, in bacteria grown in a media containing casamino acids as the solely carbon source, OppX deletion conferred a growth defect, when *ompF* was the only porin available (Figure 2.19). The need for performing these phenotypic experiments in a background deletion of other porins indicates an expected redundancy among these conserved bacterial transporters (Figure 6.7, (Vergalli et al., 2020)).

Collectively, the experiments conducted in this sub-chapter provided evidence for a regulatory axis (Figure 2.21) conserved in several Enterobacteria (Figure 2.20B) and functional in both *Salmonella* and *E. coli* (Figure 2.20A), that allows cross-talk between permeability of the OM (OmpF regulation) and transport capacity at the IM (OppX and *opp*), through regulation of the Hfq-dependet sRNA MicF.

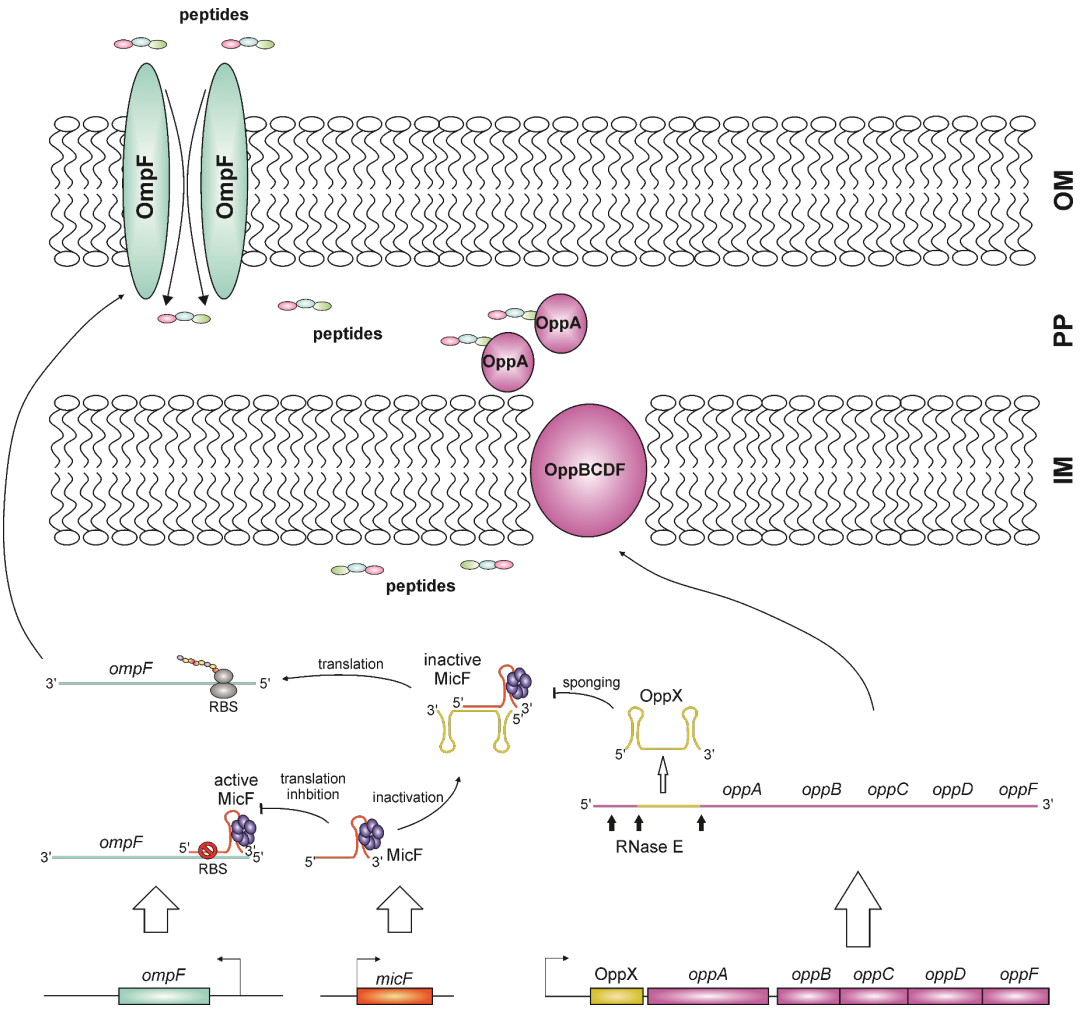

Figure 2.21 Depiction of the *opp*-OppX-MicF-*ompF* functional model.
Transcription of *ompF* and of the *opp* operon and their subsequent translation allows the OM to utilize OmpF for peptide transport into the periplasm, and the OppABCDF system for internalization into the cytosol. Transcriptional levels of MicF and OppX and their ability to bind to each other, dictates translation of *ompF*. Adapted from Matera et al., in revision.

Beyond the findings related to envelope regulation, performing RIL-seq in a pathogenic bacterium like *Salmonella* allows to better investigate how sRNAs connect the core genome with virulence functions encoded in pathogenicity islands. The growth conditions utilized in the ESP RIL-seq induce expression of the SPI-1 locus, encoding many effectors and the T3SS-1 apparatus necessary for their secretion into the host. Encoded in the SPI-1 region are also many sRNAs, such as InvR. Previous studies focused on target identification of this sRNA identified the *ompD* gene as its sole target, excluding any InvR-mediated effects on the expression levels of any SPI-1-encoded gene (Pfeiffer et al., 2007). Nevertheless, our RIL-seq data highlighted chimeras between InvR and many mRNAs encoded in the SPI-

1 locus (Table S2B-C). Additionally, *in silico* analysis predicted these interactions to mainly occur within the coding sequence of these genes, an otherwise inaccessible region in actively translating mRNAs. While in its infancy, these observations set the basis for discovering new biological mechanisms of sRNA regulation in pathogens.

2.3 RIL-seq in SPI-2 conditions and inside infected macrophages

2.3.1 Establishing RIL-seq in SPI-2 minimal media

The dataset of RIL-seq experiments performed in sub-chapter 2.1 originated from *Salmonella* grown in standard laboratory conditions, i.e., ESP in liquid rich medium (LB). While the *Salmonella* expression profile in growth phase resembles that of the invasion-program given the activation of the SPI-1 island (Kröger et al., 2013), we additionally aimed to perform RIL-seq in a specific medium that triggers the expression of the SPI-2 island (a.k.a. SPI-2 medium). Its nutrient-poor and low Mg^{2+} compositions mimic the harsh intracellular host environment (Kröger et al., 2013; Srikumar et al., 2015).

WT and *hfq*::3×FLAG strains were grown in SPI-2 medium to an OD_{600} ~0.3, a stage where the intracellular replication program is induced, and an amount of 100 ODs was harvested for cross-linking. From this step onwards, the procedure was applied analogously to RIL-seq as described in 2.1. Recovered RNA samples were checked on an RNA pico bioanalyzer for QC (Figure 2.22A-B). As for the OD_{600} 2.0 dataset, RNA peaks around 100-200 nt could be detected for all the samples (two WT and two *hfq*::3×FLAG), with the WT measuring less RNA abundance compared to the Hfq ones. Nevertheless, a lower amount of RNA seemed to be recovered from the Hfq1 replicate given the lower FU in the bioanalyzer plot (Figure 2.22, left panel, Hfq1). However, the ladder peak at 25 nt used for the quantification also gave lower FU values in this sample when compared to all the other plots (FU=5 vs FU=15, 30 and 25 for WT1, WT2 and Hfq2, respectively), making quantification of the RNA inaccurate for this specific sample. Nevertheless, cDNA library preparation was successful for all four samples, as a smear of amplicons measuring 130-200 bp in size became visible after 20 cycles of cDNA PCR amplification (Figure 2.22D). As expected, the intensity of amplified cDNA for the WTs was lower than the two *hfq*::3×FLAG samples. Libraries were pooled and sequenced as previously described and mapped to the *Salmonella* genome with the chimera-discovery RIL-seq pipeline.

Statistics of mapped fragments (single and chimeric) highlight the expected higher rate of chimera formation in the Hfq samples (~6%, Table S1B) compared to the WT (~0.15%, Table S1B).

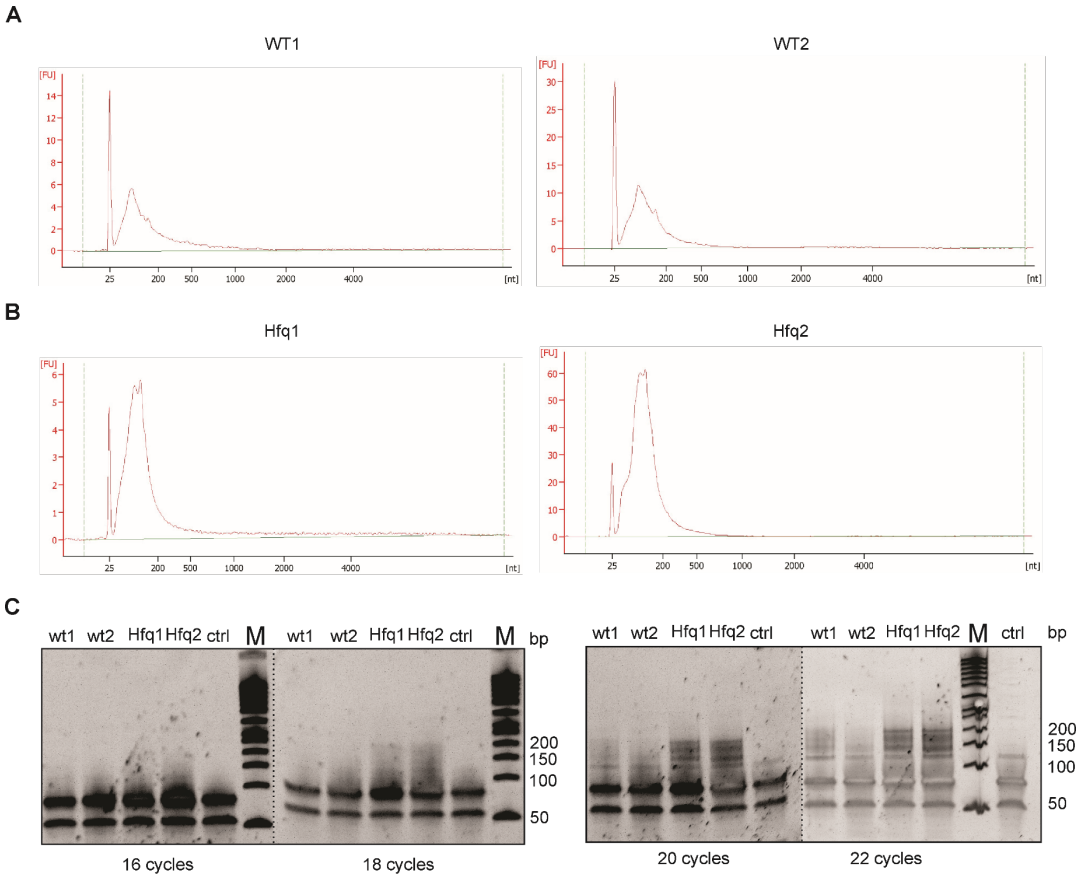

Figure 2.22 RNA profiles and cDNA libraries of the SPI-2 RIL-seq.
(**A,B**) RNA pico bioanalyzer profiles of the two replicates for the WT (WT1 and WT2) and the *hfq*::3×FLAG samples (Hfq1 and Hfq2). (**C**) cDNA libraries of each respective sample amplified with 16, 18, 20 or 22 cycles. M=50 bp DNA ladder.

2.3.2 Data overview of the SPI-2 RIL-seq

At a first glance, the interactome datasets in SPI-2 and standard growth conditions look similar. The *hfq*::3×FLAG samples show an enrichment in sRNAs, CDSs and UTRs in the chimeras over the WT control (Figure 2.23A), where sRNAs occupy mostly the second part (RNA2) of the chimeric fragments and CDS and 5'UTRs the first part (RNA1) (Figure 2.23B). Here, the sRNA ligation bias seems to be even more pronounced (~80% of the sRNA is positioned in RNA2, (Figure 2.23C)). Interestingly, 3'UTRs and

IGRs are more represented in this dataset, compared to the standard RIL-seq (15-25% of all Chimeras, (Figure 2.23A)). This observation could be partially explained by the fact that the *Salmonella* annotation for sRNAs used in the mapping has been developed over the years from datasets of experiments performed in standard growth conditions. Therefore, sRNAs or other genomic locations specifically transcribed in SPI-2 conditions could still be unannotated. However, the majority of the sequenced fragments successfully mapped to annotated regions.

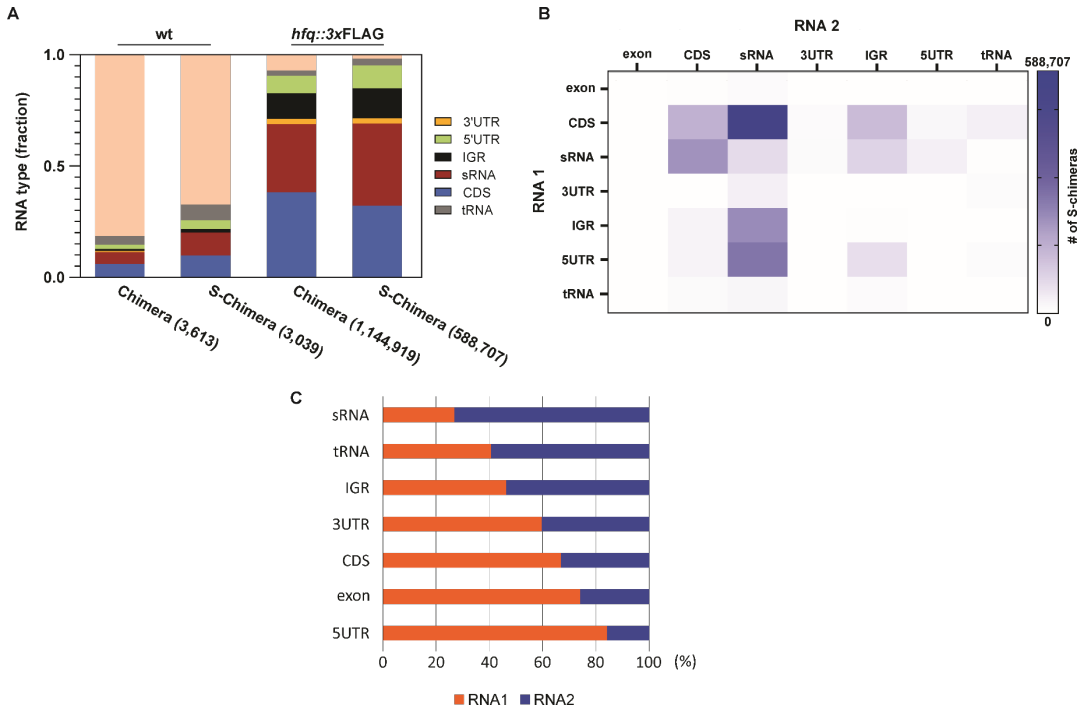

Figure 2.23 Overview of the RNAs detected with SPI-2 RIL-seq.
(**A**) Representation of RNA type fractions detected in the WT and *hfq*::3×FLAG strains (the values shown were determined from the WT1 and Hfq1 replicates). (**B**) Heat-map showing the distribution of chimeric fragments within RNA types and according to position on the hybrid read (RNA1=first read, RNA2=second read). (**C**) Ratio of positional bias (RNA1 vs. RNA2) for each RNA type.

2.3.3 The SPI-2 dataset reveals known and new sRNAs involved in infection

A deeper comparison between the SPI-2 dataset and the ESP one highlights important features with respect to biology. The repertoire of sRNAs involved in RNA interactions and therefore detected as chimeras in RIL-seq differs substantially. In ESP, ArcZ, MgrR, InvR, CyaR, CpxQ and at a moderate extent, PinT, are the most represented sRNAs in S-chimeras (Figure 2.24A, left side). Conversely, in the SPI-2 dataset, PinT, MgrR, ChiX, SdsR and IsrE dominate the sRNA network, while ArcZ seems to be having only moderate influence (Figure 2.24A, right side). Additionally, Spot42 (Spf) and STnc2080 specifically constitute a big fraction of sRNA interactions in the SPI-2 data, given that no hybrids involving these two sRNAs were detected in ESP (Table S2). These observations reflect the evidence the "sRNAome" of *Salmonella* bears high plasticity, varying according to stresses and needs for the bacteria.

2.3.3.1 PinT sRNA: the timer of virulence revealed by RIL-seq

PinT, initially known as STnc440 (Kroger et al., 2012; Pfeiffer et al., 2007), has been only recently studied in more detail. It is one of the few *Salmonella*-specific sRNAs, not present in *E. coli*, and acquired through HGT. Its expression is dictated by the PhoPQ regulators (Westermann et al., 2016), a TCS activated in the intracellular stages of *Salmonella* infection. PinT is expressed during growth in LB medium, therefore also in ESP, although its expression peaks once *Salmonella* replicates inside the host (or in media that mimic this condition, like the SPI-2 media) (Correia Santos et al., 2021; Kim et al., 2019; Westermann et al., 2016). In our RIL-seq data, PinT chimeras were detected in the SPI-2 conditions where this sRNA dominated the network (Figure 2.24A, right side) and in ESP only to a lower extent. Interestingly, the targetome of PinT in these two conditions differed drastically, except for the *ytfK* gene which scored as the top PinT pair in both datasets (Figure 2.24B). While YtfK function remains unknown, its expression is regulated by both SPI-1 master regulators, HilD and HilC

(Petrone et al., 2014). In the SPI-2 network, PinT was predicted to form interactions with known targets like *steC* (Correia Santos et al., 2021) or putative targets predicted through other methods (e.g. *ecnB,* (Westermann et al., 2016)) (Figure 2.24B). Additionally, our SPI-2 RIL-seq suggests PinT engages in duplex formation with the horizontally-acquired and PhoP-induced gene *ugtL* (Figure 2.24B), known to regulate PhoQ itself and promote *Salmonella* virulence ((Choi and Groisman, 2017, 2020)). Despite the fact that growth to ESP in LB should mimic the SPI-1 program, the top targets of PinT in this condition did not match those known to be regulated during invasion, such as *sopE, sopE2, hilA* and *rtsA* (Kim et al., 2019; Westermann et al., 2016). This discrepancy could have several explanations: i) the four cited PinT targets are known to undergo mRNA degradation upon PinT binding, making it more difficult for these chimeras to be detected via RIL-seq; ii) while ESP in LB resembles in part the invasion program of *Salmonella*, other specific ways to induce SPI-1 expression are more suitable (Lee and Falkow, 1990; Sittka et al., 2008) and would likely provide more relevant information with respect to regulation of invasion of the host.

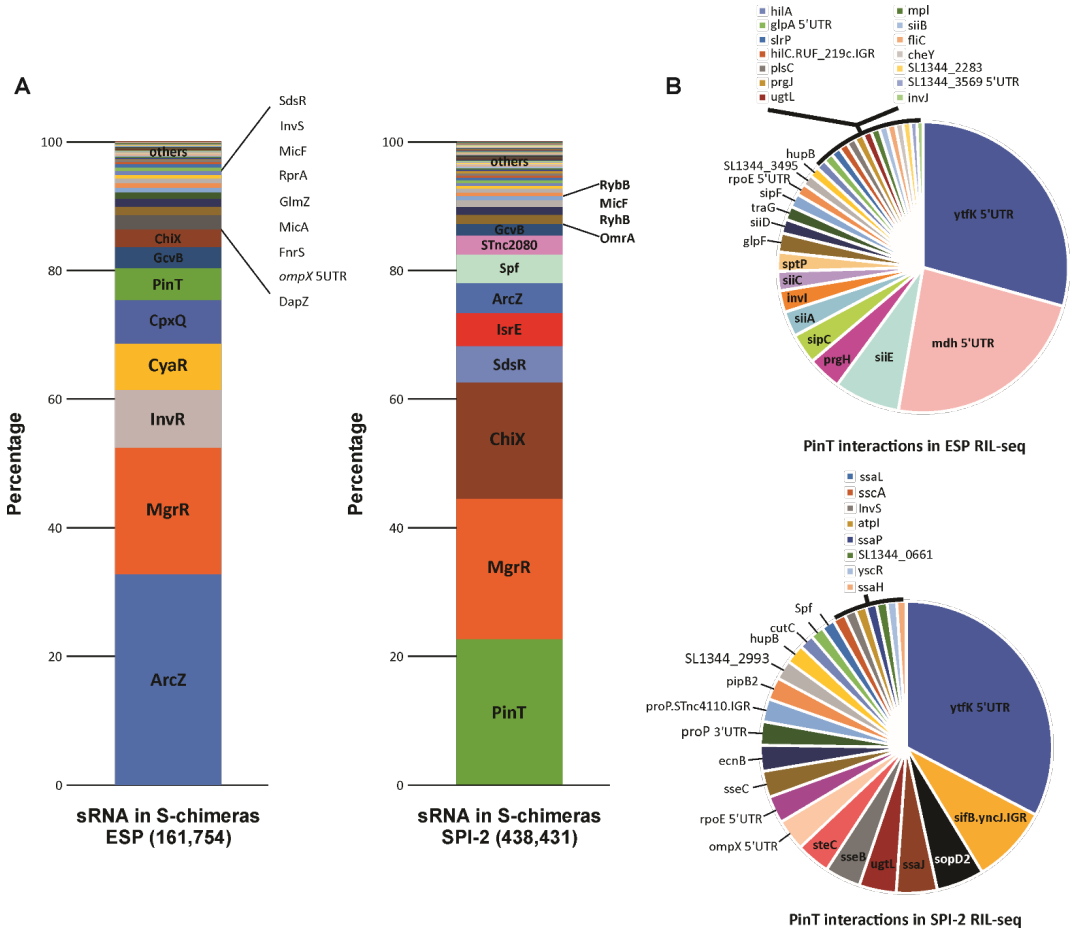

Figure 2.24 sRNAs detected in SPI-2 and highlight on PinT chimeras.
(**A**) Comparison of sRNA distribution in % between RIL-seq in ESP (on the left, plot taken from Figure 6.1) and RIL-seq in SPI-2. In brackets the total number of S-chimeric fragments is indicated. (**B**) Pie charts of PinT targets in ESP and SPI-2 and their relative abundance to the total of PinT S-chimeras in each condition.

2.3.3.2 MgrR sRNA: a core genome sRNA involved in infection

Like PinT, MgrR expression is regulated by the PhoPQ TCS (Colgan et al., 2016; Moon and Gottesman, 2009). However, MgrR is encoded in the core genome of both *E. coli* and *Salmonella*, and its function has been only studied in *E. coli* thus far. MgrR binds and regulates genes involved in sensitivity to antimicrobial peptides (AMPs), such as *eptB*, involved in lipopolysaccharide (LPS) modification. *E. coli* strains carrying an *mgrR* deletion were 10 times more sensitive to the AMP polymyxin B (Moon and Gottesman, 2009). In another *E. coli* study, MgrR was reported to downregulate the *pitA* gene that functions as a symporter with Mg^{2+} to promote accumulation of this cation

(Yin et al., 2019) when in low concentration (e.g. for *Salmonella* intracellular milieu or in SPI-2 media).

Analysis of the MgrR interactions detected by RIL-seq in both growing conditions suggests this sRNA might function in a similar way in *Salmonella.* In ESP RIL-seq, MgrR formed hybrids with the *pitA* 5'UTR (top target), *ygdQ* (target predicted in (Moon and Gottesman, 2009)) and a host of other genes such as the sRNA GcvB, the tRNA modification enzyme *trmB* and the SPI-1 master regulator *hilC* (Figure 2.25A). In SPI-2 conditions, in addition to some of the targets identified in ESP, MgrR formed hybrids with the two PhoPQ-regulated genes, *pagC* and *virK*, involved in outer membrane vesicle (OMV) secretion (Dehinwal et al., 2021; Lu et al., 2020) and resistance to cationic AMPs (Detweiler et al., 2003; Matamouros and Miller, 2015), respectively (Figure 2.25B). Interestingly, IntaRNA predicted binding sites between MgrR and many of the RIL-seq targets (Figure 6.8). Specifically, MgrR uses two distinct regions (named as seed 1 and seed 2) to bind the 5'UTR of the *pitA* mRNA (Figure 6.8A), conversely, seed 1 alone is needed for binding to the coding sequence of *ygdQ* and a very upstream region of the *pagC* 5'UTR (Figure 6.8B-C) whereas seed 2 is required for binding the mRNA of *virK* (Figure 6.8D). Whether this seed-region specificity has any influence in the output of MgrR regulation and therefore in the biology of this network is still unknown, although this feature has been previously reported for other Hfq-dependent sRNAs (e.g. GcvB (Miyakoshi et al., 2015; Sharma et al., 2011)). Collectively, this evidence suggests MgrR is a master post-transcriptional regulator of survival inside the host, exerting its function through buffering the activity of PhoPQ-dependent genes (Figure 2.25C).

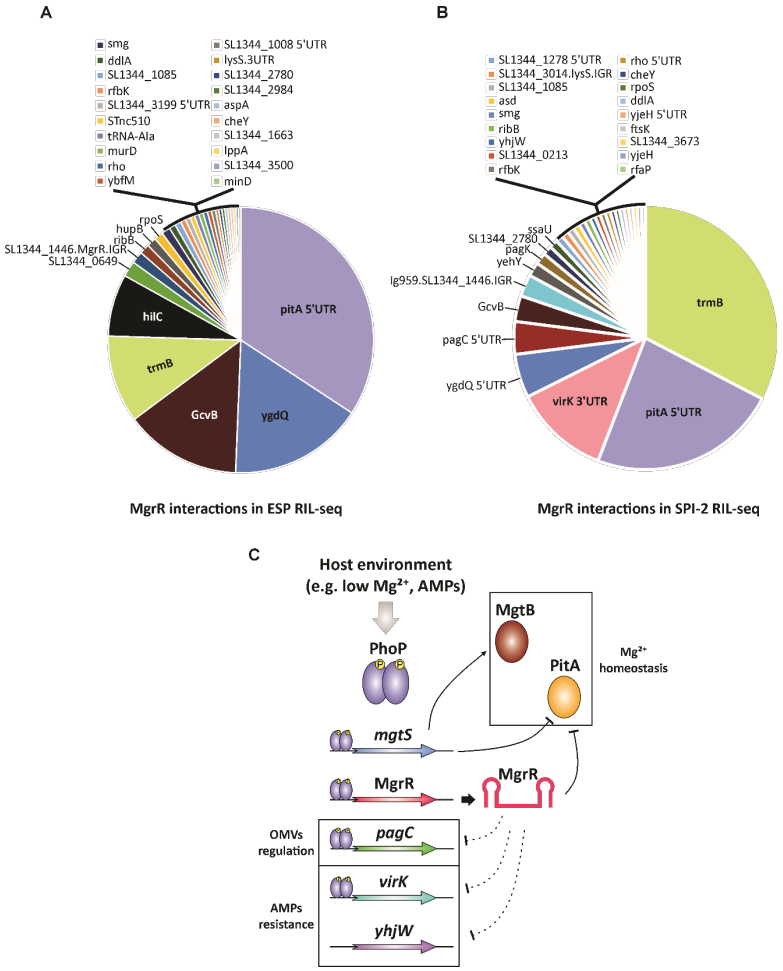

Figure 2.25 Highlight of MgrR S-chimeras and comparison between ESP and SPI-2. (**A**) Pie chart of MgrR interactions as detected from RIL-seq performed in ESP or in SPI-2 (**B**). (**C**) Cartoon showing the PhoPQ regulon and targets of MgrR (known target are indicated with a solid line, whereas RIL-seq predicted targets are indicated with a dashed line).

2.3.4 Protocol optimization for Hfq-IP with low bacterial input

The ESP and SPI-2 RIL-seq datasets discussed previously provided precious insights on the complexity and plasticity of the Hfq-mediated RNA networks. Nevertheless, while mimicking some of the *in vivo* lifestyles of *Salmonella*, these *in vitro* conditions bear intrinsic limitations. Therefore, we aimed at optimizing the RIL-seq approach in order to perform this method during a *Salmonella* infection experiment inside host macrophage cell lines. The original RIL-seq protocol requires at least 40 ODs (for *Salmonella* it equals to 40×10^9 cells) in order to perform a successful experiment. Recovering this number of bacteria from an infection assay would be impractical.

To identify the minimal cells required for a successful Hfq-IP, we performed RIL-seq with sequential cell dilutions (40×10^9, 1×10^9, 1×10^8 1×10^7 1×10^6) and stopped the procedure right before RNA extraction. The whole elution or just a small aliquot (for the sample with 40×10^9 cells as input) were separated on an SDS-PAGE gel and probed with the Flag antibody onto a western blot membrane.

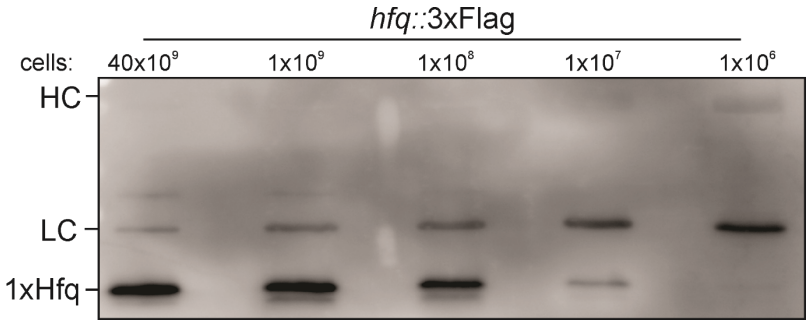

Figure 2.26 Hfq Co-IP performed with serial dilutions of bacterial cells.
Western blot showing Hfq protein after Co-IP with a Flag antibody. A 1:7 dilution of the $40x10^9$ sample was loaded, while the whole eluted fraction was loaded for the remaining samples. HC: high-chain; LC=low-chain; 1xHfq=Hfq monomer.

Figure 2.26 shows the Hfq monomer could be recovered and detected down to the 10^7 cells as bacterial input. Although a faint band could be also detected in the 10^6 sample (Figure 2.26, last lane), we considered 10^7 as the minimal cell number for an *in vivo* experiment.

2.3.5 RIL-seq applied to intra macrophage-replicating *Salmonella*

Given the still high minimal bacterial cells requirement for performing a successful Hfq-IP (~10^7 cells), *in vivo* RIL-seq was performed by sampling infected macrophages with an m.o.i. of 50, at late time points (20 h p.i.). At this stage, intracellular *Salmonella* are surviving and replicating inside the host, a condition that allows recovery of a higher number of bacteria and it recapitulates the SPI-2 medium experiment enabling a comparison between the two datasets. Pilot infections followed by CFU (colony forming units) counts were performed to assess the amount of infection flasks needed for each strain (n=8 flasks). The experiments were performed testing two lysis

strategies in parallel (Figure 2.27): i) at 20 h p.i., infected cells were harvested from the flasks, pelleted, subjected to lysis followed by RIL-seq; ii) at 20 h p.i., infected cells were harvested and blandly lysed to release the intracellular bacteria that were recovered after a two-step centrifugation (details in Materials & Methods), followed by lysis and RIL-seq. For each lysis strategy, an infection with a WT *Salmonella* strain and an *hfq*::3×FLAG strain was conducted in a single biological experiment. In parallel, the standard ESP RIL-seq was performed as a positive control.

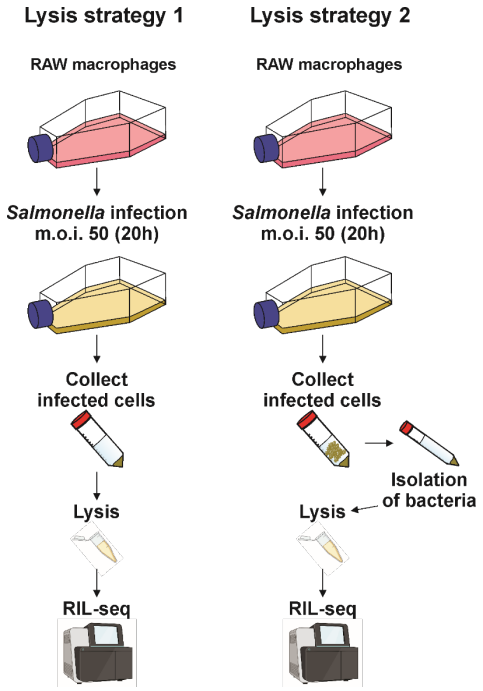

Figure 2.27 Experimental design of the two *in vivo* RIL-seq strategies.
On the left, representation of the experimental steps for strategy 1: i) infection of 8 flasks of RAW macrophages with an m.o.i. of 50 for 20 h; ii) collection of infected cells at the chosen time point with centrifugation; iii) Lysis of eukaryotic and intracellular bacterial cells followed by RIL-seq. On the right, representation as described for the left one with the addition of the "bacterial isolation" step.

A pilot experiment was initially performed to judge the possibility to IP the tagged Hfq in the undertaken approach. Loading the eluate onto an SDS-PAGE gel followed by FLAG-detection on a western blot membrane revealed that Hfq can be successfully pulled down with both lysis strategies (Figure 2.28), although strategy #1 led to a higher recovery of the Hfq monomer compared to strategy #2. Nevertheless, all samples were subjected to RNA extraction followed by bioanalyzer for QC (Figure 6.9). As expected,

the amount of recovered RNA was much lower compared to an *in vitro* RIL-seq experiment (~0.7-2 ng vs. 5-20 ng, respectively). Additionally, more RNA was recovered from the samples obtained with strategy #1, likely due to high contamination with host cell RNAs. All samples were eventually subjected to cDNA library preparation and paired-end sequencing. Analysis of the statistics after RNA sequencing enabled us to rule out any failed or poorly performing strategy. Indeed, only 5-10% of the RNA derived from both WT and *hfq*::3×FLAG samples of strategy #1 mapped to *Salmonella* genome from a total of 60-70M sequenced reads (Table S1C). Conversely, 34% and 17% of the *hfq*::3×FLAG and WT, respectively, obtained with strategy #2, mapped to the *Salmonella* genome. Additionally, the efficiency of chimera detection in the *hfq*::3×FLAG from strategy #1 scored only to 1.7%, while the same sample from #2 scored to 9.3% (Table S1C). Collectively, we concluded that despite the lower amount of RNA recovered, strategy #2 proved more successful overall probably due to less contamination with host RNA and was therefore the only analyzed dataset.

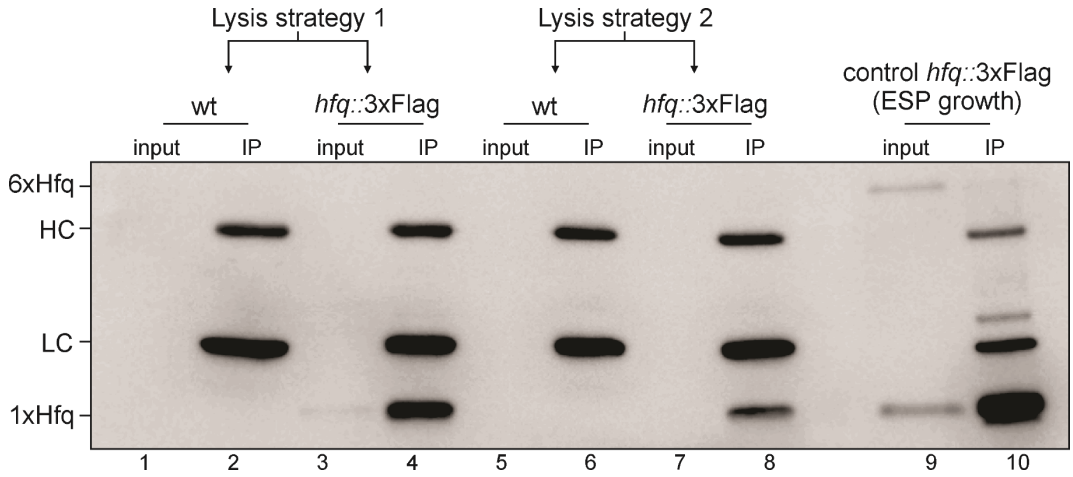

Figure 2.28 Hfq Co-IPs from pilot *in vivo* RIL-seq experiments.
Western blot probed with an anti-FLAG antibody to detect IP of *hfq*::3×FLAG vs. WT samples. Isoforms of Hfq and high-chain (HC) and low-chain (LC) of the antibody are indicated on the left. Input=aliquot of protein isolated right after lysis; IP=eluted samples after Co-IP.

2.3.6 Overview of the *in vivo* RIL-seq

Similarly to previous experiments, the *in vivo* RIL-seq enriched for chimeras in the *hfq*::3×FLAG strain compared to the WT (~1M vs ~2,000, respectively, (Figure 2.29A)). After applying Fisher's exact test and a threshold of 40, only ~190,000 chimeric fragments were retained in the Hfq sample and 125 in the WT sample. Collectively, 1,156 significant RNA-RNA interactions could be detected (Table S2E). sRNAs, CDSs and UTRs were the most represented RNA types in the chimeras (Figure 2.29A), with sRNAs and 3'UTRs mostly represented on the RNA2 of the hybrids, while CDSs and 5'UTRs mostly on RNA1, as previously described (Figure 2.29B-C).

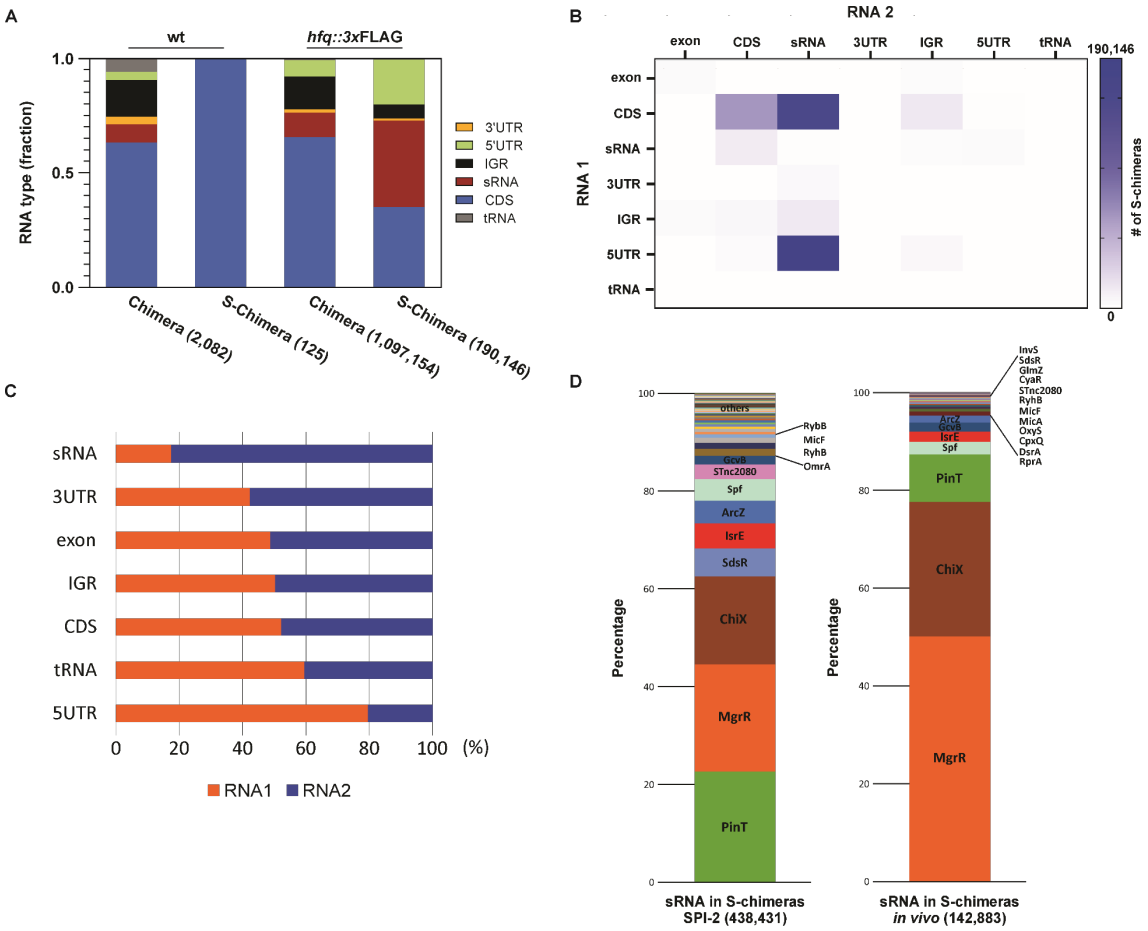

Figure 2.29 Overview of RNA types and sRNAs detected after *in vivo* RIL-seq.
(**A**, **B**, **C**) Plot representation as explained in Figure 2.23 and Figure 6.1. (**D**) Comparison of sRNA distribution between the SPI-2 (related to Figure 2.24) and the *in vivo* RIL-seq datasets.

A direct comparison of sRNAs detected in chimeras in the SPI-2 and *in vivo* datasets highlights a substantial overlap between the two experiments. However, a total of 62 sRNAs were detected in SPI-2 while only 38 *in vivo* (Figure 2.29D). Neverthless, the top 7 sRNAs making up for almost 80% of the total sRNAs in chimeras were represented by the same genes in both cases: MgrR, PinT, ChiX, Spot42, IsrE, GcvB and ArcZ (Figure 2.29D).

2.3.6.1 PinT and InvS sRNAs during infection

PinT scored as top sRNA in the *in vivo* dataset. However, a second sRNA (forming only a limited number of hybrids in both SPI-2 and *in vivo*) was InvS (Figure 2.29D). InvS, an understudied regulatory RNA, has been reported to promote *Salmonella* invasion of epithelial cells through positive regulation of a gene encoding the T3SS-1 apparatus, *prgH* (Wang et al., 2017). In our datasets, InvS consistently formed hybrids with the mRNA encoding the outer membrane protein, *mipA* (~50% of the InvS chimeras in SPI-2 and ~80% of the chimeras *in vivo*), and with the virulence regulator sRNA, PinT (~15% in SPI-2 and ~8% *in vivo*) (Figure 2.30). MipA, a.k.a. OmpV, has been recently shown to mediate adhesion of *Salmonella* to epithelial cell and promote therefore invasion of host cells ((Kaur and Mukhopadhaya, 2020)), while concomitantly activating both the adaptive and innate host immune response, making it a good candidate target for a *Salmonella* vaccine ((Kaur et al., 2021)). Given the novelty of these interactions and the clinical relevance of MipA, we decided to follow up with experimental validations.

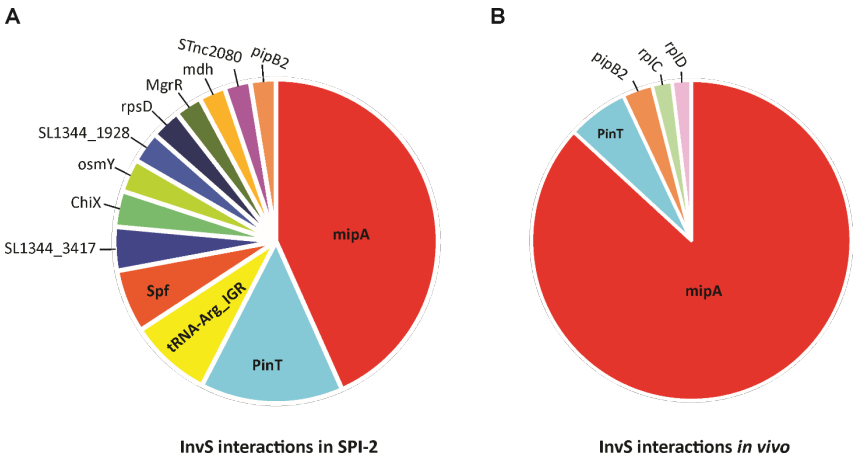

Figure 2.30 InvS interactome detected in SPI-2 and *in vivo* RIL-seq.
(**A**) Pie chart showing the distribution and abundance of the InvS chimeras as detected from the SPI-2 dataset. (**B**) Pie chart showing the distribution and abundance of the InvS chimeras as detected from the *in vivo* dataset

2.3.6.2 InvS regulates *mipA* translation in a GFP-fusion reporter system

To validate the top target of InvS as suggested from our RIL-seq experiments, we constructed a low-copy plasmid carrying the sfGFP CDS under the regulation of the *mipA* 5'UTR from nt -73 to +30 (carrying a predicted binding site according to IntaRNA), relative to *mipA* start codon (Figure 2.31A). Co-expression of a high-copy plasmid carrying the InvS sRNA with the *mipA* reporter caused a complete silencing of GFP fluorescence, pointing at a typical sRNA-mediated translation inhibition (Figure 2.31B, lower plates). No changes in GFP expression were detected with a control plasmid (ctrl) or when InvS was co-expressed with a sfGFP empty vector (Figure 2.31B, upper plates).

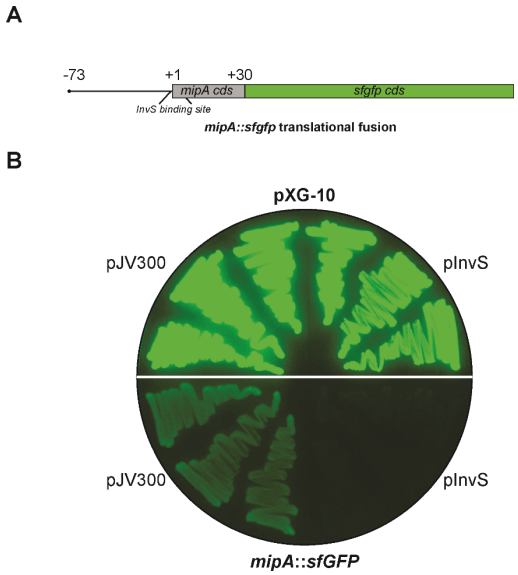

Figure 2.31 Translation reporter assay of *mipA::sfGFP*.
(**A**) Schematic representation of the translational fusion encompassing the 5'UTR of the *mipA* transcript from position -73 to +30, relative to the GUG (start codon). (**B**) LB Agar plates streaked with a Δ*invS* strain transformed with an empty pXG-10 or a *mipA::sfGFP* translational fusion (top and bottom, respectively), co-transformed with a ctrl plasmid (pJV300) or a plasmid overexpressing InvS (pInvS) (left and right, respectively).

2.3.6.3 RNA footprints validate InvS-*mipA* and InvS-PinT complexes

We investigated whether the hybrids recovered via RIL-seq and the InvS-mediated translation regulation of *mipA* formed through direct base pairing between the respective RNA counterparts. Firstly, we predicted potential binding sites between InvS and the *mipA* transcript (Figure 2.32A, left) and between InvS and PinT (Figure 2.32A, right). In both cases, a possible stable base-pair could be assigned. Particularly, InvS was predicted to bind and obstruct the RBS of *mipA* suggesting the observed translation inhibition could occur by impeding ribosome binding in the initial stages of translation (Figure 2.32A, left and green mark). Interestingly, the predicted binding between InvS and PinT would cause sequestration of both sRNAs seed sequences, suggesting a sponging mechanism from either of the two sRNAs.

We further validated these interactions by structure probing experiments. Labeled *mipA* carrying the InvS binding site was incubated with increasing concentration of cold InvS with or without the addition of Hfq and treated with the Pb^{2+} cleaving agent. A clear protection could be

observed in the region corresponding to the start codon of *mipA* (GUG) and 6 nts upstream (Figure 2.32B), confirming the IntaRNA prediction. Addition of Hfq to the reaction made the protection clearer suggesting the RNA chaperone role of this protein is additive but not necessary for the binding (Figure 2.32B, lanes 7-9). Conversely, labeled PinT was not lead (II)-protected by increasing concentration of InvS in the absence of Hfq (Figure 2.32C). However, when the RNA chaperone was added to the reaction a clear footprint appeared in the predicted seed-region at the 5'end of the PinT sRNA (Figure 2.32C, lanes 7-9, lower protection). In addition, Hfq-specific footprints at the 3'end of PinT could also be detected, confirming this sRNA binds this RBP at its terminator-like structure (Figure 2.32C, lanes 7-9, upper protection). Direct binding between the investigated RNAs could also be confirmed by band shift assays (Figure 6.10A-B). Even in this setup, Hfq was necessary for stable binding between InvS and PinT, making these two sRNAs one of the few examples where the base-pair exclusively occur with the aid of an RNA chaperone (Figure 6.10B).

Collectively, RIL-seq enabled the identification of new interactions between two sRNAs involved in infection and set the basis for further characterization of this new regulatory pathway.

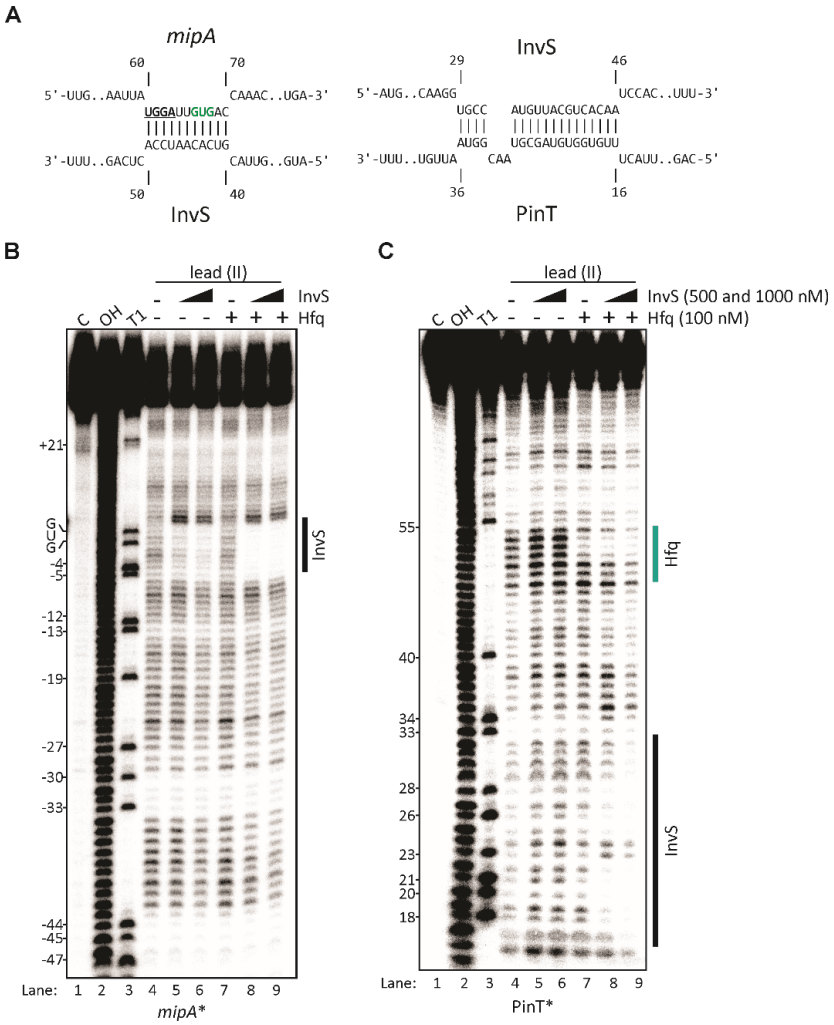

Figure 2.32 RNA structure probing validates InvS interaction with *mipA* and PinT.
(**A**) IntaRNA predictions between InvS and *mipA* (left) or InvS and PinT (right). In bold and underlined is the SD of the *mipA* mRNA, in green the GUG start codon is indicated. (**B**) RNA footprint with Pb^{2+}(lead II) as cleaving agent. Labeled *mipA* RNA or PinT (**B**) were incubated with no (-) or 500 nm and 1000 nM of cold InvS. The Hfq was absent (-) in lanes 4, 5 and 6 or added to the reaction in lanes 7, 8 and 9. Indicated on the right of each gel are the Pb^{2+}-protected regions. C: ctrl; OH: alkaline ladder; T1: RNase T1 G-ladder.

2.3.7 Discussion

In this sub-chapter, the knowledge previously gained from performing RIL-seq in standard conditions was used to extend the methodology to infection-relevant growth. The SPI-2 medium utilized was selected because of its known ability to induce expression of SPI-2 effectors and therefore reproduce the intracellular lifestyle of *Salmonella* during infection (Beuzón et al., 1999; Nikolaus et al., 2001).

SPI-2 RIL-seq was successfully performed in two biological replicates after adapting the protocol to minimal-media growing cultures (Figure 2.22 and Figure 2.23). While on a general level no substantial differences could be highlighted between this dataset and the ESP one (compare Figure 2.23 with Figure 2.3 and Figure 6.1), a deeper analysis of the detected chimeras resulted in a plethora of interesting observations. Firstly, the pool of sRNAs bound to Hfq almost completely differed from ESP. Many sRNAs known to be specifically transcribed upon activation of the SPI-2 program dominated the chimera's landscape of the dataset. Such sRNAs like PinT, IsrE and MgrR accounted for ~50% of total chimeras involving a sRNA (Figure 2.24B). Additional sRNAs without a known SPI-2 specific role could be detected too, like ChiX, SdsR and STnc2080, upholding evidence that RIL-seq enables the discovery of new functional condition-specific sRNAs (Melamed et al., 2016).

One of the advantages of performing RIL-seq in different conditions was the ability to compare the interactions of specific sRNAs consistently appearing in the different networks, as it was the case for both PinT and MgrR (Figure 2.24A). Excluding the PinT top chimera, *ytfK*, the targetome of this sRNA as revealed by RIL-seq differed drastically, suggesting that this sRNA carries an unprecedented plasticity in target recognition (Figure 2.24B). Conversely, MgrR sRNA retained many of its targets in both conditions, except for a few genes forming chimeras with this sRNA only in SPI-2 media. An *in silico* analysis of the MgrR targets and cross-comparison with studies performed in the *E. coli* homolog sRNA (Moon and Gottesman,

2009; Yin et al., 2019), allowed a better understanding of the roles of such an important sRNA (Figure 2.25A-B and Figure 6.8). Specifically, our and previous evidence suggested that MgrR carries out functions directed by the PhoPQ regulation program: i) maintainance of Mg^{2+} homeostasis; ii) regulation of virulence functions through balancing resistance to AMPs and secretion of outer membrane vescicles (Figure 2.25C).

Infection conditions that mimic *in vivo* situations have limitations due to the lack of any possible host-specific influence on the bacteria. Being able to decipher the RNA-RNA interactome of a bacterium inside its host has been the ultimate goal of the present thesis and a real challenge for the entire community. Many complexities were faced while establishing RIL-seq *in vivo*:

- The selection of the host cell type, RAW macrophage, was of primary relevance and strategic as these cell lines are easy to handle and to harvest following an infection experiment. Additionally, these pahgocytic cells do not need activation of the *Salmonella* invasion program prior to infection as they spontaneously uptake opsonized bacteria, unlike epithelial-derived cell lines (Hensel et al., 1998). Importantly, the availability of transcriptomic data on intracellular *Salmonella* in the same murine macrophages, rendered the *in vivo* RIL-seq an optimal complementary resource (Srikumar et al., 2015).
- RIL-seq is a protein-centric, IP-based approach (Melamed et al., 2018) that requires a too high amount of bacteria compared to what can be normally retrieved upon infection assays. A comprehensive pilot approach was firstly undertaken to identify the minimal bacterial cell requirements to perform a successful Hfq IP (Figure 2.26).
- Selecting the correct method to isolate bacteria from the host has proved challenging. While two different lysis approaches were tested (Figure 2.27), only the least conservative one in terms of recovered bacteria was effective (Figure 2.28 and Figure 6.9). Compromising between total amount of recovered bacteria and contamination from the host RNAs was key to success. This notwithstanding, *in vivo* RIL-

seq detected a lower absolute number of chimeras as well as a lesser amount of sRNAs involved in the netweork compared to ESP or SPI-2 RIL-seq datasets (Figure 2.24 and Table S2B-E). Future improvements to the bacterial recovery steps, such as FACS-based sorting of only bacteria-containing host cells (Westermann et al., 2016), could be integrated in order to ameliorate the detection capacity of *in vivo* RIL-seq. Additionally, affinity purification approaches such as by introduction of a His-tag in the protein of interest and pulldown with a Ni-NTA resin could prove more succesfull with lower-input materials and could be introduced as a replacement to the antibody-based precipitation (Granneman et al., 2009; Waters et al., 2017). Such refinements would potentially enable to perform this method at an m.o.i lower than 50 (not completely representative of an *in vivo* bacterial load) and at earlier time points after infection with intracellular non-/slow-replicating bacteria.

Nonetheless, direct comparison of the SPI-2 and *in vivo*-detected sRNAs in chimeras showed a substantial overlap (Figure 2.29), confirming MgrR, ChiX and PinT as the top 3 regulators in intracellular replicating *Salmonella*. Because the roles of PinT are better understood we further investigated the role of an understudied sRNA, InvS that consistently formed hybrids with PinT (Figure 2.30), apart from its top target *mipA*. Our follow up experiments could validate *mipA* as a new, direct target of InvS (Figure 2.31, Figure 2.32B and Figure 6.10A) suggesting this sRNA impairs translation initiation of the *mipA* transcript by occlusion of the RBS (Figure 2.32A, left). Additionally, *in vitro* biochemical assays suggested InvS specifically binds to the seed-sequence of PinT (Figure 2.32A, C), albeit the binding occurred only in the presence of the RNA chaperone, Hfq (Figure 2.32C and Figure 6.10B). Although further investigation is needed, it is tempting to speculate InvS acts as an RNA sponge on PinT, or vice versa. The evidence that the binding between these two RNAs results in sequestration of both's seed-sequences (Figure 2.32A, right) leaves open the possibility that both scenarios are correct. The physiological role of this interaction is still under investigation,

however, several scenarios could be taking place. For instance, the expression patterns of InvS and PinT is surprisingly similar. They are both induced in ESP, a condition mimicking the invasion step of infection, and inside macrophages (SalCom database), albeit PinT levels are generally higher than InvS. The renowned role of *mipA* in regulating *Salmonella* invasion of the host (Kaur and Mukhopadhaya, 2020; Kaur et al., 2021) and the extensively studied role of PinT as a timer of virulence (Correia Santos et al., 2021; Kim et al., 2019; Palmer et al., 2019; Westermann et al., 2016), could drag InvS into the role of yet another riboregulator that rewires the target pool of PinT when cells transition from the "invasion program" to the "intra-cellular program". In fact, factors that direct PinT sRNA to bind SPI-1 targets rather than SPI-2 targets are still unknown. InvS, while regulating *mipA* expression on its own, could buffer excessive PinT activity towards certain targets whose expression decreases over infection (e.g., the mRNAs of PinT SPI-1 targets *sopE* and *sopE2*), while promoting the binding of PinT to more abundant SPI-2-induced mRNAs. Studies of RNA sponges have taught us that binding between RNAs is competitive, making mRNA targets of the same sRNA competitors to one another. In this scenario, binding strength and RNA abundance are the two driving features for sRNA-mediated regulation. Collectively, RIL-seq in SPI-2 conditions and during infection have provided us with novel RNA regulatory circuits to further investigate and new physiology to decipher.

Chapter 3

3 Conclusions & Outlook

Half a century ago, two scientists stated that the flow of genetic information from DNA to protein is irreversible (Crick, 1970) and that information flows from DNA to RNA to protein (Watson, J.D., 1965). Although the latter statement is nowdays incorrect because of the discovery of enzymes able to revert RNA into DNA (Temin and Mizutani, 1970), the function of RNA molecules in the cell has been historically solely associated with that of supporting gene expression through mRNAs, tRNAs and rRNAs.

The discovery of non-coding RNAs in eukaryotes (Cech and Steitz, 2014) and in prokaryotes (Delihas, 2015), did justice to the functional malleability these molecules have. RNAs take part to important complexes such as the machinery that excises introns from immature eukaryotic mRNAs (snRNAs), guide RNA modifications on other RNA molecules (snoRNAs) and are involved in gene regulation (miRNAs). To deny another major dogma of molecular biology, such that catalysis in living organisms is solely performed by polypeptides, ribozymes were discovered (Kruger et al., 1982).

The initial belief was that RNA flexibility was a prerogative of eukaryotes. However, since the discovery of the very first regulatory RNA in *E. coli,* MicF, in the early 80s (Mizuno et al., 1984), it was immediately clear that non-coding RNAs engulf the bacterial kingdom, too. Since their discovery, one of the major challenges to understand sRNAs activity in bacteria, as well as miRNAs in eukaryotes, was the identification of their target RNAs (Beisel and Storz, 2010). Todays' advanced RNA-seq technologies and implemented *in silico* predictions have played a central role in transforming how we identify regulatory RNAs and decipher their targets. We by now know that in most living organisms, regulatory RNAs originates from the most diverse genomic locations, as independently

transcribed entities (Argaman et al., 2001; Wassarman et al., 2001) or processed from coding (Chao and Vogel, 2016; Miyakoshi et al., 2019; Wang et al., 2019) or non-coding (Lalaouna et al., 2015) part of the bacterial genome. Their regulation was believed to only occur at the transcriptional level by the activity of specific sigma factors or transcription regulators able to induce their expression in a condition-specific manner (Guo et al., 2014; Papenfort et al., 2006), however it is widely recognized that sRNAs activity can as well be regulated by RNA chaperones (Hör et al., 2018) and by RNA sponges (Adams et al., 2021; Miyakoshi et al., 2015). Interactome studies in living cells (Han et al., 2016; Helwak et al., 2013; Iosub et al., 2020; Melamed et al., 2016) enabled the simultaneous mapping of the whole RNA-RNA network of a cell, revealing its complexity and dynamicity. This complexity, highlighted by methods like RIL-seq, revelead that sRNAs binding to same mRNAs and carrying overlapping roles are the rule rather than the exception. The recognized inability to identify macroscopic phenotypes related to a single sRNA deletion (Westermann et al., 2016) can be explained by great redundancy among these riboregulators. RIL-seq, as well as other related methodologies, has certainly helped identifying sRNAs that take part to the same regulon(s) and therefore addressing previously unanswered biological questions. Additionally, the possibility to perform such global methods in various growth conditions, as well as during infection of the host, raised even more interest with respect to applied antibacterial therapies.

Since the discovery of penicillin by Alexander Felming in the late 20s, antibiotics have been an increadible lives-saving weapon against many pathogenic bacteria for a whole century. However, the amazingly fast adaptation strategies employed by most pathogens led to a gradual decrease in efficacy of the traditional therapies. The widespread overuse of these drugs in medical and agricultural settings together with a thus far-limited interest in producing new generation antimicrobials added fuel to the antibiotic crisis fire (Nathan, 2015, 2020; Neu, 1992). The extensive knowledge gained from studying RNA regulation in bacteria has motivated

the research community to take off the shelf the relatively old idea of using antisense oligonucleotide (ASO) -based molecules to modulate bacterial growth (Good et al., 2001; Jayaraman et al., 1981). Several modified ASOs have been developed that can base pair with RNA *in vivo* and are resistant to RNases degradation (e.g., peptide nucleic acids (PNAs) or Morpholinos (PMOs)). ASOs have been successfully employed to modulate gene expression in several bacterial models and arrest their growth by inhibition of essential genes (Sully and Geller, 2016; Vogel, 2020). While still in its infancy, this research field promises bits of solutions to the antibiotic resistance problem with the possibility of deploying ASOs as bactericidal or bacteriostatic drugs. The knowledge provided by RIL-seq or analogue methodologies in pathogenic bacteria will certainly help with establishing a more precise and targeted approach in developing ASOs as species-specific programmable antibiotics.

One of the major challenges of global RNA-seq-driven approaches is analysis and visualization of the data. The complexity of the network generated by RIL-seq is undeniable, as ~2000 interactions are usually discovered for each experimental condition (Iosub et al., 2020; Melamed et al., 2016, 2020). Development of visualization tools, as the one established in the present thesis, has offered so far a better way to explore RIL-seq networks and gain more accurate knowledge from them. However, with the ever-increasing accumulation of data points (e.g., RIL-seq over an infection time-course), visualization becomes more and more complex and would eventually require the adoption of further advanced technology. Tools such the recently-developed VRNetzer represent only one example of virtual reality (VR)-based technologies employed for an interacting visualization and exploration of large networks (Pirch et al., 2021).

RIL-seq, like all the other described methods to map RNA-RNA interactions, provides a blurry snapshot of a complex and heterogenous population of bacterial cells. Analogously to traditional bulk transcriptomics, the information relative to the single bacterial cell is lost (Imdahl and Saliba, 2020). While it is hard to believe that an

immunoprecipitation-based approach, like RIL-seq, would succed if performed at the single cell level, other similar strategies could be implemented. Hi-GRIL-seq, for instance, would be more suited since the ligation reaction is performed by expressing the T4 RNA ligase enzyme *in vivo* and does not require pulldown of a specific RBP (Zhang et al., 2017). Bacterial single-cell RNA-seq (Imdahl et al., 2020) combined with Hi-GRIL-seq could potentially represent a first step to resolve the RNA interactome network of a single bacteria. Such technology would not only highlight differential post-transcriptional behaviour of sub-population of bacteria in lab growing conditions, but it would eventually enable the discovery of valuable RNA interactions occurring in infected tissues (Nuss et al., 2017) or whole organisms.

Chapter 4

4 Materials & Methods

4.1 General equipment & consumables

Table 4.1 General equipment and machines.

Instruments/Equipment	Manufacturer	Model
Cell lysis machine	Retsch	MM 400
Centrifuge	Eppendorf	5415R and 5424
Electrophoresis apparatus	Peqlab	EV232, EV202, 250V/300V
Electroporation device	BioRad	Micro Pulser
Geiger counter	Thermo Fisher Scientific	Mini 900 Ratemeter
Gel doc system	Intas	GelStick Imager
Gel drying apparatus	BioRad	583
Imaging machine	GE Healthcare	ImageQuant LAS 4000
Incubator with shaking	New Brunswick Scientific	Innova 44
Magnetic Rack	Thermo Fisher Scientific	DynaMag 2
Phophoimager	GE Healthcare	Typhoon FLA 7000
Phosphor screen	Fujifilm	BAS-SR 2040
Plate reader	BioTek	
RT-qPCR machine	BioRad	CFX96
Sequencing gel apparatus	C.B.S. Scientific Co	SG-400-20
Thermal PCR cycler	BioRad	
UV cross-linker	Vilber	

Table 4.2 Kits and consumables.

Kit/Consumable	Manufacturer
Centrifuge tubes	Sarstedt
Cuvettes	Sarstedt
Electroporation cuvettes	Cell projects
G-25/50 MicroSpin tubes	GE Healthcare
Glass beads, 0.1 mm	Roth

Kit/Consumable	Manufacturer
PCR tubes	Thermo Fisher Scientific
Pipette tips	Sarstedt
MEGAscript T7 Transcription kit	Thermo Fisher Scientific
Phase Lock Gel tubes	5 Prime
Gel and PCR clean-up kit	Macherey-Nagel
Plasmid EasyPure	Machery-Nagel
RNA Clean & Concentrator 5	Zymo Research
NEBNext sRNA Library Prep Set	Illumina
Pierce Protein A/G magnetic beads	Thermo Fisher Scientific
Power SYBR Green RNA-to-CT 1 Step	Thermo Fisher Scientific

Table 4.3 Enzymes.

Enzymes	Manufacturer
Lysozyme	Roth
Phusion DNA polymerase	NEB
Taq DNA polymerase	NEB
T4 DNA ligase	NEB
RNase inhibitor	Promega
SUPERase IN	Thermo Fisher Scientific
Calf Intestinal Phosphatase (CIP)	NEB
T4 Polynucleotide kinase (PNK)	NEB
T4 RNA ligase I	NEB
RNase A/T1 mix	Thermo Fisher Scientific
Protease Inhibitor Cocktail Set III, EDTA-free	Merck
DNase I	Thermo Fisher Scientific
Restriction enzymes	NEB/Thermo Fisher Scientific
SuperScript II RT	Thermo Fisher Scientific

Table 4.4 Reagents and chemicals.

Reagent/Chemical	Manufacturer
Ampicillin	Roth
Carbenicillin	Roth

Reagent/Chemical	Manufacturer
Chloramphenicol	Roth
dNTPs	Thermo Fisher Scientific
ECL reagent for WB	GE Healthcare
DNA ladder	Thermo Fisher Scientific
DNA loading buffer (10x)	Thermo Fisher Scientific
GlycoBlue	Thermo Fisher Scientific
Kanamycin	Roth
Protein ladder	Thermo Fisher Scientific
Rifampicin	Fluka
Roti Hybri-Quick	Roth
Rotiphorese gel 40 (37,5:1)	Roth
Rotiphorese gel 40 (40:1)	Roth
SYBR Gold Nucleic Acid staining	Thermo Fisher Scientific
TRIzol	Thermo Fisher Scientific

Table 4.5 Antibodies.

Antibody/antisera	Source
α-FLAG (mouse)	Sigma-Aldrich
α-GroEL (rabbit)	Sigma-Aldrich
α-mouse; HRP-conjugated (goat)	Cell Signaling Technology
α-rabbit; HRP-conjugated (goat)	Cell Signaling Technology
α-OppA	Kazuei Igarashi, Chiba University, Japan
α-Omp(s)A/C/D/F	Rajeev Misra, Arizona State University, USA
α-GFP	Roche Applied Science

4.2 Media and buffers

Table 4.6 Media.

Media	Composition
Lennox broth (LB)	5 g yeast extract; 5 g NaCl; 10 g tryptone; H_2O to 1 l
LB agar	5 g yeast extract; 5 g NaCl; 10 g tryptone; 15 g agar; H_2O to 1 l
5xMM9	64 g $Na_2HPO_4x7H_2O$, 86.5 g $Na_2HPO_4x12H_2O$, H_2O to 1 l

Media	Composition
M9 + CAS	100 ml 5xMM9 stock, $MgSO_4$ 2 mM, 50 µl $CaCl_2$ 0.1 mM, 5 ml Histinidine 0.004%, Thiamine 0.5 µg/ml, Casaminoacids 0.2%
SPI-2 media	18.12 g MES; 1.25 ml 2 M KCl; 3.75 ml 1 M $(NH_4)_2SO_4$; 500 µL 0.5 M K_2SO_4; 500 µL 1 M KH_2PO_4; 4 µL 1 M $MgCl_2$; 2.78 ml Glycerol 50%; 5 ml 10% Casaminoacids; H_2O to 500 ml; pH 5.8
RPMI medium	Gibco RPMI 11640 - 21875034

Table 4.7 Buffers and solutions.

Buffers/solutions	Composition
10x SDS running buffer	30.275 g Tris; 144 g Glycin; 10 g SDS; H_2O to 1 l
10x PBS	2 g KCl; 17.7 g $Na_2HPO_4x2H_2O$; 2.72 g KH_2PO_4; H_2O to 1 l
10x TBE	216 g Tris; 110 g Boric Acid; 80 ml 0.5 M EDTA pH 8; H_2O to 2 l
10x TBS	48.22 g Tris; 175.32 g NaCl; H_2O to 2 l
1x TBS-T	100 ml 10x TBS, 10 ml 10% Tween; H_2O to 1 l
1xPBS-T	1xPBS supplemented with 0.1% Triton X-100
6% PAA 7M Urea	420 g Urea, 100 ml 10x TBE, 150 ml Rotiphorese Gel 40 (19:1); H_2O to 1 l
RNA Elution buffer	100 mM Na-Acetate pH 6.5; 0.1% SDS; 10 mM EDTA pH 8
5x Native loading dye for EMSA	0.5x TBE, 50% Glycerine; 0.2% Xylene Cyanol; 0.2% bromophenol blue
Stop mix	95% ethanol, 5% acidic phenol
Transfer buffer	3 g Tris base, 14.4 g glycine, 200 ml methoanol, H_2O to 1 l
Lower buffer (PAGE)	1.5 M Tris-HCl, pH 8.8, 0.4% (w/v) SDS
Upper buffer (PAGE)	0.5 M Tris-HCl, pH 6.8, 0.4% (w/v) SDS
SSC	3 M NaCl, 0.3 M sodium citrate, pH7
SSC-S (for northern washes)	SSC + 0.1% (w/v) SDS
Protein loading buffer (5x)	15 g SDS, 46.95 ml 1 M Tris-HCl, pH 6.8, 75 ml glycerol, 11.56 g DTT, 0.075 g bromophenol blue, H_2O to 150 ml
Sensitizer	0.2 g $Na_2S_2O_3$ x 5 H_2O; H_2O to 1 l

Buffers/solutions	Composition
Silver stain solution	2 g $AgNO_3$, 0.75 ml formaldehyde (37%), H_2O to 1 l
Stop solution (silver stain)	10 g glycine, H_2O to 1 l
Fixing solution	500 ml ethanol, 120 ml acetic acid, 0.5 ml formaldehyde (37%), H_2O to 1 l
Developing solution	60 g Na_2CO_3, 4 mg $Na_2S_2O_3$ x 5 H_2O, 0.5 ml formaldehyde (37%), H_2O to 1 l
30:1 mix	29 ml ethanol, 1 ml 3 M sodium acetate, pH 5.2/6.5

4.3 List of bacterial strains, plasmids and oligonucleotides

Table 4.8 Bacterial strains.

Name	Species	Genotype	Source
JVS-1574	*S.* Typhimurium	wild-type	Laboratory stock
JVS-1338	*S.* Typhimurium	*hfq::3xFLAG*	Laboratory stock
JVS-97	*S.* Typhimurium	Δ*micF*	Laboratory stock
JVS-8955	*S.* Typhimurium	Δ*oppA*	Laboratory stock
JVS-2244	*S.* Typhimurium	Δ*ompF*	Laboratory stock
JVS-584	*S.* Typhimurium	Δ*hfq*	Laboratory stock
JVS-12457	*S.* Typhimurium	Δ$oppX_{BS}$	This study
JVS-6999	*S.* Typhimurium	(*rluC-rne*) IG::cat	(Figueroa-Bossi et al., 2000)
JVS-7000	*S.* Typhimurium	(*rluC-rne*) IG::cat / rne-3071 (ts)	(Figueroa-Bossi et al., 2000)
JVS-12875	*S.* Typhimurium	Δ*ompC*Δ*ompD*Δ*ompA::kan*	This study
JVS-12896	*S.* Typhimurium	Δ*ompC*Δ*ompD*Δ*ompA*Δ*ompF::kan*	This study
JVS-12898	*S.* Typhimurium	Δ*ompC*Δ*ompD*Δ*ompA*Δ*micF::kan*	This study
JVS-12944	*S.* Typhimurium	Δ*ompC*Δ*ompD*Δ*oppX*Δ*ompA::kan*	This study
JVS-12881	*S.* Typhimurium	Δ*micF*+pRR5	This study
JVS-12882	*S.* Typhimurium	Δ*micF*+pMS2-MicF	This study
JVS-12883	*S.* Typhimurium	Δ$oppX_{BS}$+pRR5	This study
JVS-12884	*S.* Typhimurium	Δ$oppX_{BS}$ pMS2-OppX	This study
JVS-12792	*S.* Typhimurium	Δ*invS*	This study
YM416	*S.* Typhimurium	Δ*ompA::kan*	This study

Name	Species	Genotype	Source
YM417	*S.* Typhimurium	Δ*micF*Δ*ompA::kan*	This study
YM418	*S.* Typhimurium	Δ*ompF*Δ*ompA::kan*	This study
YM419	*S.* Typhimurium	Δ*oppX*Δ*ompA::kan*	This study
YM420	*S.* Typhimurium	Δ*ompC::kan*	This study
YM421	*S.* Typhimurium	Δ*micF*Δ*ompC::kan*	This study
YM422	*S.* Typhimurium	Δ*ompF*Δ*ompC::kan*	This study
YM423	*S.* Typhimurium	Δ*oppX*Δ*ompC::kan*	This study
JVS-12908	*E. coli* MG1655	wild-type + pJV300	This study
JVS-12909	*E. coli* MG1655	wild-type + pOppX$_{EC}$	This study

Table 4.9 Plasmids.

Name	Resistance	Backbone	Function	Source
pJV300	Amp	pZE12	ctrl plasmid	Sittka et al., 2007
pCP20	Amp/Cm		Flp recombinase	Cherepanov and Wackernagel 1995
pKD4	Amp		Km template for k.o.	Datsenko and Wanner 2000
pSUB11	Amp		3xFLAG-tag template	Uzzau et al., 2001
pKD46	Amp		λRED recombinase	Datsenko and Wanner 2000
pJL-19-1	Cm	pXG-10	*oppA* translational fusion	Sharma et al., 2007
pGM-1	Amp	pZE-12	OppX overexpression	This study
pDP31	Amp	pZE-12	MicF overexpression	Corcoran et al., 2012
pDP23	Cm	pXG-10	*ompF* translational fusion	Corcoran et al., 2012
pXG-1	Cm		ctrl for fusion	Sharma et al., 2007
pGM-14	Amp	pZE-12	OppX$_{FL}$ overexpression	This study
pOppX$_{EC}$	Amp	pZE-12	OppX$_{EC}$ overexpression	This study
pInvS	Amp	pZE-12	InvS overexpression	This study
pMipA	Cm	pXG-10	*mipA* translational fusion	This study

Table 4.10 Oligonucleotides.

Name	Sequence (5′→3′)	Function
JVO-322	CTACGGCGTTTCACTTCTGAGTTC	Salmonella 5S rRNA northern blot oligo probe
JVO-1037	GTTTTTTTTAATACGACTCACTATAGGATCGACGAAAGGCGAT	Forward oligo with T7 promoter for oppA 5'UTR (short) in vitro transcription
JVO-1038	GATGAGCGCAGTGAGTATT	Reverse oligo for OppX in vitro transcription
JVO-1113	TAATAAATAAAGTTAATGATGATAGC	MicF 5'end northern blot oligo probe
JVO-1254	CCGACAAGCAAACGTTGGTAC	Fw qRT-PCR primer for oppA *cds*
JVO-1255	TCACGGCTGACGTTCGATT	Rev qRT-PCR primer for oppA *cds*
JVO-1328	CGCAAACGCAGCAGAAATT	Fw qRT-PCR primer for ompF *cds*
JVO-1329	TTTTACTATCGCCGGTCGTTG	ompF northern blot oligo probe
JVO-1329	TTTTACTATCGCCGGTCGTTG	Rev qRT-PCR primer for ompF *cds*
JVO-1372	GTTTTTTTTAATACGACTCACTATAGGAGACACATAAAGACACCAAACT	Fw primer carrying T7 promoter for ompF in vitro transcription
JVO-1373	CCAGCTTATTACCATCTTTATTA	Rev primer for ompF in vitro transcription
JVO-1727	CTATTGGCCCGTCAAAGAGGAATTTCA	ChiX northern blot oligo probe
JVO-2359	AAGGCGATCGAACGAATCG	Fw qRT-PCR primer for oppA 5′ UTR
JVO-2360	TCAGGGTTCCTGTGCAGCA	Rev qRT-PCR primer for oppA 5′ UTR
JVO-4238	TTTTTTTTAATACGACTCACTATAGGCTATCATCATTAACTTTATTTATTAC	Forward oligo for WT MicF in vitro transcription carrying T7 promoter
JVO-4239	AAAAAAAACCGAATGCGAAGCA	Reverse oligo for MicF in vitro transcription
JVO-8540	CCTCCGACCCCTTCG	tRNAProCGG northern blot oligo probe
JVO-16451	PHO-ATCGACGAAAGGCGATCG	Fw primer to clone 5UTR oppA (from TSS) with phosp oligo
JVO-16452	GTTTTTCTAGATGGTTTTTACCAGCCTG	Rev primer with XbaI site to amplify 5UTR of oppA (binds upstream of RBS)
JVO-16457	GTATACCGAATCGACGAAAGGCGATCGAACGAATCGTCAGGTGTAGGCTGGAGCTGCTTC	Fw primer for deleting MicF binding site on oppA 5UTR – with kan seq on pKD4
JVO-16458	GTGCAGCACTGGTGGTGTACTGCCAGGTCTGTCACTGCTTGGTCCATATGAATATCCTCCTTAG	Rev primer for deleting MicF binding site on oppA 5UTR – with kan seq on pKD4
JVO-16463	GGTCTGTCACTGCTTTGCTATCAC	Salmonella OppX northern blog oligo probe
JVO-16499	ATCGACGAAAGGCGATCG	Fw primer to check del of $oppX_{BS}$
JVO-16500	GGGTTCCTGTGCAGCACT	Rev primer to check del of $oppX_{BS}$
JVO-16732	TTTTTTTTAATACGACTCACTATAGGGTATCATCATTAACTTTATTTATTAC	Forward oligo for MicF C2G mutation for in vitro transcription

Name	Sequence (5′→3′)	Function
JVO-16963	PHO- GGTATCACCGACTTTATTTAT	Rev primer swap G52C in OppX from pGM-1 as template
JVO-18065	GTTTTTTTTAATACGACTCACTATAGGAGCTCCATCG TAGATGAAT	Forward oligo with T7 promoter for oppA 5'UTR (long) in vitro transcription
JVO-18366	PHO-AGCTCCATCGTAGATGAATGTGCTAAA	Fw primer to clone $OppX_{FL}$ into pZE-12
JVO-19504	PHO-ATGAAATAATAAGAGAGCGCCAG	Fw primer to clone InvS into pZE-12
JVO-19505	GTTTTTTCTAGACCGGATGAACTGTTGATAA	Rev primer to clone InvS into pZE-12
JVO-19577	ATTGTGACGTAACATGGCAC	InvS northern oligo probe
JVO-19586	CGCACATTCCAACGGCTTGCCTGCCCCGGATGAACT GTTGGTGTAGGCTGGAGCTGCTTC	Fw primer to delete *invS*
JVO-19587	CGAAATAAAGCTCTGATTGCGCCATAACGAGGGCA CTGGCGGTCCATATGAATATCCTCCTTAG	Rev primer to delete invS
JVO-19622	GTTTTTTAATACGACTCACTATAGGGATGAAATAAT AAGAGAGCGCCAG	Fw primer for InvS in vitro transcription
JVO-19623	CCGGATGAACTGTTGATAA	Rev primer for InvS in vitro transcription
JVO-19624	GTTTTTTAATACGACTCACTATAGGGTAAAATAGTTG GGCTTACAGG	Fw primer for *mipA* in vitro transcription
JVO-19625	GCCAAGAGCCAGAAGTTTAA	Rev primer for *mipA* in vitro transcription
JVO-19626	GTTTTTTATGCATTAAAATAGTTGGGCTTACAGG	Fw primer for mipA-GFP fusion (to 10th aa) in pXG-10 with BfrBI site
JVO-19627	GTTTTTTGCTAGCGCCAAGAGCCAGAAGTTTAA	Rev primer for mipA-GFP fusion (to 10th aa) in pXG-10 with BfrBI site
JVO-19631	AAAATTACTGGTTCAGGCAG	Fw primer to check InvS deletion
JVO-19632	TTGCCGAAGTTTATAACGAT	Rev primer to check InvS deletion
JVO-19895	PHO-TCGTACACCATCAGGGTAC	MS2 fw primer with phosphate at 5′end
JVO-19896	GTTTTTTCTAGATAGCATAAATCAGCCGGGTG	Rev primer to MicF with XbaI site
JVO-19938	PHO-ATCGACATAAGGTGATCGTCT	Fw primer to clone OppX from E. coli in pZE12 with 5'Phosphate
JVO-19938	GTTTTTTCTAGATATTAACCAGCATGTGTAATCC	Rev primer to clone OppX from E. coli in pZE12 with 5'Phosphate
JVO-19989	CTATCACCGACTTATTTATT	NB oligo probe for E. Coli OppX
JVO-19999	CATAGGTCATGTCGCCATTG	*E. coli* ompF northern blot oligo probe

4.4 Common methodologies

4.4.1 *Salmonella enterica* Serovar. Typhimurium

4.4.1.1 Growth procedure

Salmonella was streaked on LB agar plates from DMSO stock aliquots and placed at 37°C O/N. Single colonies were inoculated in 2 ml of liquid media (LB) and grown overnight at 37°C shaking at 220 rpm. The main cultures were started from a 1:100 dilution of the overnight culture and grown at 37°C at 220 rpm. When needed, antibiotics were added to the overnight or to the main culture at the following concentrations:

- Chloramphenicol (Cm): 20 µg/ml
- Carbenicillin/Ampicillin (Carb/Amp): 100 µg/ml
- Kanamycin (Km): 50 µg/ml

When different growth conditions/media were used, the details were specified in the specific section.

4.4.1.2 *Salmonella* competent cells and electroporation

Salmonella competent cells were prepared by growing 10 ml (enough for one transformation) to an OD_{600} of 0.6 in LB media. Cells were evenly cooled at 4°C and washed three times with 10 ml of ice-cold water. For storage, cells were washed and resuspended with ice-cold 10% glycerol. After the last wash, pelleted cells were resuspended in 100 µl of water and used directly for transformation or stored at -80°C. For transformation, 50 µl of competent cells were incubated with 1-2 µl of plasmid DNA or 1 µg of PCR product in a pre-chilled 0.1 cm cuvette. The mix of cells and DNA was submitted to an electrical pulse (1.8 kV, 200 Ω, 25 µF), followed by immediate resuspension in 1 ml of LB and recovered at 37°C at 220 rpm. The cells were pelleted and plated on LB agar plates with the appropriate antibiotic.

4.4.1.3 Deletion or 3xFLAG-tagging of genes

Gene inactivation was performed following the published protocol for *E. coli* (Datsenko and Wanner, 2000). A *Salmonella* strain carrying the pKD46 plasmid which expressed a temperature-sensitive λRED recombinase was grown overnight at 28°C. The O/N culture was diluted 1:300 in 50 ml LB supplemented with 0.2% L-arabinose and grown until an OD_{600} of 0.5 at 28°C. Cells were prepared for electroporation as described in 4.4.1.2 and transformed with 700-1000 ng of a PCR product containing the kanamycin resistance cassette. The PCR product was generated using the pKD4 plasmid as a template and oligos carrying flanking regions to the gene of interest. Transformed cells were plated on LB agar and incubated overnight at 37°C. Deletion mutants were verified by PCR with oligonucleotides external to the deleted region. For introducing a 3xFLAG tag to a gene of interest, the same procedure was undertaken except that the pSUB11 plasmid was used as a template for PCR instead of the pKD4 (Uzzau et al., 2001).

4.4.1.4 P22 transduction

The *Salmonella* strain to transduce was inoculated in 10 ml LB media containitng antibiotic (kanamycin) and 20 µl of WT P22 phage lysate. The culture was shaken overnight at 37°C. The following day, cells were centrifuged at room temperature, 13,200 rpm and the superanatant was isolated into a glass tube. The phage-containting supernatatnt was treated mixed with 600 µl chloroform, vortexed for 10 s and used for transduction. 5-40 µl of P22 lysate was used to transduce 100 µl of the recipient strain grown to an OD_{600} 0.6. Transduction was incubated for 20 min and stopped with 10 µl 100 mM EGTA. Transductions were recovered in 800 µl LB media for 1 hour at 37°C shaking and plated on antibiotic-containting plates.

4.4.1.5 Removal of kanamycin resistance

Positive clones of a deletion/flagged strain carrying kanamycin cassette were grown overnight in 2 ml liquid LB at 37°C. Cells were made electrocompetent as described above and 100 µl were transformed with 50-100 ng pCP20 plasmid. Cells were recovered at 28°C for 2 hours and plated on chloramphenicol and ampicillin plates. Plates were incubated overnight at 28°C. Single clones from the ampicillin plate were picked and dissolved in 20 µl of sterile 0.9% NaCl. From each resuspended colony, 1 µl was spread onto an LB plate without or with ampicillin/chloramphenicol and kanamycin. Plates were incubated at 42°C overnight. Clones that are successfully cured will only grow on LB plate.

4.4.1.6 Vector cloning with restriction enzymes

Inserts were generated by PCR with primers carrying flanking regions bearing restriction sites. The insert and the plasmid were digested with the appropriate restriction enzyme(s). 75 ng of purified PCR insert and 25 ng of linearized purified plasmid were incubated with 1 U of T4 DNA ligase at room temperature for 1—2 h in a 10 µl reaction. 1 µl of ligation were used to transform electrocompetent *Salmonella* or *E. coli* cells. Plasmids are listed in Table 4.9.

4.4.1.7 Growth curves calculation

Overnight cultures of each bacterial strain were pre-grown in 10 ml of LB media until it reached an OD_{600} of 2.0. Strains were subsequently inoculated into a 96-well plate to an initial OD_{600} of 0.01 in the desired media. Measurements were performed with a BioTek Synergy H1 plate reader. The minimal media growth curves were performed in a standard M9 mineral media (Table 4.6) without Glucos or Glycerol and contained 0.2% Casaminoacids as only carbon source.

4.4.2 PCR and agarose gel electrophoresis

DNA was amplified in a PCR reaction containing the *Taq* or Phusion DNA polymerases enzymes. PCR products were purified onto a NucleoSpin Gel and PCR clean-up kit and analysed on agarose gels. Primers used in this thesis are listed in Table 4.10. For agarose gels, DNA samples were mixed with 10x DNA loading buffer and separated in a 1-2% (w/v) agarose gel in 1x TAE for a duration specific to the desired band size at 90-150 V. Gels were stained with ethidium bromide for 20 min and visualized under a UV light source.

4.4.3 RNA extraction with hot phenol

4 OD_{600} of bacterial cells were mixed with 0.2 volumes of stop mix followed by immediated freezing in liquid nitrogen. Samples were thawed and bacterial cells were centrifuged. 600 µl of 1xTE supplemented with 0.5 mg/ml of lysozyme were added to the pellets. Samples were treated with 60 µl of 10% SDS and incubated at 64°C for 2 min. 750 µl of acidic phenol and 66 µl 3M NaOAc, pH 5.2 were added and samples incubated for 6 min at 64°C and mixed once every 30 s. Samples were cooled on ice and centrifuged for 15 min at 13,200 rpm, 4°C. The acqueous phase containing nucleic acids was transferred to a PLG tube and 750 µl of clorophorm were added. After 15 min centrifugation phases were separated and the acqueous phase was mixed with 1.4 ml of 30:1 precipitation mix. RNA was precipitated for 1h at -20°C. RNA pellets were collected by 30 min centrifugation at 4°C, 13,200 rpm and washed twice with 70% ice-cold ethanol. The RNA pellets were diluted in DEPC-water and stored at -20°C.

4.4.4 P/C/I RNA extraction

1 volume of acidic Phenol/chloroform/isoamyl alcohol (P/C/I) was added to the sample. The acqueous phase was isolated by centrifugation for 15 min at 13,200 rpm at 4°C and transferred to a new tube where 3 volumes of 30:1 mix, pH 6.5 were added to precipitate the RNA. Precipitation was carried for 1 h at -20°C and RNA was isolated by 30 min centrifugation at 13,200 rpm at 4°C. RNA pellets were washed twice with 70% ethanol and diluted in DEPC-water.

4.4.5 DNase I digestion

RNA samples were digested with 1 U of DNase I per mg of RNA for 45 min at 37°C. DN-ase-digested RNAs were purified with P/C/I.

4.4.6 SDS-PAGE

Cell pellets were dissolved in 1x protein loading buffer to reach a concentration of 0.01 OD/µl. Samples were boiled for 5-10 min at 95°C and kept a room temperature until loaded. 0.05 or 0.1 ODs were separated on a 10-12% SDS polyacrylamide gel in 1x SDS running buffer for 1-3 h at 40 mA. Gels were either stained with silver staining or used for western blot.

4.4.7 Western blot

Bacterial cultures grown in the desired conditions were collected by centrifugation for 3 min a 13,000 g at 4°C, and the pelleted cells were dissolved in 1x protein loading dye to a final concentration of 0.01 OD/µl. The samples were heated up for 10 min at 95°C, and 0.1 ODs were separated on a 10-12% SDS-PAGE gel. Proteins were transferred onto a methanol-activated (90 sec) PVDF membrane for 90 min at 0.34 A, using a semi-dry

blotter in transfer buffer (25 mM Tris-HCl pH 8.3, 190 mM glycine, 20% methanol). Membranes were blocked with 5% milk for 1 h at room temperature and rinsed in 1xTBS-Tween buffer (20 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.1% Tween20). After blocking, membranes were incubated with the primary antibodies diluted in 1xTBS-T buffer containing 3% BSA for 2 hours at room temperature or O/N at 4°C. The membranes were washed three times for 15 min with agitation in 1xTBS-T buffer at room temperature. Membranes were then incubated for 1 h at RT with HRP-linked secondary antibodies, diluted in 1xTBS-T containing 3% BSA, and washed three times for 15 min with 1xTBS-T. The membranes were developed using the Amersham ECL Prime reagents and signals were detected on a LAS4000 and Imager 600 (GE Healthcare). Bands were quantified using EMBL ImageJ software.

4.4.8 Silver staining

Protein samples were separated onto a 12.5% SDS-PAGE gel followed by incubation of the gel in fixing solution for 3h. The gel was washed twice with a 50% ethanol solution for 20 min, sensitized for 1 min and washed three times in water for 20 s. Then, the gel was incubated in silver staining solution for 20 min and washed twice with water for 20 s. The gel was deleveoped for 15-35 min in developer solution and stopped with glycin when the band reached the optimal intensity. The gel was rinsed with water and imaged with a scanner.

4.4.9 RNA stability assay

Cells were grown to an OD_{600} of 2.0 in LB and 2 ml were collected as the 0 time point (untreated). The remaining cultures were treated with 500 µg/ml rifampicin to stop transcription. Samples were collected at 2, 4, 8 and 16 min post treatment. RNA samples were purified using TRizol extraction and

DNase treated with DNase I. Samples were analysed using northern blot gels.

4.4.10 Northern blot

Bacterial cultures were snap-frozen in liquid nitrogen after the addition of 0.2 vol/vol of stop solution (95% ethanol and 5% phenol). Total RNA was isolated using the TRIzol reagent (Invitrogen), DNAse I treated and precipitated in cold EtOH at -20°C. RNA samples were quantified using a NanoDrop. 5 to 10 µg of total RNA was denatured at 95°C for 5 min in RNA loading dye (95% v/v formamide, 10 mM EDTA, 0.1% w/v xylene cyanole, 0.1% w/v bromophenol blue) and separated on a 6% polyacrylamide/7 M urea gel in 1xTBE buffer for 2 hours at 300 volts. RNA was transferred onto a Hybond-XL nylon membrane (GE Healthcare) with electro-blotting at 50 volts for 1 h at 4°C. The membrane was crosslinked at 120 mJ/cm2 with UV light and pre-hybridized for 10 min in Rapid-Hyb buffer (Amersham). A [32P]-labeled probe was added onto the membrane and hybridized at 42°C overnight with rotation. The membrane was washed three times for 15-min with 5x SSC/0.1% SDS (first wash), 1x SSC/0.1% SDS (second wash) and 0.5x SSC/0.1% SDS (third wash) buffers at 42°C. Air dried membranes were then exposed onto a phosphor screen and signals were visualized on a Typhoon scanner and quantified with the EMBL ImageJ software.

4.4.11 *In vitro* transcription and RNA labelling

400 ng of the desired DNA fragment was PCR-amplified from *Salmonella* genomic DNA and used as a template in a T7 transcription reaction using the MEGAscript T7 Transcription kit. Transcription was performed by incubating the reaction at 37°C for at least 6 hours (up to 16 hours). The size and integrity of RNA was analysed on a denaturing polyacrylamide gel. RNA bands corresponding to the correct expected size were excised from the gel and eluted in RNA elution buffer (0.1 M sodium acetate, 0.1 % SDS,

10 mM EDTA at 4°C overnight), isolated with Phenol:Clorophorm:Isoamyl (P:C:I) and precipitated in EtOH. An amount of 50 pmol of RNA was dephosphorylated with 10 units of calf intestinal phosphatase (CIP) in a 50 µL reaction at 37°C for 1 h. CIP-treated RNA was extracted with P:C:I and EtOH precipitated. 20 pmol of the dephosphorylated RNA was 5'-labelled with 2 µL of ^{32}P-γ-ATP (10 µCi/µL) using 1 unit of T4 polynucleotide kinase (PNK) for 1 h at 37°C in a 20 µL reaction. RNA was purified from unincorporated nucleotides with microspin G-50 columns (GE Healthcare) according to the manufacturer´s instructions. Purified and labelled RNA was separated on a 6% polyacrylamide gel and the correct sized band was extracted as described earlier. This step ensures that any degraded RNA fragments is excluded from the RNA pool used for subsequent reactions.

4.4.12 Electrophoretic mobility shift assays (EMSA)

0.04 pmol of radio-labelled RNA was used for each reaction mix. Labelled RNA was denatured at 95°C for 1 min and cooled on ice for 5 min. 1xStructure buffer (10 mM Tris-HCl pH 7.0, 0.1 M KCl, 10 mM $MgCl_2$) was added and the RNA was re-natured at 37°C for 10 min. 1 µg of Yeast RNA (Ambion) was added to each reaction and the labelled RNA was added to tubes containing increasing concentration of unlabelled RNA. When Hfq was included in the assay, the protein was pre-incubated with the boiled and re-natured RNA for 15 min at 37°C. Binding reactions were incubated at 37°C for 20 min, stopped by adding 5x RNA native loading buffer and separated on a native 6% polyacrylamide gel at 4°C in 0.5% TBE at constant current of 40 mA for 3-4 h. Gels were dried and signals detected on a Typhoon FLA 7000 phosphoimager and quantified with EMBL ImageJ software. Quantified bands were plotted using Prism to generate a dose-dependent binding curve and estimate a dissociation constant (k_d). For the competition binding assays, pre-formed complexes characterized by one labelled and one cold RNA species were incubated with increasing concentration of cold competitor.

4.4.13 RNA structure probing

Labeled and cold RNAs were prepared as described for the EMSA. The reactions were prepared as follows: 0.4 pmol of labeled RNA were denatured as described earlier and incubated with increasing concentration of unlabeled RNA partner for 15 min at 37°C in the presence of 1xStructure buffer and 1 µg of yeast RNA in 10 µL. RNAs were cleaved with 2 µL of 25 nM Lead-Acetate for 90 sec at 37°C. To stop the cleavage reaction, 12 µL of GL II RNA loading dye were added to each tube. 10 µL of each sample were boiled at 95°C for 3 min, loaded on an 8-10% PAA 7 M Urea gel, and separated for 3 hours at 45 Watts. For the CTRL lane, 1 pmol of labeled RNA was denatured at 95°C in 10µL of water and cooled on ice upon adding 10 µL of GL II RNA loading dye. For the OH ladder, 1 pmol of labeled RNA was denatured at 95°C for 5 min in 1xAlkaline buffer in a 10 µL reaction. For the T1 ladder, 1 pmol of RNA was denatured in H_2O for 1 min at 95°C followed by addition of 1 µL of RNase T1 enzyme and incubated for 3 min at 37°C. All reactions were stopped as mentioned before.

4.4.14 Translational gfp reporter assay

Plasmid carrying a gfp-reporter were constructed by restriction enzyme cloning strategy using the pXG-10 plasmid as a backbone. The 5'UTR of the investigated mRNA was cloned upstream of and in frame to the gfp coding sequence. Translation reporter plasmids with or without insert were transformed to the appropriate recipient strains. Strains carrying a gfp-reporter plasmid were grown as described above and streaked on LB-agar followed by UV-exposure to visualize GFP expression. For western blot analysis, strains were grown in LB to an OD_{600} of 2.0 and samples were collected. Samples were separated onto a polyacrilamide SDS-PAGE gel and immunodetected with a GFP-specific antibody.

4.4.15 Reverse transcription qPCR (RT-qPCR) analysis

RNA was extracted as described for northern blot analysis. The absence of gDNA contamination was tested through PCR-based amplification and RT-qPCR was performed using the Power SYBR Green RNA-to-CT1-Step kit (Thermo Fisher Scientific) and the CFX96 system (Bio-Rad). 5S ribosomal RNA was quantified as housekeeping gene. Data were analyzed using the ΔΔCT method (Livak and Schmittgen, 2001). Oligos used for this experiment are listed in Table 4.10.

4.5 Uncommon methodologies

4.5.1 RIL-seq protocol and analysis

RIL-seq experiments were performed following the protocol described in (Melamed et al., 2018) with a few modifications to adapt the steps to *Salmonella*. WT strain or a strain carrying a chromsomally flag-tagged *hfq* gene were grown in LB medium until OD_{600} of 2.0. A total amount of 80 ODs was collected by centrifugation, washed once in ice-cold 1xPBS, and cross-linked under a 256 nm UV light source. Cross-linked cells were pelleted and resuspended in 1 ml ice-cold 1xPBS, transferred into a 1 ml Eppendorf tube and pelleted once again by centrifugation. If not immediately used for the following steps, the pellets were snap frozen in liquid nitrogen. On the day of the experiments, pellets were thawed on ice and resuspended in 800 µl of lysis buffer (NP-T buffer: 50 mM NaH_2PO_4, 300 mM NaCl, 0.05% Tween, pH 8.0) supplemented with protease inhibitor (1:200) and RNase inhibitor (final concentration of 0.1 U/µL). Lysis was performed by adding 800 µl of 0.1 mm glass beads and shaking for 10 min at 30 Hz in the Retsch MM400 machine. Lysates were cleared by a centrifugation for 3 min at 13,000 g followed by another centrifugation step of 20 min at 13,000 g to remove cell debris and remaining beads. Cleared cell lysates were incubated with 3 µg of anti-Flag (M2 monoclonal antibody) bound protein A/G magnetic beads for 2 hours at 4°C with rotation followed by three washing steps with 200 µl of lysis

buffer. Samples were digested with an RNase A/T1 mix for 5 min at 22°C in an RNase inhibitor-free lysis buffer. Samples were washed three times with 200 µl of lysis buffer supplemented with 3.25 µL (per sample) of SUPERase In RNase inhibitor (Thermo-Fisher, 20 U/µL). The trimmed ends of the Hfq-bound RNAs were cured by PNK treatment for 2 hours at 22°C with agitation, followed by 3 washing steps at 4°C with 200 µl of normal lysis buffer. This step allows the 5′ end to be phosphorylated and the 3′ to have a free –OH. Proximal Hfq-bound RNAs were ligated by treatment with T4 RNA ligase I enzyme in the following reaction: 8 µL T4 ligase buffer, 7.2 µL DMSO, 0.8 µL ATP (100 mM), 32 µL PEG 8000, 1.2 µL RNase inhibitor, 23.6 µL of water, 140 units of T4 RNA ligase I enzyme. Samples were incubated O/N at 22°C with agitation, followed by three steps of washing with 200 µl lysis buffer, at 4°C. The RNAs were eluted from beads with a proteinase K digestion for 2 hours at 55°C followed by LS Trizol extraction, as per manufacturer indications. Purified RNAs were resuspended in 7 µL of nuclease-free water and quality controlled on a Bioanalyzer pico RNA chip before proceeding with cDNA library preparation.

Library preparation was conducted using the sRNA NEBNext kit for Illumina, with few modifications. An amount of 3 µL of RNA or H_2O (for the negative control) was mixed with 1 µL of 3′ SR Adaptor (pre-diluted 1:10) and incubated at 70°C for 2 min. 6.5 µL of 3′ ligation mix (5 µL of 3′ ligation buffer, 1.5 µL of 3′ enzyme mix) were added to each reaction and incubated at 25°C for 1 hour. 2.75 µL of SR reaction mix (2.5 µL of water, 0.25 µL of SR RT primer) were added to the tubes which were then incubated for three sub sequential steps of: 75°C for 5 min, 37°C for 15 min, 25°C for 15 minutes. This step allows for binding and inactivation of the non-ligated excess of 3′ adaptor by the SR RT primer, preventing subsequent formation of 3′ and 5′ primer dimers. Pre-denatured 5′adaptor (pre-diluted 1:10) was added to each tube together with 5′ ligation mix (0.5 µL of 5′ ligation reaction buffer, 1.25 µL 5′ ligation enzyme mix) and samples were incubated at 25°C for 1 hour. First strand cDNA synthesis was performed by adding 5 µL of cDNA synthesis mix (4 µL first strand buffer, 0.5 µL Murine RNase inhibitor, 0.5 µL

SuperScript II RT). The reactions were incubated at 50°C for 1 hour followed by 15 min at 70°C. 10 µL of cDNA was PCR amplified with barcoded NEB index primers and SR primer in a 50 µL reaction (25 µL LongAmp Taq 2x mix, 12.5 µL nuclease-free water, 1.25 µL SR primer, 1.25 µL index primer). The PCR cycling program was set as follow: 30 s at 94°C initial denaturation, 15 s at 94°C, 30 s at 62°C and 70°C for 16, 18, 20 or 22 cycles, and a final elongation of 5 min at 70°C. PCR products were AMPure XL beads purified and checked on a DNA Bioanalyzer (or on a 6% polyacrylamide gel) to estimate size distribution and amount of DNA fragments. Once the optimal amplification cycle was identified, cDNAs were re-amplified and purified as described above. Amplified cDNAs were equally pooled and sequenced on an Illumina NextSeq 500 platform. Sequenced fragments were mapped to Salmonella enterica subsp. Serovar Typhimurium str. SL1344 genome including one chromosome (NC_016810.1) and three plasmids (pCol1B9 NC_017718.1, pRSF1010 NC_017719.1, and pSLT NC_017720.1).

Sequenced fragments that mapped to two different loci were defined as chimeric. Otherwise, mapped sequenced fragments were defined as single if they mapped to the same transcript or within 1,000 nts. The ratio of chimeras out of the total number of sequenced fragments that mapped as single and chimera is indicated in Table S1 for all the RIL-seq experiments. Fisher´s exact test was applied to assign each chimera an Odds Ratio value and a *p*-value corrected for multiple hypotheses testing. Chimeras with a *p*-value ≤0.05 were considered as representing a putative interacting RNA pair and further investigated. These chimeras were defined as S-chimeras. A threshold number of chimeric fragments was further considered to filter out lowly represented S-chimeras. This number was heuristically defined for each experiment, considering the average depth of sequencing in the experiments and the number of chimeras identified in the control libraries (WT). The reads mapping to ribosomal genes were filtered out *in silico* prior to any further analysis. While in general the computational pipeline followed Melamed et al. (Melamed et al., 2016, 2018), some modifications were made to match the cDNA library preparation and the *Salmonella*

Typhimurium str. SL1344 spec. For the analyses, an *ad hoc* annotation employed in (Holmqvist et al., 2016) was used here. Briefly, gene annotations from NCBI were used for the genomic features such as tRNAs, rRNAs and CDSs. Transcriptional units (TUs) were defined according to TSS annotation (Kröger et al., 2013) and Rho-independent terminator prediction with RNIE (Gardner et al., 2011). For the plasmids, a BioCyc annotation was used (Karp et al., 2019). In Melamed et al. intergenic regions within transcript (IGT) were distinguished from other intergenic regions (IGR), while in our data they were under one annotation (IGR). In addition, in the current analysis the data of the two or more replicate experiments were not unified, but analyzed separately.

4.5.2 RNase E inactivation assay

A *Salmonella* LT2 strain carrying a thermo-sensitive allele of *rne* and its isogenic WT control were grown overnight at 28°C in LB media. The next day strains were diluted 1:100 in fresh LB media and grown to an OD_{600} 2.0 at 28°C. An aliquot of bacteria was taken an shifted to 44°C for 30 min while the remaining culture was kept for 30 min at 28°C. After this time. 4 OD_{600} of cells were harvested and RNA was extracted as described in 4.4.3. RNA samples were analysed via northern blot.

4.5.3 MS2-pulldown

MS2 pulldowns were performed with *Salmonella* strains carrying a plasmid-driven expression of a MS2-tagged RNA (MicF or OppX). The experiments were performed as previously described (Smirnov et al., 2016). Shortly, 200 OD_{600} of *Salmonella* cultures carrying the MS2-tagged RNAs or vector control plasmids collected at early stationary phase were used for each pulldown. Cells were lysed and ran through a MS2-MBP coated amylose beads. RNAs were eluted with 12 mM maltose and samples extracted with the P:C:I chemical (acqueous phase used for RNA extraction, organic phase for

protein extraction). RNA samples were analysed on a northern blot gel, whereas protein samples were analysed with silver staining (4.4.8).

4.5.4 *Salmonella* infection procedure

T75 cell flasks were seeded with 200,000 murine RAW macrophage cells at passage 16. On the day of the infection, eukaryotic cells were counted and overnight bacterial cultures were harvested to an amount that would determine a multiplicity of infection (m.o.i.) of 50 for all the needed flasks (# bacterial cells = (m.o.i. x # of eukaryotic cells x # of flasks)/bacterial concentration. Before infection, harvested bacteria were opsonised with 10% of mouse serum for 20 min at room temperature. Opsonised bacteria were resuspended in RPMI media to the correct dilution to ensure correct m.o.i. and inoculated in cells-containing flasks. Flasks were centrifuged at 250 g for 10 min at room temperature to synchronise the infection. Flasks were incubated for 30 min at 37°C to allow infection to take place. The RPMI from the flasks was replaced with fresh RPMI containing 100 µg/ml of gentamycin (high-gentamycin) to kill extracellular bacteria and flasks were incubated at 37°C for 30 min. RPMI medium was replaced once more with RPMI containing 10 µg/ml of gentamycin (low-gentamycin) and incubated for the desired time post infection (20 hours in the experiment presented here). At 20 h, falsks were washed once with ice-cold 1xPBS and incubated with 10 ml 1xPBS-T to detach adhering cells. Harvested cells were incubated for 10 min at room temperature to ensure mild lysis of only eukaryotic cells. Cells were centrifuged at 250 g for 10 min at 4°C to separate eukaryotic cell debris from intracellular bacteria. The bacteria-containing supernatant was collected and centrifuged at 4,500 g for 20 min at 4°C. Bacterial pellet was snap-frozen in liquid nitrogen and RIL-seq was performed as described in 4.5.1.

4.5.5 Gradient fractionation

Gradient experiments were performed as described in (Smirnov et al., 2016). Salmonella WT and □oppX were grown at 37°C to an OD600 of 2.0 and lysed in lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM MgCl2, 1 mM DTT, 1 mM PMSF, 0.2% Triton X-100, 20 U/ml DNase I (Thermo Fisher Scientific), 200 U/ml SUPERase-IN (Life Technologies)) on a Retsch MM400 machine with 0.1 mm glass beads. Lysates were cleared and the supernatants centrifuged at 100,000 g for 17 h at 4°C through linear 10-40% (wt/vol) glycerol gradients pre-formed in a Beckman SW40Ti tubes. The gradients were fractionated in 20 equal fractions, plus a pellet fraction. Each fraction was deproteinzed with 1% SDS, RNA was extracted with the P:C:I solvent and ethanol precipitated.

4.5.6 Size-exclusion chromatography

Salmonella strains were grown as described in Gradient fractionation. Cells were resuspended in lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM KCl, 10 mM MgCl2, 1 mM DTT, 1 mM PMSF, 0.2% Triton X-100, 20 U/ml DNase I (Thermo Fisher Scientific), 200 u/ml SUPERase-IN (Life Technologies)) and lysate was injected into the Äkta pure 25 column system. Gel filtration program was selected, and a Superose 6 increase 10/300 GL column (GE Healthcare) was used for size-exclusion experiments. The column flow rate was fixed to 0.25 ml/min. UV measurements (280 nm and 254 nm) were automatically performed from the machine in real-time. The column was equilibrated with 2 CV of 1x Lysis buffer at the same flowrate. The lysate was then applied into the inlet syringe system and fractions were collected over time. Fractions 29 to 48 were collected, each containing 380 µL of sample. 290 µL of each fraction was deproteinized with 25 µL of 10% SDS by hand shaking for 20 sec. For the RNA extraction, an equal volume (300 µL) of P:C:I solvent was added to each fraction followed by the adding of 400 µL of chloroform. Samples were vortexed for 30 s and let rest for 5 min at room

temperature. Samples were centrifuged at 13,000 rpm at 4°C for 15 min and the acqueous phase was ethanol/sodium acetate-precipitated at -20°C for 1 hour. Northern blotting was performed on DNase I-digested RNA samples.

Chapter 5

5 Bibliography

Acuña, L.G., Barros, M.J., Peñaloza, D., Rodas, P.I., Paredes-Sabja, D., Fuentes, J.A., Gil, F., and Calderón, I.L. (2016). A feed-forward loop between SroC and MgrR small RNAs modulates the expression of eptB and the susceptibility to polymyxin B in Salmonella Typhimurium. Microbiology (Reading) *162*, 1996–2004.

Adams, P.P., Baniulyte, G., Esnault, C., Chegireddy, K., Singh, N., Monge, M., Dale, R.K., Storz, G., and Wade, J.T. (2021). Regulatory roles of Escherichia coli 5′ UTR and ORF-internal RNAs detected by 3′ end mapping. Elife *10*, e62438.

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.

Andersen, J., Delihas, N., Ikenaka, K., Green, P.J., Pines, O., Ilercil, O., and Inouye, M. (1987). The isolation and characterization of RNA coded by the *micF* gene in *Escherichia coli*. Nucleic Acids Res. *15*, 2089–2101.

Andrews, J.C., and Short, S.A. (1985). Genetic analysis of *Escherichia coli* oligopeptide transport mutants. J. Bacteriol. *161*, 484–492.

Andrews, J.C., and Short, S.A. (1986). opp-lac Operon fusions and transcriptional regulation of the *Escherichia coli* trp-linked oligopeptide permease. J. Bacteriol. *165*, 434.

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. Current Biology *11*, 941–950.

Ascano, M., Hafner, M., Cekan, P., Gerstberger, S., and Tuschl, T. (2012). Identification of RNA-protein interaction networks using PAR-CLIP. Wiley Interdiscip Rev RNA *3*, 159–177.

Bar, A., Argaman, L., Altuvia, Y., and Margalit, H. (2021). Prediction of Novel Bacterial Small RNAs From RIL-Seq RNA-RNA Interaction Data. Front Microbiol *12*, 635070.

Barchinger, S.E., and Ades, S.E. (2013). Regulated proteolysis: control of the *Escherichia coli* σE-dependent cell envelope stress response. In Regulated Proteolysis in Microorganisms, pp. 129–160.

Bauriedl, S., Gerovac, M., Heidrich, N., Bischler, T., Barquist, L., Vogel, J., and Schoen, C. (2020). The minimal meningococcal ProQ protein has an intrinsic capacity for structure-based global RNA recognition. Nat Commun *11*, 2823.

Beisel, C.L., and Storz, G. (2010). Base pairing small RNAs and their roles in global regulatory networks. FEMS Microbiol Rev *34*, 866–882.

Beuzón, C.R., Banks, G., Deiwick, J., Hensel, M., and Holden, D.W. (1999). pH-dependent secretion of SseB, a product of the SPI-2 type III secretion system of Salmonella typhimurium. Mol Microbiol *33*, 806–816.

Blocker, A., Komoriya, K., and Aizawa, S.-I. (2003). Type III secretion systems and bacterial flagella: Insights into their function from structural similarities. Proceedings of the National Academy of Sciences *100*, 3027–3030.

Bomjan, R., Zhang, M., and Zhou, D. (2019). YshB promotes intracellular replication and is required for *Salmonella* virulence. J. Bacteriol. *201*, e00314-19, /jb/201/17/JB.00314-19.atom.

Brantl, S. (2007). Regulatory mechanisms employed by cis-encoded antisense RNAs. Curr Opin Microbiol *10*, 102–109.

Brawn, L.C., Hayward, R.D., and Koronakis, V. (2007). Salmonella SPI1 effector SipA persists after entry and cooperates with a SPI2 effector to regulate phagosome maturation and intracellular replication. Cell Host Microbe *1*, 63–75.

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.

Cech, T.R., and Steitz, J.A. (2014). The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. Cell *157*, 77–94.

Chao, Y., and Vogel, J. (2016). A 3′ UTR-derived small RNA provides the regulatory noncoding arm of the inner membrane stress response. Mol. Cell *61*, 352–363.

Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., and Vogel, J. (2012). An atlas of Hfq-bound transcripts reveals 3′ UTRs as a genomic reservoir of regulatory small RNAs: Hfq-dependent small RNAs from 3′ UTRs. The EMBO Journal *31*, 4005–4019.

Chao, Y., Li, L., Girodat, D., Förstner, K.U., Said, N., Corcoran, C., Śmiga, M., Papenfort, K., Reinhardt, R., Wieden, H.-J., et al. (2017). In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathways. Molecular Cell *65*, 39–51.

Chen, S., Lesnik, E.A., Hall, T.A., Sampath, R., Griffey, R.H., Ecker, D.J., and Blyn, L.B. (2002). A bioinformatics based approach to discover small RNA genes in the Escherichia coli genome. Biosystems *65*, 157–177.

Choi, J., and Groisman, E.A. (2017). Activation of master virulence regulator PhoP in acidic pH requires the Salmonella-specific protein UgtL. Sci Signal *10*, eaan6284.

Choi, J., and Groisman, E.A. (2020). Horizontally acquired regulatory gene activates ancestral regulatory system to promote Salmonella virulence. Nucleic Acids Res *48*, 10832–10847.

Chou, J.H., Greenberg, J.T., and Demple, B. (1993). Posttranscriptional repression of Escherichia coli OmpF protein in response to redox stress: positive control of the micF antisense RNA by the soxRS locus. J. Bacteriol. *175*, 1026.

Coleman, J., Green, P.J., and Inouye, M. (1984). The use of RNAs complementary to specific mRNAs to regulate the expression of individual bacterial genes. Cell *37*, 429–436.

Colgan, A.M., Kröger, C., Diard, M., Hardt, W.-D., Puente, J.L., Sivasankaran, S.K., Hokamp, K., and Hinton, J.C.D. (2016). The Impact of 18 Ancestral and Horizontally-Acquired Regulatory Proteins upon the Transcriptome and sRNA Landscape of Salmonella enterica serovar Typhimurium. PLoS Genet *12*, e1006258.

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.D., and Vogel, J. (2012). Superfolder GFP reporters validate diverse new mRNA targets of the classic porin regulator, MicF RNA. Mol. Microbiol. *84*, 428–445.

Correia Santos, S., Bischler, T., Westermann, A.J., and Vogel, J. (2021). MAPS integrates regulation of actin-targeting effector SteC into the virulence control network of Salmonella small RNA PinT. Cell Rep *34*, 108722.

Crick, F. (1970). Central Dogma of Molecular Biology. Nature *227*, 561–563.

Dalebroux, Z.D., and Miller, S.I. (2014). Salmonellae PhoPQ regulation of the outer membrane to resist innate immunity. Curr. Opin. Microbiol. *17*, 106–113.

Darnell, R.B. (2010). HITS-CLIP: panoramic views of protein-RNA regulation in living cells. Wiley Interdiscip Rev RNA *1*, 266–286.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A *97*, 6640–6645.

Dehinwal, R., Cooley, D., Rakov, A.V., Alugupalli, A.S., Harmon, J., Cunrath, O., Vallabhajosyula, P., Bumann, D., and Schifferli, D.M. (2021). Increased Production of Outer Membrane Vesicles by Salmonella Interferes with Complement-Mediated Innate Immune Attack. MBio *12*, e0086921.

Deighan, P., Free, A., and Dorman, C.J. (2000). A role for the Escherichia coli H-NS-like protein StpA in OmpF porin expression through modulation of micF RNA stability. Mol. Microbiol. *38*, 126–139.

Delihas, N. (2015). Discovery and characterization of the first non-coding RNA that regulates gene expression, *micF* RNA: A historical perspective. WJBC *6*, 272.

Denham, E.L. (2020). The Sponge RNAs of bacteria - How to find them and their role in regulating the post-transcriptional network. Biochim Biophys Acta Gene Regul Mech *1863*, 194565.

Detweiler, C.S., Monack, D.M., Brodsky, I.E., Mathew, H., and Falkow, S. (2003). virK, somA and rcsC are important for systemic Salmonella enterica serovar Typhimurium infection and cationic peptide resistance. Mol Microbiol *48*, 385–400.

Dimastrogiovanni, D., Fröhlich, K.S., Bandyra, K.J., Bruce, H.A., Hohensee, S., Vogel, J., and Luisi, B.F. (2014). Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. Elife *3*.

Egger, L.A., Park, H., and Inouye, M. (1997). Signal transduction via the histidyl-aspartyl phosphorelay. Genes to Cells *2*, 167–184.

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

Fröhlich, K.S., and Gottesman, S. (2018). Small regulatory RNAs in the enterobacterial response to envelope damage and oxidative stress. Microbiol Spectr *6*, 4.

Fröhlich, K.S., Papenfort, K., Fekete, A., and Vogel, J. (2013). A small RNA activates CFA synthase by isoform-specific mRNA stabilization. EMBO J *32*, 2963–2979.

Fröhlich, K.S., Haneke, K., Papenfort, K., and Vogel, J. (2016). The target spectrum of SdsR small RNA in Salmonella. Nucleic Acids Res *44*, 10406–10422.

Gama-Castro, S., Jimenez-Jacinto, V., Peralta-Gil, M., Santos-Zavaleta, A., Penaloza-Spinola, M.I., Contreras-Moreira, B., Segura-Salazar, J., Muniz-Rascado, L., Martinez-Flores, I., Salgado, H., et al. (2007). RegulonDB (version 6.0): gene regulation model of *Escherichia coli* K-12 beyond transcription, active (experimental) annotated promoters and Textpresso navigation. Nucleic Acids Research *36*, D120–D124.

Giacomodonato, M.N., Uzzau, S., Bacciu, D., Caccuri, R., Sarnacki, S.H., Rubino, S., and Cerquetti, M.C. (2007). SipA, SopA, SopB, SopD and SopE2 effector proteins of Salmonella enterica serovar Typhimurium are synthesized at late stages of infection in mice. Microbiology (Reading) *153*, 1221–1228.

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.

Gonzalez, G.M., Hardwick, S.W., Maslen, S.L., Skehel, J.M., Holmqvist, E., Vogel, J., Bateman, A., Luisi, B.F., and Broadhurst, R.W. (2017). Structure of the Escherichia coli ProQ RNA-binding protein. RNA *23*, 696–711.

Good, L., Awasthi, S.K., Dryselius, R., Larsson, O., and Nielsen, P.E. (2001). Bactericidal antisense effects of peptide–PNA conjugates. Nat Biotechnol *19*, 360–364.

Gorski, S.A., Vogel, J., and Doudna, J.A. (2017). RNA-based recognition and targeting: sowing the seeds of specificity. Nat Rev Mol Cell Biol *18*, 215–228.

Grabowicz, M., and Silhavy, T.J. (2017). Envelope stress responses: an interconnected safety net. Trends Biochem. Sci. *42*, 232–242.

Grabowicz, M., Koren, D., and Silhavy, T.J. (2016). The CpxQ sRNA negatively regulates Skp to prevent mistargeting of β-barrel outer membrane proteins into the cytoplasmic membrane. MBio *7*, e00312-00316.

Granneman, S., Kudla, G., Petfalski, E., and Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. Proceedings of the National Academy of Sciences *106*, 9613–9618.

Guo, M.S., Updegrove, T.B., Gogol, E.B., Shabalina, S.A., Gross, C.A., and Storz, G. (2014). MicL, a new σE-dependent sRNA, combats envelope stress by repressing synthesis of Lpp, the major outer membrane lipoprotein. Genes Dev. *28*, 1620–1634.

Han, K., Tjaden, B., and Lory, S. (2016). GRIL-seq provides a method for identifying direct targets of bacterial small regulatory RNA by *in vivo* proximity ligation. Nat. Microbiol. *2*, 16239.

Haraga, A., Ohlson, M.B., and Miller, S.I. (2008). Salmonellae interplay with host cells. Nat Rev Microbiol *6*, 53–66.

Harris, J.C., Dupont, H.L., and Hornick, R.B. (1972). Fecal leukocytes in diarrheal illness. Ann Intern Med *76*, 697–703.

Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the Human miRNA Interactome by CLASH Reveals Frequent Noncanonical Binding. Cell *153*, 654–665.

Hensel, M., Shea, J.E., Waterman, S.R., Mundy, R., Nikolaus, T., Banks, G., Vazquez-Torres, A., Gleeson, C., Fang, F.C., and Holden, D.W. (1998). Genes encoding putative effector proteins of the type III secretion system of *Salmonella* pathogenicity island 2 are required for bacterial virulence and proliferation in macrophages. Molecular Microbiology *30*, 163–174.

Holmqvist, E., Unoson, C., Reimegård, J., and Wagner, E.G.H. (2012). A mixed double negative feedback loop between the sRNA MicF and the global regulator Lrp. Mol. Microbiol. *84*, 414–427.

Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., Backofen, R., and Vogel, J. (2016). Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking in vivo. EMBO J. *35*, 991–1011.

Holmqvist, E., Li, L., Bischler, T., Barquist, L., and Vogel, J. (2018). Global Maps of ProQ Binding In Vivo Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3′ Ends. Mol Cell *70*, 971-982.e6.

Hope, I.A., and Struhl, K. (1986). Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell *46*, 885–894.

Hör, J., Gorski, S.A., and Vogel, J. (2018). Bacterial RNA Biology on a Genome Scale. Molecular Cell *70*, 785–799.

Hör, J., Matera, G., Vogel, J., Gottesman, S., and Storz, G. (2020). Trans-acting small RNAs and their effects on gene expression in *Escherichia coli* and *Salmonella enterica*. EcoSal Plus *9*.

Hoyos, M., Huber, M., Förstner, K.U., and Papenfort, K. (2020). Gene autoregulation by 3′ UTR-derived bacterial small RNAs. ELife *9*, e58836.

Imdahl, F., and Saliba, A.-E. (2020). Advances and challenges in single-cell RNA-seq of microbial communities. Current Opinion in Microbiology *57*, 102–110.

Imdahl, F., Vafadarnejad, E., Homberger, C., Saliba, A.-E., and Vogel, J. (2020). Single-cell RNA-sequencing reports growth-condition-specific global transcriptomes of individual bacteria. Nat Microbiol *5*, 1202–1206.

Iosub, I.A., van Nues, R.W., McKellar, S.W., Nieken, K.J., Marchioretto, M., Sy, B., Tree, J.J., Viero, G., and Granneman, S. (2020). Hfq CLASH uncovers sRNA-target interaction networks linked to nutrient availability adaptation. Elife *9*, e54655.

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

Jayaraman, K., McParland, K., Miller, P., and Ts'o, P.O.P. (1981). Selective inhibition of Escherichia coli protein synthesis and growth by nonionic oligonucleotides complementary to the 3′ end of 16S rRNA. Proceedings of the National Academy of Sciences *78*, 1537–1541.

Johansen, J., Eriksen, M., Kallipolitis, B., and Valentin-Hansen, P. (2008). Down-regulation of outer membrane proteins by noncoding RNAs: unraveling the cAMP-CRP- and sigmaE-dependent CyaR-*ompX* regulatory case. J. Mol. Biol. *383*, 1–9.

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.

Karp, P.D., Billington, R., Caspi, R., Fulcher, C.A., Latendresse, M., Kothari, A., Keseler, I.M., Krummenacker, M., Midford, P.E., Ong, Q., et al. (2019). The BioCyc collection of microbial genomes and metabolic pathways. Brief. Bioinformatics *20*, 1085–1093.

Kaur, D., and Mukhopadhaya, A. (2020). Outer membrane protein OmpV mediates Salmonella enterica serovar typhimurium adhesion to intestinal epithelial cells via fibronectin and α1β1 integrin. Cell Microbiol *22*, e13172.

Kaur, D., Gandhi, S., and Mukhopadhaya, A. (2021). Salmonella Typhimurium Adhesin OmpV Activates Host Immunity To Confer Protection against Systemic and Gastrointestinal Infection in Mice. Infect Immun *89*.

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. (2005). Detection of 5′- and 3′-UTR-derived small RNAs and cis-encoded antisense RNAs in Escherichia coli. Nucleic Acids Research *33*, 1040–1050.

Keene, J.D., Komisarow, J.M., and Friedersdorf, M.B. (2006). RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. Nat Protoc *1*, 302–307.

Kim, K., Palmer, A.D., Vanderpool, C.K., and Slauch, J.M. (2019). The Small RNA PinT Contributes to PhoP-Mediated Regulation of the Salmonella Pathogenicity Island 1 Type III Secretion System in Salmonella enterica Serovar Typhimurium. J Bacteriol *201*, e00312-19.

Klein, G., and Raina, S. (2017). Small regulatory bacterial RNAs regulating the envelope stress response. Biochem Soc Trans *45*, 417–425.

König, J., Zarnack, K., Luscombe, N.M., and Ule, J. (2012). Protein-RNA interactions: new genomic technologies and perspectives. Nat Rev Genet *13*, 77–83.

Konovalova, A., Mitchell, A.M., and Silhavy, T.J. (2016). A lipoprotein/β-barrel complex monitors lipopolysaccharide integrity transducing information across the outer membrane. ELife *5*, e15276.

Kroger, C., Dillon, S.C., Cameron, A.D.S., Papenfort, K., Sivasankaran, S.K., Hokamp, K., Chao, Y., Sittka, A., Hebrard, M., Handler, K., et al. (2012). The transcriptional landscape and small RNAs of Salmonella enterica serovar Typhimurium. Proceedings of the National Academy of Sciences *109*, E1277–E1286.

Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S.K., Hammarlöf, D.L., Canals, R., Grissom, J.E., Conway, T., Hokamp, K., et al. (2013). An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. Cell Host Microbe *14*, 683–695.

Kruger, K., Grabowski, P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T.R. (1982). Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. Cell *31*, 147–157.

Kudla, G., Granneman, S., Hahn, D., Beggs, J.D., and Tollervey, D. (2011). Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. Proceedings of the National Academy of Sciences *108*, 10010–10015.

Lalaouna, D., Carrier, M.-C., Semsey, S., Brouard, J.-S., Wang, J., Wade, J.T., and Massé, E. (2015). A 3′ external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. Mol. Cell *58*, 393–405.

LaRock, D.L., Chaudhary, A., and Miller, S.I. (2015). Salmonellae interactions with host processes. Nat Rev Microbiol *13*, 191–205.

Lee, C.A., and Falkow, S. (1990). The ability of Salmonella to enter mammalian cells is affected by bacterial growth state. Proceedings of the National Academy of Sciences *87*, 4304–4308.

Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources. Cell *157*, 624–635.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-ΔΔC(T)) method. Methods *25*, 402–408.

Lu, J., Li, L., Pan, F., Zuo, G., Yu, D., Liu, R., Fan, H., and Ma, Z. (2020). PagC is involved in salmonella pullorum OMVs production and affects biofilm production. Vet Microbiol *247*, 108778.

Massé, E., Vanderpool, C.K., and Gottesman, S. (2005). Effect of RyhB small RNA on global iron use in Escherichia coli. J Bacteriol *187*, 6962–6971.

Matamouros, S., and Miller, S.I. (2015). S. Typhimurium strategies to resist killing by cationic antimicrobial peptides. Biochimica et Biophysica Acta (BBA) - Biomembranes *1848*, 3021–3025.

McClelland, M., Sanderson, K.E., Spieth, J., Clifton, S.W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., et al. (2001). Complete genome sequence of Salmonella enterica serovar Typhimurium LT2. Nature *413*, 852–856.

Melamed, S., Peer, A., Faigenbaum-Romm, R., Gatt, Y.E., Reiss, N., Bar, A., Altuvia, Y., Argaman, L., and Margalit, H. (2016). Global mapping of small RNA-target interactions in bacteria. Mol. Cell *63*, 884–897.

Melamed, S., Faigenbaum-Romm, R., Peer, A., Reiss, N., Shechter, O., Bar, A., Altuvia, Y., Argaman, L., and Margalit, H. (2018). Mapping the small RNA interactome in bacteria using RIL-seq. Nat. Protoc. *13*, 1–33.

Melamed, S., Adams, P.P., Zhang, A., Zhang, H., and Storz, G. (2020). RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles. Mol. Cell *77*, 411-425.e7.

Miyakoshi, M., Chao, Y., and Vogel, J. (2015). Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. EMBO J. *34*, 1478–1492.

Miyakoshi, M., Matera, G., Maki, K., Sone, Y., and Vogel, J. (2019). Functional expansion of a TCA cycle operon mRNA by a 3′ end-derived small RNA. Nucleic Acids Research *47*, 2075–2088.

Mizuno, T., Chou, M.Y., and Inouye, M. (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proceedings of the National Academy of Sciences *81*, 1966–1970.

Møller, T., Franch, T., Højrup, P., Keene, D.R., Bächinger, H.P., Brennan, R.G., and Valentin-Hansen, P. (2002). Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. Mol Cell *9*, 23–30.

Moon, K., and Gottesman, S. (2009). A PhoQ/P-regulated small RNA regulates sensitivity of Escherichia coli to antimicrobial peptides: sRNA regulates sensitivity of E. coli to antimicrobial peptides. Molecular Microbiology *74*, 1314–1330.

Moon, K., and Gottesman, S. (2011). Competition among Hfq-binding small RNAs in Escherichia coli. Mol Microbiol *82*, 1545–1562.

Nathan, C. (2015). Cooperative development of antimicrobials: looking back to look ahead. Nat Rev Microbiol *13*, 651–657.

Nathan, C. (2020). Resisting antimicrobial resistance. Nat Rev Microbiol *18*, 259–260.

Neu, H.C. (1992). The crisis in antibiotic resistance. Science *257*, 1064–1073.

Nikolaus, T., Deiwick, J., Rappl, C., Freeman, J.A., Schröder, W., Miller, S.I., and Hensel, M. (2001). SseBCD proteins are secreted by the type III secretion system of Salmonella pathogenicity island 2 and function as a translocon. J Bacteriol *183*, 6036–6045.

Nuss, A.M., Beckstette, M., Pimenova, M., Schmühl, C., Opitz, W., Pisano, F., Heroven, A.K., and Dersch, P. (2017). Tissue dual RNA-seq allows fast discovery of infection-specific functions and riboregulators shaping host–pathogen transcriptomes. Proc Natl Acad Sci USA *114*, E791–E800.

Pain, A., Ott, A., Amine, H., Rochat, T., Bouloc, P., and Gautheret, D. (2015). An assessment of bacterial small RNA target prediction programs. RNA Biol *12*, 509–513.

Palmer, A.D., Kim, K., and Slauch, J.M. (2019). PhoP-Mediated Repression of the SPI1 Type 3 Secretion System in Salmonella enterica Serovar Typhimurium. J Bacteriol *201*, e00264-19.

Panja, S., Schu, D.J., and Woodson, S.A. (2013). Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. Nucleic Acids Res *41*, 7536–7546.

Panja, S., Santiago-Frangos, A., Schu, D.J., Gottesman, S., and Woodson, S.A. (2015). Acidic Residues in the Hfq Chaperone Increase the Selectivity of sRNA Binding and Annealing. J Mol Biol *427*, 3491–3500.

Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J.C.D., 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., Pfeiffer, V., Lucchini, S., Sonawane, A., Hinton, J.C.D., and Vogel, J. (2008). Systematic deletion of *Salmonella* small RNA genes identifies CyaR, a conserved CRP-dependent riboregulator of OmpX synthesis. Mol. Microbiol. *68*, 890–906.

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

Papenfort, K., Sun, Y., Miyakoshi, M., Vanderpool, C.K., and Vogel, J. (2013). Small RNA-mediated activation of sugar phosphatase mRNA regulates glucose homeostasis. Cell *153*, 426–437.

Petrone, B.L., Stringer, A.M., and Wade, J.T. (2014). Identification of HilD-Regulated Genes in Salmonella enterica Serovar Typhimurium. Journal of Bacteriology *196*, 1094–1101.

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.

Pfeiffer, V., Papenfort, K., Lucchini, S., Hinton, J.C.D., 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.

Pirch, S., Müller, F., Iofinova, E., Pazmandi, J., Hütter, C.V.R., Chiettini, M., Sin, C., Boztug, K., Podkosova, I., Kaufmann, H., et al. (2021). The VRNetzer platform enables interactive network analysis in Virtual Reality. Nat Commun *12*, 2432.

Raden, M., Ali, S.M., Alkhnbashi, O.S., Busch, A., Costa, F., Davis, J.A., Eggenhofer, F., Gelhausen, R., Georg, J., Heyne, S., et al. (2018). Freiburg RNA tools: a central online resource for RNA-focused research and teaching. Nucleic Acids Res. *46*, W25–W29.

Raivio, T.L. (2014). Everything old is new again: an update on current research on the Cpx envelope stress response. Biochim. Biophys. Acta *1843*, 1529–1541.

Ramani, N., Hedeshian, M., and Freundlich, M. (1994). MicF antisense RNA has a major role in osmoregulation of OmpF in Escherichia coli. J. Bacteriol. *176*, 5005–5010.

Rice, J.B., and Vanderpool, C.K. (2011). The small RNA SgrS controls sugar-phosphate accumulation by regulating multiple PTS genes. Nucleic Acids Res. *39*, 3806–3819.

Said, N., Rieder, R., Hurwitz, R., Deckert, J., Urlaub, H., and Vogel, J. (2009). In vivo expression and purification of aptamer-tagged small RNA regulators. Nucleic Acids Research *37*, e133–e133.

Schu, D.J., Zhang, A., Gottesman, S., and Storz, G. (2015). Alternative Hfq-sRNA interaction modes dictate alternative mRNA recognition. EMBO J *34*, 2557–2573.

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., Hoffmann, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., Chabas, S., Reiche, K., Hackermüller, J., Reinhardt, R., et al. (2010). The primary transcriptome of the major human pathogen Helicobacter pylori. Nature *464*, 250–255.

Sharma, C.M., Papenfort, K., Pernitzsch, S.R., Mollenkopf, H.-J., Hinton, J.C.D., 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.

Silhavy, T.J., Kahne, D., and Walker, S. (2010). The Bacterial Cell Envelope. Cold Spring Harb. Perspect. Biol. *2*, a000414–a000414.

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., Hinton, J.C.D., 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.

Smirnov, A., Förstner, K.U., Holmqvist, E., Otto, A., Günster, R., Becher, D., Reinhardt, R., and Vogel, J. (2016). Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. Proc. Natl. Acad. Sci. U.S.A. *113*, 11591–11596.

Smirnov, A., Wang, C., Drewry, L.L., and Vogel, J. (2017). Molecular mechanism of mRNA repression in trans by a ProQ-dependent small RNA. EMBO J *36*, 1029–1045.

Srikumar, S., Kröger, C., Hébrard, M., Colgan, A., Owen, S.V., Sivasankaran, S.K., Cameron, A.D.S., Hokamp, K., and Hinton, J.C.D. (2015). RNA-seq Brings New Insights to the Intra-Macrophage Transcriptome of Salmonella Typhimurium. PLoS Pathog *11*, e1005262.

Stecher, B., Robbiani, R., Walker, A.W., Westendorf, A.M., Barthel, M., Kremer, M., Chaffron, S., Macpherson, A.J., Buer, J., Parkhill, J., et al. (2007). Salmonella enterica serovar typhimurium exploits inflammation to compete with the intestinal microbiota. PLoS Biol *5*, 2177–2189.

Steele-Mortimer, O., Brumell, J.H., Knodler, L.A., Méresse, S., Lopez, A., and Finlay, B.B. (2002). The invasion-associated type III secretion system of Salmonella enterica serovar Typhimurium is necessary for intracellular proliferation and vacuole biogenesis in epithelial cells. Cell Microbiol *4*, 43–54.

Stein, E.M., Kwiatkowska, J., Basczok, M.M., Gravel, C.M., Berry, K.E., and Olejniczak, M. (2020). Determinants of RNA recognition by the FinO domain of the Escherichia coli ProQ protein. Nucleic Acids Res *48*, 7502–7519.

Storz, G., Vogel, J., and Wassarman, K.M. (2011). Regulation by small RNAs in bacteria: expanding frontiers. Mol Cell *43*, 880–891.

Sully, E.K., and Geller, B.L. (2016). Antisense antimicrobial therapeutics. Current Opinion in Microbiology *33*, 47–55.

Takayanagi, K., Maeda, S., and Mizuno, T. (1991). Expression of micF involved in porin synthesis in Escherichia coli: two distinct cis-acting elements respectively regulate micF expression positively and negatively. FEMS Microbiology Letters *83*, 39–44.

Temin, H.M., and Mizutani, S. (1970). Viral RNA-dependent DNA Polymerase: RNA-dependent DNA Polymerase in Virions of Rous Sarcoma Virus. Nature *226*, 1211–1213.

Thomason, M.K., Bischler, T., Eisenbart, S.K., Förstner, K.U., Zhang, A., Herbig, A., Nieselt, K., Sharma, C.M., and Storz, G. (2015). Global Transcriptional Start Site Mapping Using Differential RNA Sequencing Reveals Novel Antisense RNAs in Escherichia coli. J Bacteriol *197*, 18–28.

Tree, J.J., Granneman, S., McAteer, S.P., Tollervey, D., and Gally, D.L. (2014). Identification of bacteriophage-encoded anti-sRNAs in pathogenic *Escherichia coli*. Mol. Cell *55*, 199–213.

Udekwu, K.I., Darfeuille, F., Vogel, J., Reimegård, J., Holmqvist, E., and Wagner, E.G.H. (2005). Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. Genes Dev. *19*, 2355–2366.

Updegrove, T.B., Kouse, A.B., Bandyra, K.J., and Storz, G. (2019). Stem-loops direct precise processing of 3′ UTR-derived small RNA MicL. Nucleic Acids Res *47*, 1482–1492.

Urban, J.H., and Vogel, J. (2007). Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Research *35*, 1018–1037.

Uzzau, S., Figueroa-Bossi, N., Rubino, S., and Bossi, L. (2001). Epitope tagging of chromosomal genes in Salmonella. Proc Natl Acad Sci U S A *98*, 15264–15269.

Vergalli, J., Bodrenko, I.V., Masi, M., Moynié, L., Acosta-Gutiérrez, S., Naismith, J.H., Davin-Regli, A., Ceccarelli, M., van den Berg, B., Winterhalter, M., et al. (2020). Porins and small-molecule translocation across the outer membrane of Gram-negative bacteria. Nat. Rev. Microbiol. *18*, 164–176.

Vogel, J. (2003). RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. Nucleic Acids Res. *31*, 6435–6443.

Vogel, J. (2020). An RNA biology perspective on species-specific programmable RNA antibiotics. Mol Microbiol *113*, 550–559.

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

Wang, C., Chao, Y., Matera, G., Gao, Q., and Vogel, J. (2019). The conserved 3′ UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. Nucleic Acids Research gkz1168.

Wang, L., Cai, X., Wu, S., Bomjan, R., Nakayasu, E.S., Händler, K., Hinton, J.C.D., and Zhou, D. (2017). InvS Coordinates Expression of PrgH and FimZ and Is Required for Invasion of Epithelial Cells by Salmonella enterica serovar Typhimurium. J Bacteriol *199*.

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.

Waters, S.A., McAteer, S.P., Kudla, G., Pang, I., Deshpande, N.P., Amos, T.G., Leong, K.W., Wilkins, M.R., Strugnell, R., Gally, D.L., et al. (2017). Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. EMBO J. *36*, 374–387.

Watson, J.D. (1965). Molecular Biology of the Gene. (W.A. Benjamin, Inc. New York).

Westermann, A.J., Förstner, K.U., Amman, F., Barquist, L., Chao, Y., Schulte, L.N., Müller, L., Reinhardt, R., Stadler, P.F., and Vogel, J. (2016). Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions. Nature *529*, 496–501.

Winter, S.E., Thiennimitr, P., Winter, M.G., Butler, B.P., Huseby, D.L., Crawford, R.W., Russell, J.M., Bevins, C.L., Adams, L.G., Tsolis, R.M., et al. (2010). Gut inflammation provides a respiratory electron acceptor for Salmonella. Nature *467*, 426–429.

Yin, X., Wu Orr, M., Wang, H., Hobbs, E.C., Shabalina, S.A., and Storz, G. (2019). The small protein MgtS and small RNA MgrR modulate the PitA phosphate symporter to boost intracellular magnesium levels. Mol. Microbiol. *111*, 131–144.

Zhang, A., Schu, D.J., Tjaden, B.C., Storz, G., and Gottesman, S. (2013). Mutations in Interaction Surfaces Differentially Impact E. coli Hfq Association with Small RNAs and Their mRNA Targets. Journal of Molecular Biology *425*, 3678–3697.

Zhang, Y.-F., Han, K., Chandler, C.E., Tjaden, B., Ernst, R.K., and Lory, S. (2017). Probing the sRNA regulatory landscape of *P. aeruginosa*: post-transcriptional control of determinants of pathogenicity and antibiotic susceptibility. Mol. Microbiol. *106*, 919–937.

Chapter 6

6 Appendix

6.1 Appendix Figures

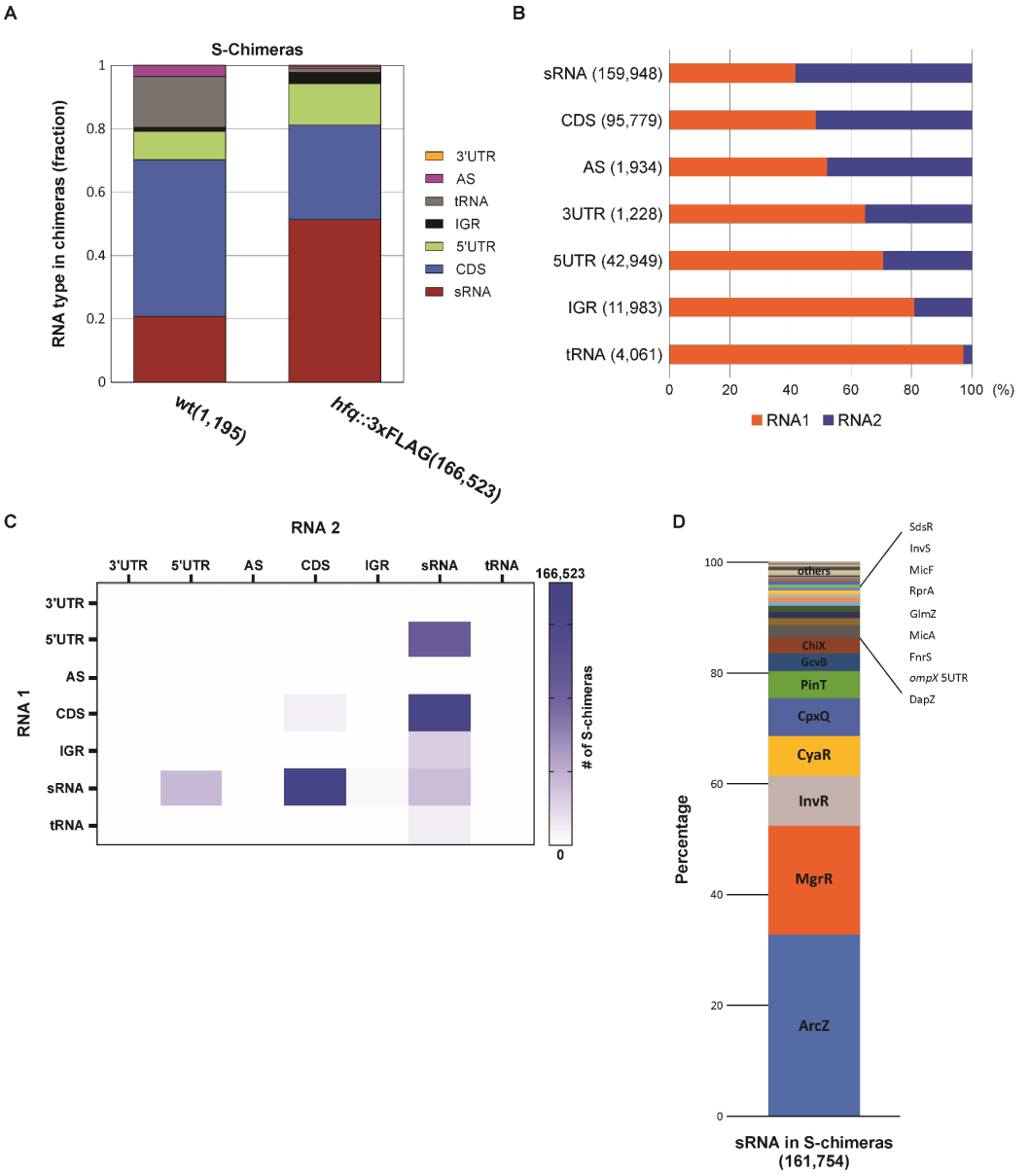

Figure 6.1 Analysis of RIL-seq RNA distribution.
(**A**) Percentage of sequenced chimeric fragments corresponding to S-chimeras plotted by RNA type for W1 and H1, respectively WT and *hfq*::3×FLAG. The total number of sequenced chimeric fragments corresponding to S-chimeras is indicated in brackets. (**B**) Two-sided plot showing distribution of chimeric fragments for each RNA type through RNA1 and RNA2. The plot is representative of the H1 replicate. (**C**) Heat map plotting number of chimeric fragments corresponding to S-chimeras according to the position in the read (RNA1 or RNA2). Due to limited RNA type annotation of the plasmids, fragments mapping to plasmids were excluded from the count in A, B and C. (**D**) Percentage of chimeric fragments corresponding to S-chimeras for individual sRNAs (relative to S-chimeras including at least one sRNA, based on H1 library). Interactions involving both chromosomal and plasmid-derived genes were included in this analysis.

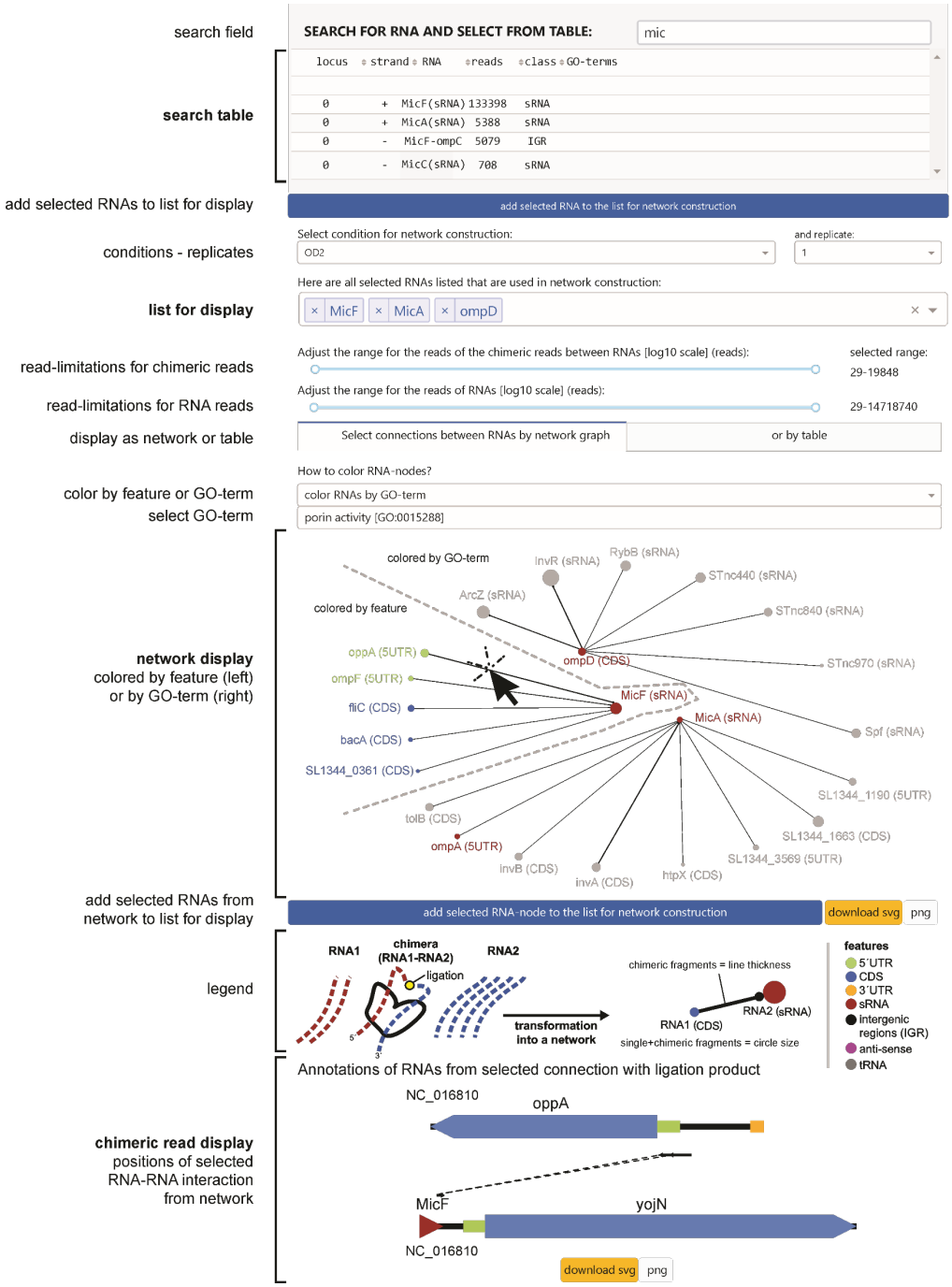

Figure 6.2 RIL-seq browser screenshot.
For demonstrative purposes, the MicF, MicA and *ompD* RNAs were selected and shown in the RIL-seq browser (http://resources.helmholtz-hiri.de/rilseqset/). The interaction investigated in this study (MicF/OppX (*oppA* 5'UTR)) is indicated with a black arrow.

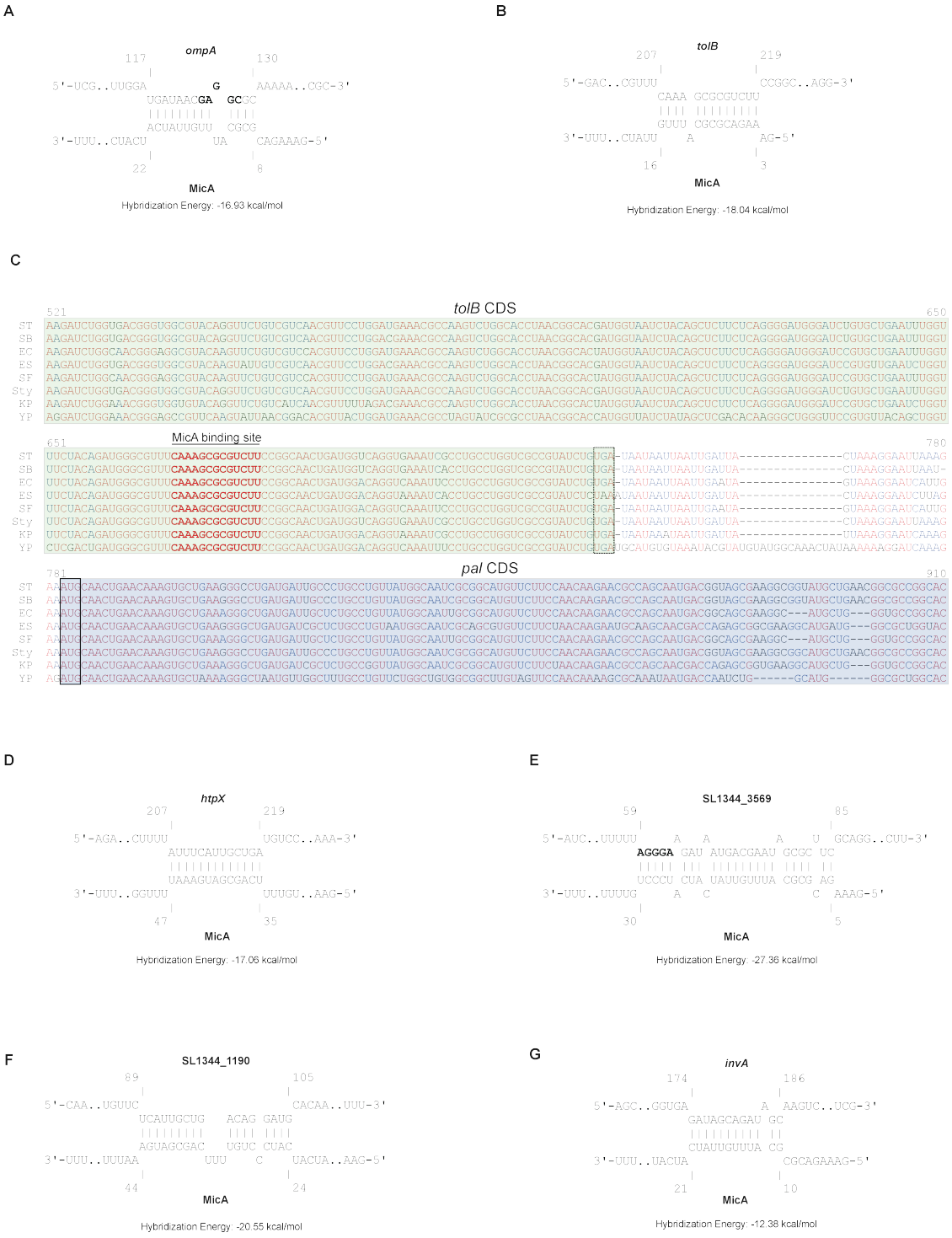

Figure 6.3 Putative MicA-targets identified *via* RIL-seq.
(**A, B**) Prediction of RNA-RNA hybrids formation through IntaRNA between MicA-*ompA* and MicA-*tolB*, respectively. (**C**) Sequence alignment across Enterobacteriacae of the *tolB-pal* bi-cistronic mRNA. Highlighted in bold is the MicA binding site as predicted by IntaRNA. Boxed areas represent the stop and the start codons of *tolB* and *pal*, respectively. (**D, E, F, G**) IntaRNA predictions of hybrids between MicA and the targets identified with RIL-seq, *tolB, ompA, htpX, SL1344_3569, SL1344_1190* and *invA*, respectively. For (**A**) and (**E**) the SD sequence is highlighted in bold.

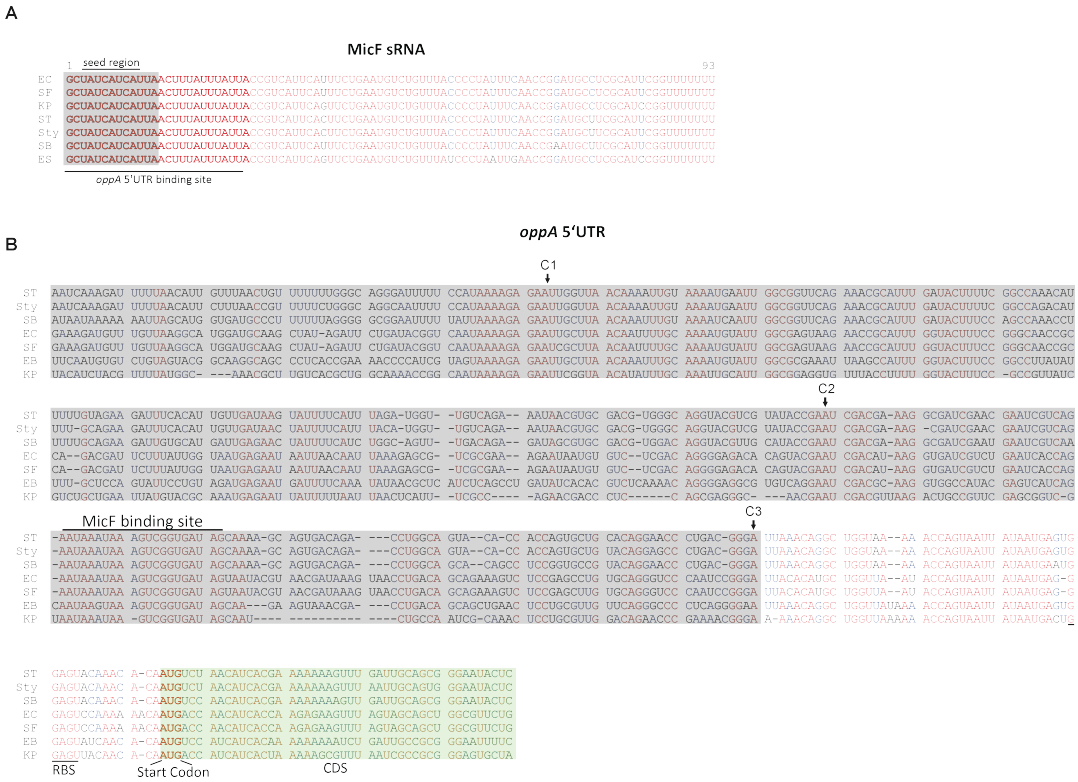

Figure 6.4 Conservation and alignment of MicF and the *oppA* locus.
(**A**) Sequence alignment of the MicF sRNA. In bold, the 25 nts (containing the seed sequence of the sRNA, indicated above the sequence) involved in the interaction with the *oppA* 5′UTR. (**B**) Sequence alignment of the *oppA* 5′UTR. Arrows indicate putative RNase E cleavage sites (C1, C2 and C3). The conserved region involved in binding to MicF is indicated above the sequence. The boxed green area indicates the initial part of the coding sequence of *oppA* mRNA. Boxed in grey is the area released upon processing at C3.

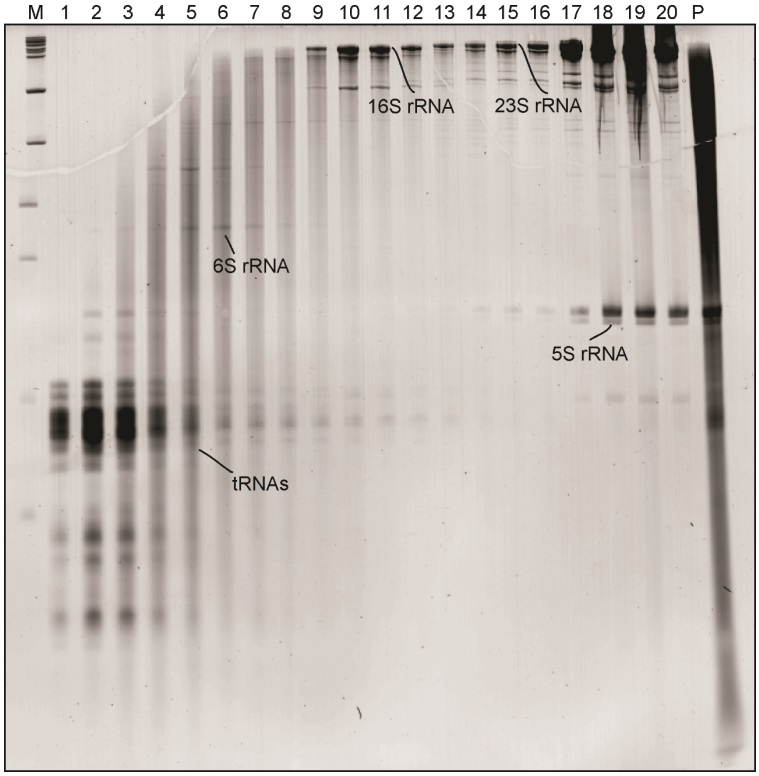

Figure 6.5 RNA quality control of a gradient profiling experiment.
After fractionation and RNA extraction, RNA samples were loaded (equal volume each fraction) on a 6%PAA gel and stained with ethidium bromide. M=marker, P=pellet, each number corresponds to each in-gradient fraction (1 to 20). Abundant RNA species visible on the gel are indicated with a black line.

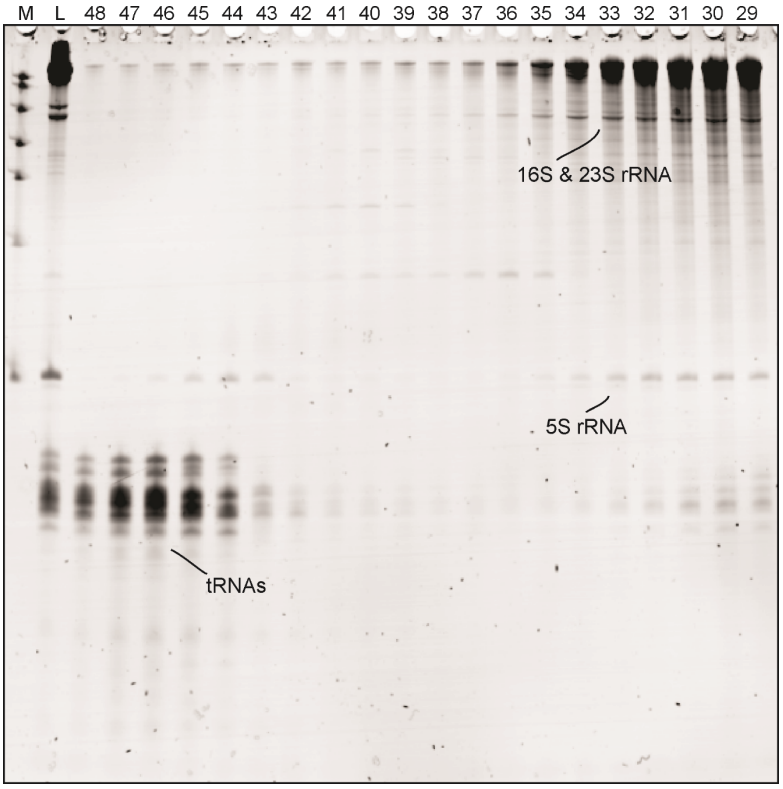

Figure 6.8 RNA quality control of a size-exclusion chromatography (SEC) experiment. After collection of eluted fractions and RNA extraction, samples were loaded (equal volume each fraction) on a 6%PAA gel and stained with ethidium bromide. M=marker, L=lysate, each number corresponds to each in-SEC fraction (48 to 29). Abundant RNA species visible on the gel are indicated with a black line.

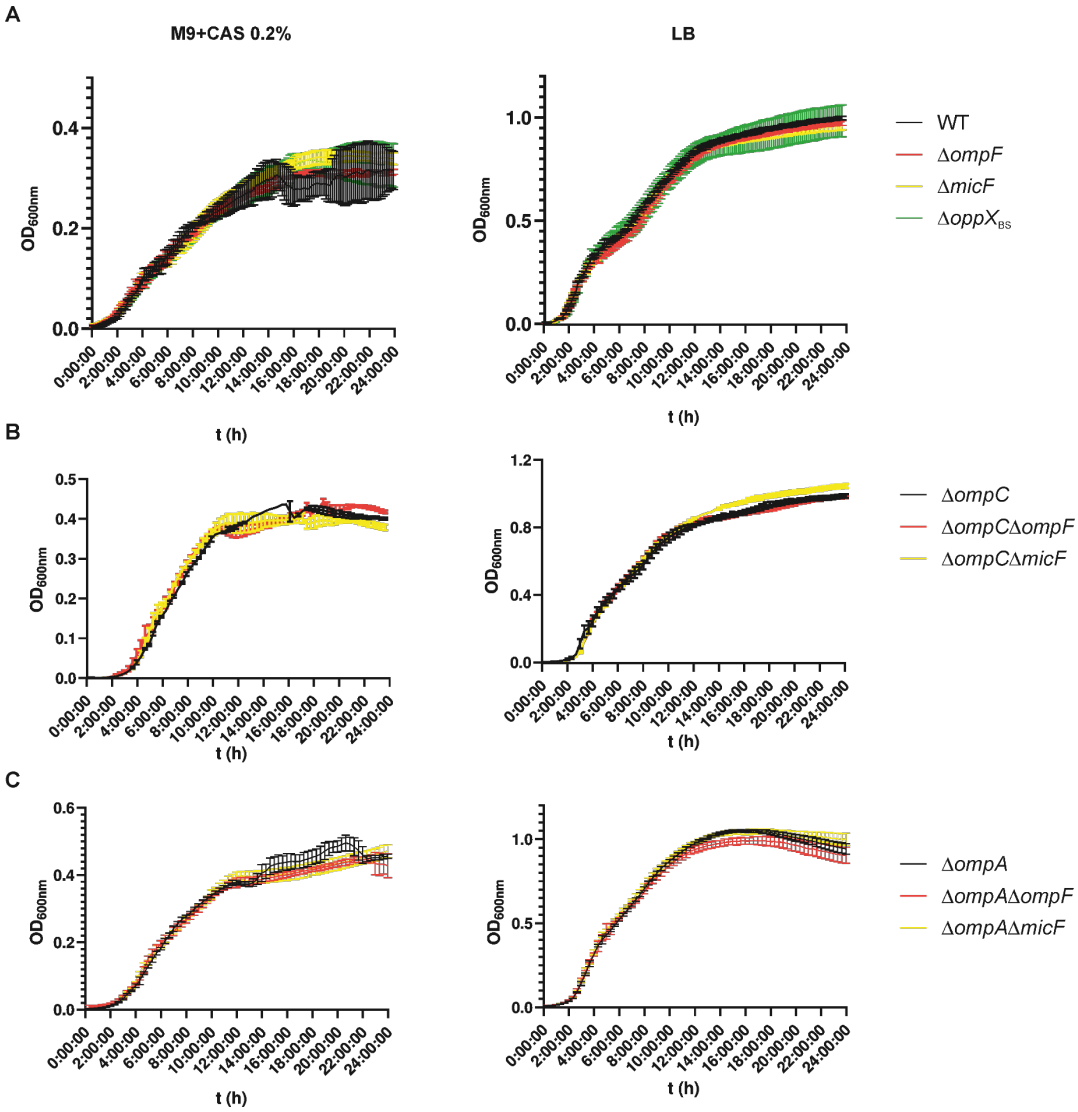

Figure 6.11 Growth curve nalysis of OMPs mutants in poor or rich media.
(**A**) OD600 measurement of 24h growth for WT, ompF, micF and oppXBS deletion mutants. Growth was tested in both M9 supplemented with 0.2% casamino acids as solely carbon source (left) or in LB (right). (**B**, **C**) Same experiments as described in (**A**) with mutants respectively indicated on the right side of the graphs. Mean and ±SD values of independent experiments are plotted (n=3).

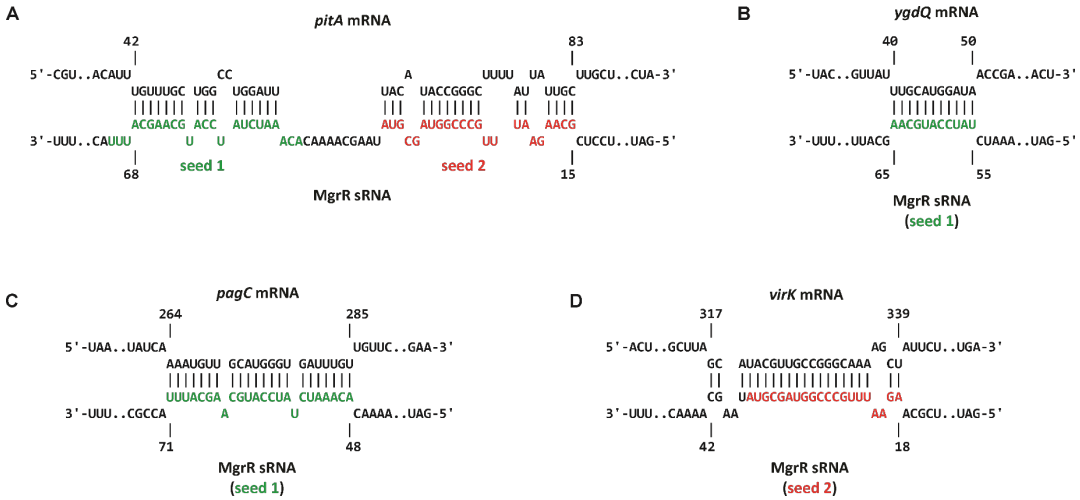

Figure 6.14 Predictions of base pair between MgrR and its RIL-seq targets.
IntaRNA prediction between the MgrR sRNA sequence and the pitA mRNA (**A**), ygdQ (**B**), pagC (**C**) and virK (**D**). "Seed 1" and "seed 2" are highlighted in green or bold, respectively and indicated below the predicted duplex.

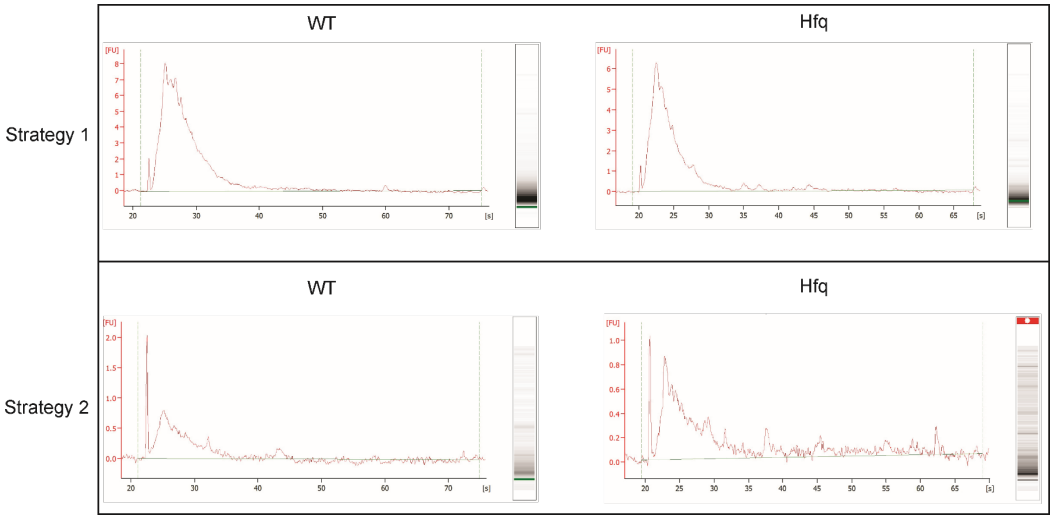

Figure 6.17 Bioanalyzer QC of the RNA recovered from in vivo RIL-seq.
RNA pico bioanalyzer run for the WT or the hfq::3×FLAG Salmonella strains of RNA recovered with Strategy 1 (upper panels) or with Strategy 2 (lower panels). The x-axis indicates time of elution from the bioanalyzer chip, proportional to RNA size; the y-axis indicates the intensity of the bands (FU=fluorescence units), proportional to RNA abundance.

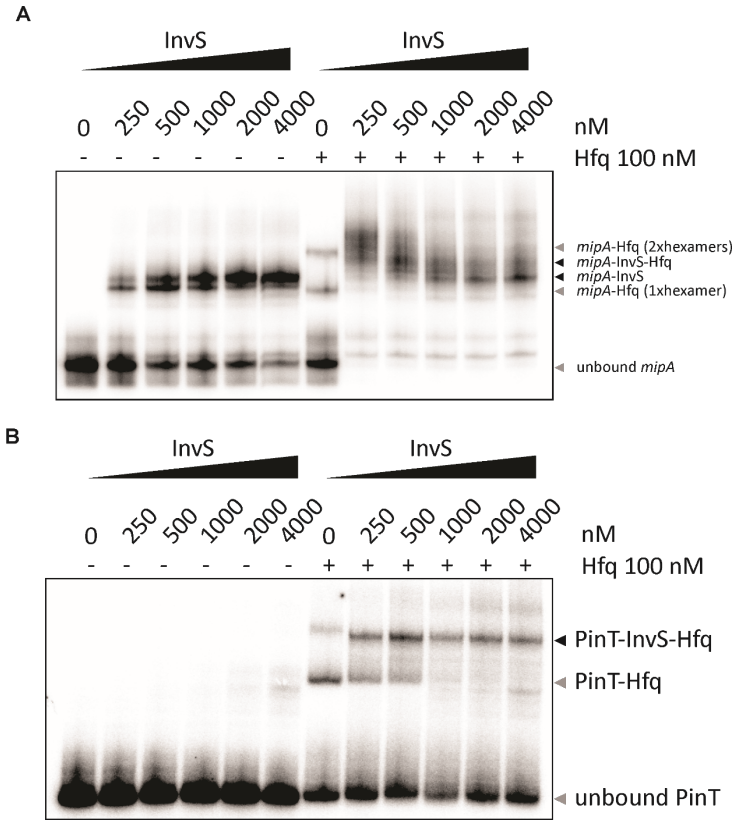

Figure 6.20 Band shift assays of InvS, PinT and mipA, with or without Hfq.
(**A**) EMSA gels of labeled mipA mRNA fragments carrying the InvS binding site, incubated with increasing concentrations (indicated above the gel) of cold InvS. Reactions were performed without (-, left gel) or with (+, right gel) Hfq at a concentration of 100 nM. Unbound mipA or mipA in complex with InvS and/or Hfq are indicated on the right side of the gels. (**B**) EMSA gels of labeled PinT, performed as described in (**A**).

6.2 List of abbreviations

Table 6.1 List of abbreviations.

Abbreviation	Meaning
aa	amino acid
APS	ammonium persulfate
bp	base pair
CAS	casamino acids
cDNA	complementary DNA
CDS	coding sequence
CLASH	UV-crosslinking, ligation and sequencing of hybrids
CLIP-seq	crosslinking immunoprecipitation and RNA-seq
ctrl	control
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide
dRNA-seq	differential RNA-seq
dsRNA	double-stranded RNA
DTT	dithiothreitol
EDTA	ethylene diamine tetraacetic acid
gDNA	genomic DNA
Grad-seq	gradient profiling by sequencing
LB	Lennox broth
M	Marker
MBP	maltose binding protein
m.o.i.	multiplicity of infection
mRNA	messenger RNA
MS	mass spectrometry
NB	northern blot
nt	nucleotide
OD	optical density
ORF	open reading frame
P	pellet
Pb^{2+}	lead acetate
P/C/I	phenol/chlorophorm/isoamyl alcohol
PAA	polyacrylamide
PAGE	polyacrylamide gel electrophoresis

Abbreviation	Meaning
PCR	polymerase chain reaction
PNPase	polynucleotide phosphorylase
RBP	RNA-binding protein
RBS	ribosome binding site
RIL-seq	RNA interaction by ligation and sequencing
RIP-seq	RNA immunoprecipitation and RNA-seq
RNA	ribonucleic acid
RNase	ribonuclease
RNA-seq	RNA sequencing
rRNA	ribosomal RNA
RT-qPCR	reverse transcription quantitative PCR
SD	Shine-Dalgarno
SD	standard deviation
SDS	sodium dodecyl sulfate
sRNA	small RNA
TCS	two-component system
tRNA	transfer RNA
TSS	transcriptional start site
UTR	untranslated region
w/v	weight/volume
WB	western blot
WT	wild type

6.4 Publications

6.4.1 Published manuscripts

Hör, J.*, **Matera, G.***, Vogel, J., Gottesman, S., and Storz, G. (2020). Trans-acting small RNAs and their effects on gene expression in Escherichia coli and Salmonella enterica. EcoSal Plus 9.

Miyakoshi, M., **Matera, G**., Maki, K., Sone, Y., and Vogel, J. (2019). Functional expansion of a TCA cycle operon mRNA by a 3′ end-derived small RNA. Nucleic Acids Research 47, 2075–2088.

Wang, C., Chao, Y., **Matera, G**., Gao, Q., and Vogel, J. (2019). The conserved 3′ UTR-derived small RNA NarS mediates mRNA crossregulation during nitrate respiration. Nucleic Acids Research 48, 2126–2143.

6.4.2 Accepted manuscripts

Matera, G., Altuvia, Y., Gerovac, M., El Mouali, Y., Margalit, H., Vogel, J. (2021). Global RNA interactome of *Salmonella* discovers a 5'UTR sponge for the MicF small RNA that connects membrane permeability to transport capacity. Molecular Cell, *accepted*.

6.4.3 Manuscripts in preparation

Matera, G., Venturini E., Vogel J. RIL-seq reveals changes in the interactome of Hfq during a *Salmonella* infection time course. I*n preparation.*

7 Acknowledgements

During the years of my Ph.D. many people played a major role in helping and supporting me in several ways, and for that I am very grateful. Specifically, I would like to thank

Prof. Dr. Jörg Vogel for welcoming me into his group as a Master student first, and as a Ph.D. student in a second place. I am grateful for the support you have given me during the years, for the exciting scientific environment you have exposed me to and for always pushing me further to achieve better results.

Dr. Franziska Faber and Prof. Dr. Kai Papenfort for taking part to my thesis committee and for providing me with the necessary support and suggestions during our meetings. I would also thank Prof. Dr. Christoph Sotriffer for agreeing to be the chairperson of my thesis defence.

Prof. Hanah Margalit, Dr. Yael Altuvia and other members of the Margalit lab for the precious scientific collaboration on the RIL-seq project and for their support during my (short) stay in Israel.

Dr. Gisela Storz and Dr. Susan Gottesman for allowing me to visit their labs and for the amazing collaboration on the *Salmonella*/*E. coli* sRNA review.

All the members of the Vogel lab (current and past) and members of other labs from IMIB/ZINF/HIRI which constantly helped me during the toughest times of my Ph.D.. Every single one of you has, in some way, taught me something precious that will stay with me forever.

All the people from the institute that I encountered during my Ph.D. time. With you it has been, albeit not the easiest, the greatest journey of my life. Particularly, special thanks go to: Elisa, Svetlana, Hilde, Jens, Falk, Youssef, Valentina, Manu, Marie, Philipp, Eli, Kotaro, Wael, Elise and many more. Without you I would have never pulled through the toughest moments.

The “Los Candidos” band members: Fabian, Sergio and Austin. It has been such a great fun sharing with you our love for music and I am so proud we have managed to entertain people with a few small gigs.

The GSLS for providing all the necessary tools to successfully complete my Ph.D.

Above everything, my family, for constantly believing in me and sharing their love and support albeit being 1,500 km away. It has never been easy from either side.

8 Affidavit/Eidesstattliche Erklärung

I hereby confirm that my thesis entitled "Global mapping of RNA-RNA interactions in *Salmonella* via RIL-seq" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Place, Date Signature

Hiermit erkläre ich an Eides statt, die Dissertation „Globale Analyse der RNA-RNA-Interaktionen in *Salmonella* mitells RIL-seq" eigenständig, d.h. insbesondere selbständig und ohne Hilfe eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen un Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

Ort, Datum Unterschrift